Universidad Rey Juan Carlos

Facultad de Ciencias Jurídicas y Sociales
Programa: Economía Aplicada
Departamento Coordinador: Economía Aplicada I

FINANCIAL ASSET PRICING UNDER KNIGHTIAN UNCERTAINTY

Tesis para la obtención del grado de Doctor

Director: Prof. Dr. Jesús Huerta de Soto
Codirector: Prof. Dr. Miguel Ángel Alonso Neira
Autor: David Howden

Madrid, España
Curso Académico 2008-2009

ACKNOWLEDGEMENTS

Several unique parties all contributed in forming the ideas for this thesis.

My own director, Jesús Huerta de Soto, through his support and encouragement is who I owe my greatest intellectual debt. The "función empresarial" provides the basis of my final asset pricing model, and an ignorance of the extensive literature would have disallowed the fruitful results it is hoped are yielded in this work.

My own group of friends and colleagues in Madrid has contributed through the great discussions that we have had. Particularly, Antonio Zanella and Philipp Bagus, each of whom has left a mark on many of the ideas within must be given due acknowledgment. Both contributed greatly to both the enjoyable time spent working on this project, the celebrations of the discoveries of its core findings, and provided a continual source of encouragement throughout its preparation.

One anonymous individual provided me with the ability to work on this project continually and without other obligations for the full year it required to write. Without her, I could not have finished this work as thoroughly or quickly as I did.

Giovanna Pittilino in a great feat, made the Spanish prologue understandable.

In particular, the Ludwig von Mises Institute, and the wonderful people who make it happen, provided much support. Not only are they directly responsible for allowing me much time and freedom as a summer fellow in both 2008 and 2009 to work on these concepts, the ideas propagated through its website have, since the year 2000, allowed me to expand my mind as an Austrian economist. I would not be in the place I am today without the hard work and dedication of these people, in particular (and in no order): Lew Rockwell, Doug French, Jeff Tucker, Mardi Rockwell, Kristy Holmes, Norma Willock, Pat Barnett, Barbara Pickard, Joe Salerno, Judy Thommesen, Kathy White, Tom Woods, and Mark Thornton.

Lastly, but never least, my own family has done more to shape my character and mind than any other group could. To my Dad who continually stressed critical thinking and education, and my Mom who ever encouraged an optimistic outlook and strove to make me give my best effort at every task, I am forever thankful. Above all, my brothers Michael and Jamie, and my sister Sarah, have helped shape every thought that develops in my own mind. This dissertation is dedicated to all of them.

BIOGRAPHY

David John Howden was born in London Ontario Canada on September 22nd, 1981. He was raised on a beef farm outside of Watford, Ontario, Canada, until the age of 18 when he left for University.

The valedictorian of his elementary school, Warwick Central School, he later graduated East Lambton Secondary School in 2000. He attended Wilfrid Laurier University, graduating with an honors degree in Business Administration with a specialization in Finance in 2004. A member of the winning team for the prestigious Integrated Case Exercise at Wilfrid Laurier University in 2003, he later went on to work for Mackenzie Financial Corporation in Toronto Canada, and Custom House Currency Exchange in Victoria Canada.

Having graduated from the Universidad Rey Juan Carlos in Madrid Spain with a Masters of Economics in 2008, he was accepted into the PhD program at the same University shortly after. This dissertation fulfills the requirements of the PhD program.

A summer research fellow at the Ludwig von Mises Institute in 2008 and 2009, he was the winner of the Douglas E. French award in Austrian economics during his first year. The same year, he was a co-winner of the Society for the Development of Austrian Economics' annual award for best doctoral student paper.

He is, to date, the author or coauthor of eight published and refereed academic papers:

1. with Philipp Bagus. (forthcoming). "*Qualitative* Easing in Support of a Tumbling Financial System: A Look at the Eurosystem's Recent Balance Sheet Policies." *Economic Affairs*.
2. (forthcoming). "Fama's Efficient Market Hypothesis and Mises' Evenly Rotating Economy: Comparative Constructs." *The Quarterly Journal of Austrian Economics*.
3. with Philipp Bagus. (forthcoming). "The Legitimacy of Loan Maturity Mismatching: A Risky, But Not Fraudulent, Undertaking." *The Journal of Business Ethics*.
4. with Philipp Bagus. (forthcoming). "The Federal Reserve System and Eurosystem's Balance Sheet Policies During the Financial Crisis: A Comparative Analysis." *Romanian Economic and Business Review.*
5. (forthcoming). "Knowledge Shifts and the Business Cycle: When Boom Turns to Bust." *Review of Austrian Economics*.
6. (forthcoming) "Single Trial Probability Applications: Can Subjectivity Evade Frequency Limitations?" *Libertarian Papers*.
7. (2009). "Money as a Medium of Exchange and Its Evolution: An Elaboration on Mengerian Monetary Economics." *Procesos de Mercado: Revista Europea de Economía Política*, 6(1): 199-210.
8. (2008). "Stability of Gold Standard and Its Selected Consequences: A Comment." *Procesos de Mercado: Revista Europea de Economía Política*, 5(1): 159-175.

Table of Contents

VOLUME I

The

EFFICIENT MARKET HYPOTHESIS
and the
CAPITAL ASSET PRICING MODEL

New Perspectives on Old Theories

PART I
THE EFFICIENT MARKET HYPOTHESIS AND THE CAPITAL ASSET PRICING MODEL

PART II:
THEORETICAL PROBLEMS WITH THE MODELS

VOLUME II

Production and Value: Microfoundations of Finance

BOOK I
ACTION AND ITS WORLD

PART II
ACTION AND ITS STRUCTURES

PART III
ACTION AND ITS VALUE

PRÓLOGO

Este trabajo ha sido escrito originalmente para rectificar los grandes problemas que se existen en el mundo financiero. De hecho, la crisis hoy en día ha confirmado concretamente que los problemas y errores con el pensamiento corriente han hecho sus efectos conocidos en casi todas las esquinas de nuestro mundo. Con miles de billones de euros de perdidos ya, un alternativo a los ideas y teorías que se existen y tienen mucho apoyo están buscado por mucha gente – ambos los practicantes y académicos de las finanzas. Con una gama muy ancha, los problemas se necesitan una solución fundamental para explicarse la fuente verdad de los problemas que se han resultado. Como los errores han sido, en esto caso, tan egregio, la necesidad por un alternativo es evidente.

Cuando escribiendo este trabajo, se emerge que una solución no será completa si el enfoque fue solo una crítica de los problemas y sus fuentes. Aunque la identificación de los problemas fundamentales está un buen punto para empezar en la carrera que va a las soluciones, al mismo tiempo será solo una empezar. Para identificar las fuentes de los problemas y no ofrecer un alternativo dejaré una deficiencia imperdonable en el análisis. Por esta razón, el trabajo presente tiene dos volúmenes. El primer volumen es la critica que necesitamos antes que comenzamos nuestra nueva teoría sobre como a determinar los precios de bienes financieros. Ciertamente, el cuerpo de literatura sobre la determinación de los precios de bienes financieras se basado principalmente en dos conceptos: la hipótesis de la eficiencia de los mercados (HEM) y el capital asset pricing model (CAPM), se haciendo esos los dos el enfoque de la crítica.

Más o menos, la HEM se dice que los mercados son eficientes si reflectó plenamente y correctamente toda la información actual y relevante concerniendo un precio de una acción. Por lo tanto, por un conjunto de información, precios no afectaron de revelando cada información desde eso conjunto a los participantes del mercado. Es más, la eficiencia sobre esa conjunto de información se implica que será imposible que hacer beneficios basado en la información se contiene. En su forma más débil, la HEM se afirma que los precios reflectaron plenamente la información contendido en la secuencia histórica de precios. Como resulta, inversores no se pueden hacer más rentas desde analizando la pauta pasado de los precios de una acción. En la otra mano, la forma de HEM mas fuerte se dice que toda la información conocido a cualquier participante del mercado ya se reflejado en su precio. Por lo tanto, ningún de los participantes se pueden beneficiar de aun la información previlejada como ya lo han incorporado a la constelación de los precios.

Los economistas han reconocido que Samuelson (1965) and Mandelbrot (1966) han probado que si los flujos de información son sin fricciones y sin costas, los precios de mañana debe ser cambiado solo desde información nueva. Como la información no puede ser conocido adelante, los precios se deben seguir un "camino aleatorio" (random walk). Esta resurgimiento del camino aleatorio que se origina con el matemático francés Louis Bachelier (1900) se marcado el fallo mas egregio del la literatura de HEM. Según a Burton Malkiel (1992: 741), "[p]robablemente no otra hipótesis en cualquier de los dos campos, la economía o las finanzas, ha sido usado mas extensivamente." Las implicaciones malas resultando directamente desde HEM se quedo clara.

Desagraciadamente, mientras el uso de un camino aleatorio basado en alguna forma de eficiencia del mercado ha probado fácil analíticamente, esos procesos estocásticos se fallan a decir la historia verdad sobre la determinación de los precios se llegue día a día en el mercado. Siete fallos específicos se hacen la crítica en la primera volumen.

Concretamente, una vista lineal de tiempo se hace cambiados de valores endógenos imposible solo como resulta del pasaje del tiempo (cambiados en la preferencia del tiempo, por ejemplo). Segundo, la incertidumbre Knightiano es imposible dentro la HEM, porque solo deviaciones estadísticas desde la estructura presente de las expectaciones es posible. Tercero, el abandono de la función empresarial no se puede explicar como la información se propagar tras el mercado. Cuarto, la HEM ha sido siempre un resultado buscando por una teoría – la propuesta metodología ha sido 180 grados opuestos y se deje la teoría en una fundación muy débil. Quinta, la bifurcación entre la información y conocimiento ha ocurrido, que se hace la interpretación subjetiva de la información imposible (y, por lo tanto, no se permite por las oportunidades de rentabilidad no descubrió antes). Sexta, la HEM se asume la eficiencia estática que enfoque en los precios presentes, pero abandone las expectativas sobre la futura. Finalmente, la HEM se asume que precios son independientes y aleatorios, y como resulta, se permite participantes quien se crean actuar en una manera aleatoria similar.

El capital asset pricing model es un término que se refiera a la colectiva de modelas que se proba a maximizar las rentas de un portfolio por una cantidad de varianza dado, como sugerido de Markowitz (1952; 1958). CAPM es un modelo equilibrio que se relata los precios de acciones a los datos exógenos, dado las preferencias de inversores. El rasgo básico es que las preferencias de los inversores son definidos de la media y varianza de la renta expectativa de un portfolio, y que las rentas juntas de los bienes individuos de un portfolio se determinan el riesgo que se resultara. La teoría de separación de Tobin (1958) se implica que los individuos elegirán entre un bien con riesgo, y uno sin riesgo (inicialmente dinero, pero usualmente una bono sin riesgo, i.e., de corto tiempo del gobierno), para llegar a un bien sintetizado y único. La implicación es que se existe una frontera donde todos los inversores se comparten que se represente las combinaciones eficientes de bienes sin y con riesgo que maximizar las rentas por un nivel de riesgo dado.

La suposición de creencias de inversores homogéneos que Tobin (1958) se necesita para llegar a esa conclusión fue relajado y olvidado después Sharpe (1964) y Litner (1965) se desarrollan formalmente que hoy en día es la fundación del capital asset pricing model.

El CAPM ha llegado como un rasgo importante en muchos de los modelos modernos del mundo de las finanzas. Como Rubinstein (2002: 1044) nos informe, "las ideas en el articulo [de Markowitz] han llegado tan central a la economía financiera que no pueden ser removidos." Dado que el CAPM es tan aceptado en las ideas finanzas del pensamiento corriente, siete criticas son hechos para refutar la corazón del modelo.

En primer lugar, el tiempo es valorado linealmente, y, por lo tanto, se abandona los cambios endógenos en los valores. Segundo, el incertidumbre es tratado como el riesgo, y perfiles de riesgo-renta del futuro están asumidos a ser conocidos probabilísticamente en avanza. Tercero, los individuos están tratados como "price takers", removiendo la función empresarial. Cuarto, la metodología inductiva se hace resultados no concluyentes que se necesitan continuos probando para mantener su validez. Quinta, el CAPM se asume que la elección se arbitrar entre dos variables – riesgo y renta – mientras abandonando los otros variables importantes a las preferencias individuales. Sexta, el uso de un bono libre de riesgo se asume una substitución buena para dinero, mientras abandonando el papel de dinero en la mitigación de incertidumbre. Finalmente, el riesgo como la deviación estándar de los precios es erróneo, por que los inversores son solo preocupados con el precio cuando se actualiza la venta de su acción, y se realiza un beneficio o una pérdida.

La crisis del presente ha dado a muchos inversores perdidas doloridas como un resultado directo de siguiendo esas dos teorías. De hecho, fueron importantes, si encubierto, rasgos del boom reciente.

La influencia de HEM y la idea que los precios se según un camino aleatorio se da inversores un sentido de seguridad falsa sobre los precios. Implicando que los precios están determinados correctamente a cualquier momento de tiempo, inversores fueron enfrentados con la realidad que vendiendo una acción sería una mala decisión de inversión. Se quedan claro que los fondos de índices tenido una prevalencia en el mercado que se permitirán los inversores a participan en el boom con la paz del mente que los precios fueron correctos, considerando toda la información disponible. El crash reciente, como muchos otros antes, ha revelado las conclusiones erróneas que la HEM ha hecho.

Igualmente, la conclusión de CAPM que un inversor puede hacer mas rentabilidad solo de adquiriendo mas riesgo fue abusado desde inversores cuando se pesado sus portfolios con bienes de más riesgo. Buscando por rentas altas, esos inversores se pueden dormir fácilmente conociendo que los precios de bienes fueron correctos a cualquier momento de tiempo. Diversificación se vista como un método para asegurar rentas continuamente altas, sin teniendo a sufrir cualquier perdidas severas en el medio a largo plazo. La suposición que el perfil de riesgo-renta será constante, conjunto con la suposición que clases de bienes tienen covarianza que también son constante o probabilísticamente conocido en el futuro, han resultado en portfolios diversificados que se funciona bien durante condiciones específicos, pero cuando nuevo situaciones surgido, se cree un crisis detrimento. Las pérdidas que se ocurren podría ser evitados sin el uso de eso modelo incorrecto.

En el segundo volumen de eso trabajo, construyo una teoría para determinar los precios de bienes financieras basado en una fundación *a priori*. Esa fundación ha contestada los problemas identificados en el primer volumen, con la esperanza de produciendo un estructura más coherente para determinar los valores de los bienes financieras basado en micro-fundaciones.

Empezando con la función empresarial, demuestro que esa función combina tres funciones: la mitigación de riesgo, previendo incertidumbre, y la provisión de recursos. Un individual con todas las tres funciones a un máximo nivel será el que refriéremos como la "función empresarial pura" (nota que si las dos primeros funciones son realizadas plenamente, que la tercera función será eliminada porque no hay pérdidas económicas). No único individual puede tener esas tres funciones perfectamente. Como resulta, la función empresarial pura no puede existir naturalmente, y entonces, individuales se necesitan hacer una síntesis para obtenerla. Esos intentos están que comúnmente refriéremos como "empresas." Empresas se existe para combinar individuales como están requeridos para producir bienes queridos en el futuro. A la alcance que la función empresarial pura se hecho con éxito, rentas dentro de una empresa serian maximizadas.

Si eso es como la empresa se hace rentas, ¿cómo determinamos el valor que un individual da a la porción de esas rentas tras las compras de acciones?

Tenemos tomaba el común modelo de dividendo descontado (MDD) de preciando acciones y lo cambiado para mover nuestras tres funciones empresariales nuevas al frente. El MDD se asume que el valor presente de una acción es el valor presente de todos los dividendos del futuro esperado a ser pagado. Al mismo tiempo, se asume cualquiera de dos cosas: a) la acción se comprada para siempre, o b) al tiempo que lo vende, el flujo de dividendos se desploma al precio futuro de la acción y será equivocado en valor al flujo que obtendré si un inversor ha comprado la acción para siempre. Se afectado desde la proporción de dividendos pagaba, y la tasa de crecimiento de la acción (y segundamente desde la costa de oportunidad de inversión, normalmente dado come el tasa de interés sin riesgo).

La función empresarial que se mitiga el riesgo se determina el nivel de dividendo pagado en el presente. Se trabajan a maximizar rentabilidad dentro una sistema cerrada (pero cambiando) que es determinado del pronósticos de los individuales con la función empresarial de previendo incertidumbre. Esa segunda clase de la función empresarial se determina la tasa de crecimiento de la empresa, y por lo tanto, el valor de sus acciones y también se proporcionen el nivel de rentabilidad que la primera función empresarial puede hacer en el presente. Finalmente, la cantidad de recursos ofertado desde la tercera función empresarial se determinare que será la costa de oportunidad de una acción. Con mas inversores invistiendo dinero, la tasta de interés en préstamos se disminuya y viceversa.

El componente final que se contribuye de los precios es el factor de tiempo con respecto a nuestras expectaciones. Inversores no tienen expectativas constantes, ni se planifican a comparar una acción para siempre. Vemos que las compran para un periodo de tiempo finito dictado tras su preferencia de tiempo, y también de la expectativa de máximo rentabilidad de la acción. Así, el valor futuro de la acción al momento que espera a venderlo sería distinto del valor de los dividendos expectativos que lo hace (de hecho, tendré por la definición de un venta porque mercados se requiere participantes con preferencias y valores reversadas). El componente de tiempo será una determinante mejor en la determinación de los valores de bienes en el presente.

Eso trabajo puramente teorético ha hecho una fundación para mas trabajos en el futuro a desarrollar esa rama de la economía austriaca en la dirección de las finanzas. Las pies de las paginas ha comentado por maneras pruebas empíricas se lo puede ser hecho. Por ejemplo, la estructura para haciendo las decisiones bajo condiciones de incertidumbre desarrollado en el segundo volumen se permite una aproximación para el nivel de incertidumbre percibido desde mediando la cantidad de ahorras de dinero.

Se espera que el trabajo presente serviré dos funciones. En un lado, a causar descreditó de los modelos del pensamiento corriendo para determinando los precios de valores, imperfecto como están. En el otro lado, si se puede hacer una chispa de revolución en las teorías austriacas de finanza e inversión, el segundo volumen será un éxito. Aunque es mi propio gol que la segunda se obtendré, esa tarea probablemente se necesitare a esperar hasta el primer ya ha comenzado. Espero que esas ideas y conceptos del segundo volumen no se puede esperar hasta tal un tiempo para ganar aceptación.

Referencias

Bachelier, Louis. 1900. Théorie de la Spéculation*. Annales Scientifiques de l'École Normale Supérieure*: 21-86.

Lintner, John. 1965. The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics* 47(1): 13-37.

Malkiel, Burton G. 1992. Efficient Market Hypothesis. *New Palgrave Dictionary of Money & Finance*, (*eds*.) Peter Newman, Murray Milgate and John Eatwell. London: The Macmillan Press Ltd.

Mandelbrot, Benoit. 1966. Forecasts of Future Prices, Unbiased Markets, and Martingale Models. *Securities Prices: A Supplement, Journal of Business* 39(1): 242-255.

Markowitz, Harry. 1952. Portfolio Selection. *The Journal of Finance* 7(1): 77-91.

Markowitz, Harry. 1958. *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley & Sons, Inc.

Rubinstein, Mark. 2002. "Portfolio Selection": A Fifty-Year Retrospective. *The Journal of Finance* 57(3): 1041-1045.

Samuelson, Paul A. 1965. Proof That Properly Anticipated Prices Fluctuate Randomly. *Industrial Management Review* 6(2): 41-49.

Sharpe, William F. 1964. Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk. *Journal of Finance* 19(3): 425-442.

Tobin, J. 1958. Liquidity Preference as Behavior Towards Risk. *The Review of Economic Studies* 25(2): 65-86.

PROLOGUE

This thesis was originally written to rectify the large issues that exist in the financial world. In fact, the present crisis has confirmed that the problems and errors in mainstream finance theory have made their effects known in almost all the corners of our world. With trillions of dollars lost through recent events, an alternative to these theories has much support among many individuals – investors and economists alike. With wide ranging implications, these problems require a fundamental solution to explain the true source of the financial crisis that has resulted. As the errors have been, in this specific case, so egregious, the necessity for an alternative is evident.

When writing this work, it emerged that no solution would be complete if the focus was solely on a critique of these old problems and their sources. Even though the identification of the fundamental problems is the correct place to start the road to analyzing these problems, it is, after all, only a start. To identify the sources of these problems and not offer an alternative would leave an unpardonable deficiency in the thesis. For this reason, the present work proceeds in two volumes.

The first volume is the critique necessary before we start our new theory concerning pricing financial assets. Indeed, the current body of asset pricing literature is based primarily upon two concepts – the efficient market hypothesis (EMH) and the capital asset pricing model (CAPM), making these the focal point of the critique.

Loosely stated, the EMH states that markets are efficient if they fully *and* correctly reflect all current and relevant information concerning a security's price. Hence, for a given set of information, prices would be unaffected by revealing any information from this set to participants. Furthermore, efficiency regarding this information set implies that it is impossible to earn profits trading on the basis of information contained in it. In its weakest form, EMH claims that prices fully reflect the information contained in the historical sequence of prices. Thus, investors cannot profit by analyzing the past pattern of securities' prices. At the other end of the spectrum, the strong form of EMH asserts that all information known to any market participant is fully reflected in its price. Hence, no individual can benefit from even the most privileged of information as it has already been incorporated into the constellation of prices.

Samuelson (1965) and Mandelbrot (1966) have been recognized by economists to have proven that if information flows without impediment and this is costless, tomorrow's prices may be altered only by new information. As information cannot be known in advance, prices must follow a "random walk."This revival of the random walk model of stock prices which dates back to the French mathematician Louis Bachelier (1900) marks the most egregious failing of the EMH literature.

According to Burton Malkiel (1992: 741), "[p]robably no other hypothesis in either economics or finance has been more extensively used." The wide reaching implications resulting from EMH become evident.

Unfortunately, while the use of a random walk model based on some form of market efficiency has proven analytically easy, these stochastic processes fail to tell the true story of price determination arising day-to-day on the market. Seven specific failings are critiqued in volume I.

In particular, a linear view of time rules out valuational changes endogenous to the passing of time (changes in time preference, for example). Second, Knightian uncertainty is ruled out of EMH, allowing for only statistical deviations from the present structure of expectations. Third, the neglect of the entrepreneurial function cannot explain how information is spread through the market. Fourth, EMH has always been a result searching for a theory – the methodological approach has been 180 degrees backwards and left the theory on a weak foundation. Fifth, a bifurcation between information and knowledge has occurred, which rules out the subjective interpretation of information (and hence, allow for undiscovered profit opportunities). Sixth, EMH assumes a static efficiency which focuses on present prices, but neglects forward looking expectations. Finally, EMH assumes prices are independent and random, and hence, allows for the individuals who create them to act in a similarly random manner.

The capital asset pricing model is a term used to refer to the collective of models which seek to maximize returns in a portfolio for a given amount of variance as proposed by Markowitz (1952; 1958). CAPM is an equilibrium model which relates asset prices to exogenous data, given the preferences of investors. The basic feature is that these investor preferences are defined by the mean and variance of a portfolio return, and that the joint returns of the portfolio's composite assets determine the resultant risk. Tobin's (1958) separation theorem implied that individuals choose between a risky asset and a riskless asset (originally cash, and later a risk-free bond) to arrive at a unique composite asset. The implication is that there exists a frontier which all investors will share which represents the efficient combinations of risky and riskless assets that maximize returns for a given level of risk.

The assumption of homogeneous investor beliefs necessary for Tobin to arrive at this conclusion was relaxed and forgotten after Sharpe (1964) and Litner (1965) formally developed what today is the basis of the capital asset pricing model.

The CAPM has become a salient feature in most modern finance models. As Rubinstein (2002: 1044) reports, "the ideas in [Markowitz's] paper have become so interwoven into financial economics that they can no longer be disentangled." Given that the CAPM is such an accepted part of mainstream finance, a seven point critique has been provided to refute the core of the model.

First, time is viewed and valued linearly, and hence, neglects endogenous changes to its value. Second, uncertainty is treated as risk, and future risk-return profiles are assumed to be probabilistically known in advance. Third, individuals are treated as price-takers, removing the entrepreneurial element. Fourth, the methodological approach has led to inconclusive results requiring continual testing for validity. Fifth, CAPM assumes choice is arbitraged between two variables – risk and return – while neglecting other variables important to individuals' preferences. Sixth, the use of a risk-free bond is assumed a substitute for cash, while neglecting money's role in uncertainty mitigation. Finally, risk as the standard deviation of return is erroneous, as investors are only concerned with the selling price when they actually sell an asset and realize a profit or loss.

The current crisis has given many investors painful loses as a direct result of following these two theories. Indeed, they were salient, if mostly latent, features of the recent boom.

The influence of EMH and the idea that prices follow a random walk gave investors a false sense of security concerning prices. By implying that prices are correctly priced at any given time, investors were faced with the reality that selling a stock, no matter what its current value was, would be a poor investment decision. Index funds became prevalent in the market allowing investors to costlessly partake in the investment boom with the peace of mind that prices were correctly considering all available information. The recent crash, like many crashes before, has revealed the erroneous conclusions that EMH has provided.

Likewise, the conclusion of CAPM that an investor can earn greater return only by taking on more risk was abused by investors as they weighted their portfolios with riskier assets. Searching for higher profits, these investors could sleep easy knowing that asset prices were largely correct at any given time. Indeed, diversification was seen as a way to ensure that profits could remain continually high, without having to suffer any severe losses in the short-run. The assumption of a constant risk-return profile, combined with the assumption that asset classes have covariances that are also constant or probabilistically known over time, resulted in diversified portfolios that functioned well under historical conditions, but as new situations arose, led to a detrimental breakdown. The losses that ensued could have been wholly avoided without the use of this erroneous model.

In the second volume of this work, I have built up a theory of pricing financial assets based on *a priori* foundations. This has addressed the problems raised in the previous volume, in hopes of producing a more coherent framework for pricing financial assets based on micro-foundations.

Starting with the entrepreneur, I have shown that the entrepreneurial function involves three functions: mitigation of risk, bearing of uncertainty and the provision of resources. An individual endowed with all three to the maximum extent would be what we refer to as a pure entrepreneur (note that if the first two functions are fulfilled, the third function is eliminated due to the lack of economic losses). No one individual can have these functions perfectly endowed in them. As a result, pure entrepreneurs who cannot exist naturally are synthesized by individuals. These attempts are what we commonly refer to as "firms." Firms exist to combine individuals as they are required to produce goods desired by individuals in the future. To the extent that a pure entrepreneur is successfully synthesized, profits within a firm will be maximized.

If this is how a firm derives its profits, what determines the value an individual places on a portion of those profits through security purchases?

We have taken the dividend discount model (DDM) of pricing securities and altered it to place these entrepreneurial functions to the fore. The DDM assumes that the present value of a stock is the present value of all dividend payments it is expected to pay out. At the same time, it assumes that either: a) a security is purchased to be held forever, or b) that at the time when the security is sold, the stream of future dividends will collapse into the future price and will be equivalent in value to that which would obtain if an investor held the stock forever. It is affected by the dividend payout ratio, and the growth rate of the stock (and secondarily by the prevailing opportunity cost of investment, commonly cited as the risk-free rate).

Risk mitigating entrepreneurs are those who determine the present dividend payment. They work by maximizing profits within a closed (but changing) system that uncertainty bearing entrepreneurs determine through their forecasts. As such, this second entrepreneurial class determines the growth rate of the firm, and hence, the value of its stock as well as providing a maximum profit that the risk-mitigating entrepreneurs can earn in the present. Lastly, the amount of resource provided by entrepreneurs will determine what the opportunity cost of investing in a security is. As more entrepreneurs invest resources, the rate of return on loans decreases and vice versa.

One final component contributing to the determination of these prices is the time factor of expectations. Investors do not hold constant expectations, nor do they plan on holding a stock forever.

Instead, we see that they hold stocks for a finite amount of time dictated by their time preference, and also the expected maximization of stock value. Thus, the future value of the stock at the date they expect to sell it will differ from the value of the future expected dividends that would accrue to it (in fact, it would have to by the definition of a sale as markets require participants with reverse preferences). The time component becomes a major determinate in determining present asset values.

This purely theoretical work has laid a foundation for future work to develop this branch of Austrian economics in the direction of finance. Footnotes have pointed to the way in which empirical testing of these theories may be undertaken. The framework for decision-making under uncertainty developed in volume II allows for a proxy for perceived uncertainty by measuring current cash holdings, for example.

It is hoped that the present work will serve a two-fold purpose. On the one hand, if it can cause disrepute of the previously popular, but fundamentally flawed, asset pricing models, volume I will have been a success. On the other hand, if it can spark a revolution in Austrian finance and investment theory, volume II will have been a success. Although it is my own personal overriding goal that the second occurs, this task will likely wait until the first is well underway. I hope that the ideas in the second volume will not have to wait until such a time to gain acceptance.

References

Bachelier, Louis. 1900. Théorie de la Spéculation. *Annales Scientifiques de l'École Normale Supérieure*: 21-86.

Lintner, John. 1965. The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics* 47(1): 13-37.

Malkiel, Burton G. 1992. Efficient Market Hypothesis. *New Palgrave Dictionary of Money & Finance*, (eds.) Peter Newman, Murray Milgate and John Eatwell. London: The Macmillan Press Ltd.

Mandelbrot, Benoit. 1966. Forecasts of Future Prices, Unbiased Markets, and Martingale Models. *Securities Prices: A Supplement, Journal of Business* 39(1): 242-255.

Markowitz, Harry. 1952. Portfolio Selection. *The Journal of Finance* 7(1): 77-91.

Markowitz, Harry. 1958. *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley & Sons, Inc.

Rubinstein, Mark. 2002. "Portfolio Selection": A Fifty-Year Retrospective. *The Journal of Finance* 57(3): 1041-1045.

Samuelson, Paul A. 1965. Proof That Properly Anticipated Prices Fluctuate Randomly. *Industrial*

Management Review 6(2): 41-49.

Sharpe, William F. 1964. Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk. *Journal of Finance* 19(3): 425-442.

Tobin, J. 1958. Liquidity Preference as Behavior Towards Risk. *The Review of Economic Studies* 25(2): 65-86.

VOLUME I

The

EFFICIENT MARKET HYPOTHESIS

and the

CAPITAL ASSET PRICING MODEL

New Perspectives on Old Theories

All things are subject to the law of cause and effect.
This great principle knows no exception.

- Carl Menger (1840 - 1921)

BOOK I

THE
EFFICIENT MARKET HYPOTHESIS
AND
THE CAPITAL ASSET PRICING MODEL

I. INTRODUCTION

Economics is the youngest of all sciences. So opened Mises in his magnum opus *Human Action*. He would have this work published in 1949, slightly before the formal advent of the science of finance. Whether one thinks finance belongs in its own realm, or should be properly known forever as a subset of economics – financial economics – is not relevant to the present work.[1] However, while over the 2000 year history of formal economic theorizing there would be many ingenious and enlightening ideas discovered and lost, we conclude this is fundamentally less so in the financial realm.

Formally commencing in the 1950s, two early ideas would take the financial realm by storm and become the bedrock of most future developments in the field. In fact, the concepts of the *efficient market hypothesis* and the *capital market pricing model* are now so ingrained as to be totally inseparable from modern finance thought. However, we also see that relatively little theoretical attention has been given to either of these ideas. This issue stems from one specific area where the financial realm suffers a distinct disadvantage *vis á vis* the realm of economics proper: methodology. The consequences are both evident, and manifold.

Economics has enjoyed a long, storied debate as to what the proper methodology to be used in the field should be. As a result, almost every method imaginable has been attempted, and debated, with the consequence that many different competing ideas have existed concurrently. Finance, on the other hand, has existed solely in a methodological form solely dominated by empirical observations, mathematical models, and inductive rationales. Indeed, as Ross (1987: 29) views the critical difference between the two fields:

> Finance uses the modeling framework constructed in economics, but, within this scaffolding, finance has taken a different methodological perspective. It is wrong to characterize finance, or financial economics to be formal, as simply another of the specialty areas of economics... While finance is specialized in its focus on the financial markets, the differences between economics and finance only begin there. The principle distinction is one of methodology rather than focus.

---

1 Markowitz (1991: 476) views himself as being, at least partly, responsible for bringing finance under the umbrella of economics, “at the time I defended my dissertation, portfolio theory was not part of Economics. But now it is.” Although formal theorizing may have been distinct at one point, finance was always, and will always be, intimately dependent on the underlying realm of economics it describes.

This unified methodological perspective has left little time, or opportunity, for theoretical ideas to move to the forefront and challenge the accepted ideas. However, more and more, the fundamental building blocks of finance are being criticized *from an empirical position*. A giant disconnect has formed between the academic world marked by these models and ideas, and the real financial world that these academics owe their existence to. As Coase once remarked, “if you torture the data long enough, it will confess to anything.” These empirical critiques have been scathing at times, however, they have lacked the theoretical punch to put these ideas into the category of history – they both still enjoy a reputation as being necessary for further work. Indeed, both models are still taught heavily in universities, *even by professors who don't believe in their validity*. One of the most significant issues that creates this issue is that the theories themselves are questioned, *but not their individual ideas*. We find that a house is only as strong as its foundation, and that it is time to reassess the validity and theoretical correctness of both these theories: the efficient market hypothesis and the capital asset pricing model (henceforth referred to as EMH and CAPM).

Part one of this work will provide a brief history of these two concepts. The line of development that went into their existence must be reviewed to discover what basis their theoretical concepts find themselves. Both ideas enjoy colorful pasts, but are found lacking in strong theoretical rationalization. A brief overview of the influence that these two concept have had over the past 50 years of finance theory. The emphasis that has been placed on both EMH and CAPM becomes evident, and the need for a review of their theoretical correctness thenceforth becomes apparent.

The second part looks specifically at the theoretical issues found lacking with these theories. Four specific concerns consistently shared with each model will be assessed. Starting with the concept of time, we can see glaring issues as to its treatment. As time is the one element that unites all actors together, this creates the first, in a long line, of problems for later developments. In fact, as a result of this incorrect view of time, we see that two related concepts are misinterpreted. The first is the difference between risk and uncertainty. Both theories have used the idea of risk to be applicable to the world of human action. However, we see that the future is unknowable in many regards; a Knightian fog engulfs us as actors. As a result, our futures are not defined by *unknown* events that are *knowable* in advance, but instead have an *unknowable* component that is *unknowable* in advance. The repercussions for both are grand. The misconception of time excludes the concept of the entrepreneur from entering either picture. As both ideas are central to asset pricing theory, we see this as a huge omission. The entrepreneur is the force that drives the economy forward, into the unknown future. Hence, any price movements are a direct result of their actions. We must understand their role fully to grasp the process that creates and moves asset prices. A brief overview of method is given with a criticism of the empirical and mathematical methods used, almost exclusively, in finance. The alternative of the praxeologic approach is given, with an overview of our limits as physical scientists.

Part II progresses with specific issues concerning the efficient market hypothesis. In particular, three theoretical points are disputed. First is the conception of information in the economy. The concept of information as being costlessly and instantly dispersed to all individuals in the market is found wanting. Efficiency is also found to be a misguided concept in EMH. Any concept of efficiency must incorporate a concept of time intimately into it. This initial failure in EMH has led to erroneous conclusions as to the role of efficiency in the market. Finally, the specific determinants of price are examined. EMH views future price movements as only able to occur as a result of fresh information. However, we find the explanation of price movements not so simple. Factors exist which alter prices even when lacking the existence of fresh market data. As these are all central to the EMH, we find the issues inherent in them to be of utmost importance.

CAPM receives the same treatment, with the focus on three specific points. The first is the idea of choice homogeneity. Since Markowitz (1952) the choice investors face in the financial realm has been unquestionably deemed to be between risk and return. However, we see that investors exert a multitude of criteria when choosing investments, with a desire for low risk and high return only being one part of the choice. Furthermore, CAPM assumes a passive role for the providers of investments. We see however that the competitive process continually pushes firms to introduce new products, with differentiating features, to make their product unique in the market, and hence, increase their value to consumers. The financial realm is no different, despite how CAPM presupposes it to function.

Second, the idea of the risk-free asset is questioned. As the existence of a risk-free asset is necessary for the reduction of risk in a portfolio, we find that the lack of one in reality has severe consequences for the model. Risk is a fundamental part of our inter-temporal world, this lack of theoretical completeness almost certainly results from the prior mistreatment of time. Finally, we take the position that all the previous criticisms were misguided, and view the conception of risk within CAPM *as if it were valid*. We find that the actual computation is found wanting, and hence, misstates the risk that investors face, provided that they actually faced the risk CAPM portrays.

Finally we will conclude this work. Both EMH and CAPM are found severely theoretically flawed. As they are inextricably intertwined in the work of finance academia the repercussions are devastating. Many theories must be rethought with this new insight in mind. In fact, we find no truer words today than those of Menger (2006: 47), written over 140 ago:

> In order to avoid any justifiable doubts on the part of experts, we must not, in such an enterprise, neglect to pay careful attention to past work in all the fields of our science thus far explored. Nor can we abstain from applying criticism, with full independence of judgment, to the opinions of our predecessors, and even to doctrines until now considered definitive attainments of our science. Were we to fail in the first task, we would abandon lightly the whole sum of experience collected by the many excellent minds of all peoples and of all times who have attempted to achieve the same end. Should we fail in the second, we would renounce from the beginning any hope of a fundamental reform of the foundations of our science. These dangers can be evaded by making the views of our predecessors our own, though only after an unhesitating examination, and by appealing from doctrine to experience, from the thoughts of men to the nature of things.

The world of finance is to be found in such a state today as Menger found economics proper during his life. The importance and impact of his words have resonated since, and leave the realm of economics forever changed. It is with this in mind that we embark upon a discerning journey into the depths of finance theory with a critical mind. It is hoped that this present work will spark a new interest towards a more theoretically sound assessment and understanding of finance theory, leading us to higher learning.

II. A BRIEF HISTORY OF FINANCIAL ECONOMICS

Markets might be as old as civilization itself, but the financial conception of markets is a relatively new development. For much of human history, wealth existed in a merely physical form – land or the privileges inherent in it. Throughout the Middle Ages lending money was viewed as usury, condemned by the church, and also by governments as a result. The development of the early exchanges – Antwerp in the 16th century, Paris in 1720, or London in 1792 – changed this conception. Capital was to take on a more financial form, transaction costs reduced, and real options (also financial options) created for individuals. With the advent of these increased avenues came a need for a theory to describe the events that were unfolding. The establishment of a powerful (and more importantly, stable), legal system during the 18th and 19th centuries was a crucial enabler of financial wealth, and its ensuing transfer.

For much of its history, finance and economics were two separate and distinct fields. In the early 1960s, several events occurred that would change this distinction, and move finance under the realm of economics. Two of the prime events were the application of intellectual arbitrage to existing economic concepts, and the paradigm shift whereby "financial markets were 'converted' in the eyes of economists from a 'casino' to an idealized market" (Harrison 1997: 174).

The intellectual arbitrage would "squeeze" new ideas from old ones by applying them in new ways. But, like any arbitrage opportunity, the "profits" are fleeting, and large scale ideas have been fewer and farther between. The diminishing returns from these re-hashed ideas have brought a wealth of new literature, but not a wealth of new knowledge. By definition, this approach of arbitraging opportunities in academia must eliminate itself,[2] and today we see more and more inductive observations of the market, and less and less core deductive reasoning. However, this concept of intellectual arbitrage made the realm of finance legitimate in the eyes of economists, who deemed the study of financial markets prior to the 1950s as not sufficiently academic.[3]

Much of this negativeness towards the field of finance could be attributed to economists who had suffered throughout the Great Depression. Keynes (1939: 66) would note that the management of

---

2 Arbitrage opportunities only eliminate themselves in a purely static world. As innovations continue to be created, and "big-bang" ideas are discovered, new areas to arbitrage become apparent.

3 Not only was finance insufficiently academic, many academics were inept to deal with it. One particularly humorous story arose when Harry Markowitz presented his 1955 Ph. D. dissertation on portfolio selection to Milton Friedman. Friedman's response was "Harry, I don't see anything wrong with the math here, but I have a problem. This isn't a dissertation in economics, and we can't give you a Ph. D. in economics for a dissertation that 's not economics. It's not math, it's not economics, it's not even business administration." (as quoted in Bernstein 1992: 60) Markowitz, suddenly alarmed after being initially confident his degree would be awarded, would be awarded his Ph. D..

stock exchange investments of any kind is a low pursuit. Irving Fisher, shortly after his famous comment of the stock market reaching a “permanently high plateau” would suffer in the largest prolonged stock market decline in the history of the United States. In fact, the stock market crash of 1929 would represent the antithesis of the orderly world that economists would prefer to study. As Fabozzi, Focardi and Kolm (2006: 2) remind us, as recently as the 1970s both French and Italian languages referred to investing as “jouer à la Bourse”, or “giocare in Borsa”, that is to say, “gambling in the exchanges.” Also adding to the neglect was the fact that the financial markets in everyday life were of a less prominent role in the pre-1960 world than they would become after.

The lack of understanding as to how financial markets worked would prove to be a source of frustration for economists, even to the present day.

The change to modeling the market as an idealized form also presented a significant challenge. As the financial markets were deemed to be too chaotic to apply the standard economic theories, the market was changed in the eyes of the academics to suit their theories. If their concepts could not apply to a world where down was seemingly up, and up was down, then they would change their conception of the world to one of orderliness. Financial markets were thus viewed as calm markets, where people rationally absorbed and reasoned all the relevant information, *before* making an informed decision.

The main principle behind mainstream academic literature in the post-WWII period was the idea that financial markets are able to keep prices in line with underlying economic fundamentals efficiently. What this means is that prices of assets at any given time accurately reflect the rational sentiment and valuations of all market actors, and that these rational views are fully reflected in prices that accurately portray this. Most literature during this period has used this implicit assumption of rational expectations and market efficiency. In fact, the role of 'irrational' action, seems to be largely down-played.[4]

However, the fact that actors do not always act in a purely 'rational' manner, at least in the neo-classical utility-maximizing sense, cannot be dismissed so easily. In fact, approaches that assume this position are gaining popularity. The static models that these mainstream approaches are based on however are still very much *de rigeur*, a failing which even the most hardened empiricist must, and at times does, recognize. This flaw however, is typically swept away, with the justification that models are necessarily simplifications of real-life and that these simplifying assumptions are therefore necessary.

The influence of Friedman (1953b: 15) is all too real; “the relevant question to ask about

[4] For example, the influence of herding behavior, or mass psychology as they affect actors' decisions.

“assumptions” of a theory is not whether they are descriptively “realistic”... but whether they are sufficiently good approximations for the purpose at hand.” Over time, academics would grow to view financial markets as being the ideal form they longed for. As Harrison (1997: 180) writes:

> Here were perfect markets – a market where the power of arbitrage was supreme; where thousands of individuals with millions of dollars in incentives were pursuing information and pouncing on arbitrage opportunities. The traded good was almost as generic as a widget; there was a plethora of publicly available information; there was easy entry and exit and trading was relatively costless and free from other frictions... What more inviting place for economists to venture?”[5]

To understand how the world of financial economics came to be engulfed in this state of affairs, it is important to first review the two main theories that would bring it about.[6]

---

5 Also adding to the attraction of finance as an area of study to economists are two additional factors. First, the abundance of empirical data makes research that much easier. Second, finance is typically relegated to business schools in Universities, which typically offer higher salaries then economics departments (Bernstein 1997: 182). This monetary incentive should not be overlooked when viewing the current prevalence of finance professors.

6 Modigliani and Miller's (1958) work on how firm financing is related to total firm value should also be included as a seminal work during this time period. However, as it not strictly about financial markets directly, we find it outside the scope of the present work.

III. THE EFFICIENT MARKET HYPOTHESIS

1. A History of EMH

The roots of EMH can be traced to Louis Bachelier (1870-1946), a French mathematician, who published his Ph.D. thesis in 1900 entitled *Théorie de la spéculation.* The focus was the application of Brownian motion to stock-prices, an approach which Bachelier is commonly credited with pioneering.[7] The thesis, although being published in the prestigious French journal, *Annales Scientifiques de l'École Normale Supérieure*, and receiving the distinctive mark of *honorable*, was to be forgotten in the annals of Finance for over half a decade.[8] Despite providing a theory that would have profound repercussions on the realm of finance for more than a century, his idea was a case of the right idea, at the wrong time. Strangely, Einstein would introduce the same concept to physics five years later to ovations and accolades. But, as Fabozzi, Focardi and Kolm (2006: 4) point out, "[u]nfortunately for Bachelier, his reasoning was too economic to satisfy mathematicians and too mathematical to satisfy economists."

In fact, it would not be until American mathematician Leonard Jimmie Savage forwarded a copy to his University of Chicago colleague Paul Samuelson, that the article, and ideas, would re-emerge. The previous 50 years had seen very few quantitative approaches to finance,[9] but with the re-introduction of Bachelier's thesis, a movement was created that would spread quickly. Samuelson began circulating the dissertation to his students and colleagues, one of whom was Eugene Fama. The timing of this re-introduction was optimal, as the computer, and statistical calculation programs, were also becoming more widespread during this period. In fact, it was largely the advent of the computer that further research in this area (Fama 1970: 390). As most research was now almost purely empirically based, the new ability to sift through mountains of data greatly increased academics' ability to further results in this area. Fama (*ibid.:* 389) would note that by the time Samuelson started

---

7 Brownian motion is the idea of random (stochastic) particle movement, originally developed by the Scottish botanist Robert Brown (Pas 1971).

8 Girlich (2002) notes that Bachelier was influenced mainly by mathematicians and physicists, particularly C. F. Gauss and G. T. Fechner. At the turn of the century when Bachelier published his thesis, there was little work done on securities or stock exchanges, and almost all of what was done was through faculties of law. Bachelier remains the first academic to publish a thesis on speculation from a stochastic point of view.

9 See Working (1934), Cowles and Jones (1937), Kendall and Hill (1953), Roberts (1959), Cootner (1962), Mandelbrot (1963), and for several notable articles concerning the random walk approach to finance during this period. Random walk theorists generally posit that price series are random departures, and hence unrelated, from previous prices. They also note that as a given price at any point in time reflects all past information, any future deviations in price must be caused by future information. As this information is largely unknown and unpredictable, price changes must also be unknown and random.

researching the possibility of random walk models, “there existed a large body of empirical results in search of a rigorous theory.”

Fama (*ibid*.: 383) would later note that “[t]hough we proceed from theory to empirical work, to keep the proper historical perspective we should note to a large extent the empirical work in this area preceded the development of the theory.” That such a statement was deemed relevant in 1970 (actually 1969 when it was first written) speaks volumes as to the methodological changes that have engulfed the academic world of mainstream financial economics today. Similarly, Bernstein (1992: 42) reports that in the 1952 edition of the Journal of Finance, Markowitz's article on portfolio diversification was the only article to contain a single mathematical equation.

Fama had his doctoral thesis published in 1965, entitled *The Behavior of Stock Market Prices*, and the landscape of financial economics would be forever changed. Later that same year, Samuelson would publish his empirical 'proof' of this thesis, and solidify the idea of 'efficient markets', along with the idea of random future price movements, into the minds of economists.[10]

[10] Efficiency as it relates to the asset markets would not be more fully defined until Fama's (1970: 383) article which stated, “[a] market in which prices always “fully reflect” available information is called “efficient”.” Later, Malkiel (2003: 5) would define efficient markets as when “they do not allow investors to earn above-average returns without accepting above-average risks.” The distinction between informational efficiency, and efficiency of returns is an important distinction that will be looked at later.

2. Fama's EMH

In Fama's dissertation (1965b), we find that the original aim was to refute the “chartist techniques” used in stock market analysis.[11] This chartist approach is know today under the guise of technical analysis, which is the counter-party to the idea of fundamental analysis. The former holds the tenet that endogenous information is what moves markets and determines future directions, the latter, that data external to the market provides such impetus.[12]

In stark contrast to the search for repetitive events that technical analysts seek, Fama (*ibid.*: 34) concluded that “the path of the price level of a security is no more predictable than the path of a series of cumulative random numbers.” Furthermore, he cited (*ibid.*: 34) the fact that the empirical evidence in favor of such a postulate was in such abundance, that “the counterarguments of the chart reader will be completely lacking in force if they are not equally well supported by empirical work.”

Laying the foundation for Fama's dissertation were two underlying assumptions (1965b: 35):

[1] Successive price changes are independent (i.e., random walk)
[2] The price changes conform to some probability distribution

By the initial assumption, Fama assumes that a given price change in a time period is completely separate and independent of any price change in a previous period. Or, in other terms, that knowledge of the history of prices is of no use in predicting the probability of a future price change. Fama (*ibid*., 35) notes that this is an unreasonable assumption, one that would never exist in reality if held in its strict form. However, he accepted this as an acceptable restriction, even mirroring Friedman's (1953b, 15) earlier words that “assumptions don't matter.” He (1965b, 35) justifies this decision by stating:

> [T]he independence assumption is an adequate description of reality as long as the actual degree of dependence in the series of price changes is not sufficient to allow the past history

11 Given his (1965a: 15) view of chartists as being equivalent to astrologers, we can see the emphasis he placed on this motive.

12 For example, the idea that volume or price level is an indicator of future movements. Technical analysis concerns itself with primarily market generated data of this sort, in hopes of finding repetitive patterns that will hold some predictive value. According to Fama (1965b), “if through careful analysis of price charts one develops an understanding of these “patterns,” this can be used to predict the future behavior of prices and in this way increase expected gains.”

of the series to be used to predict the future in a way which makes expected profits greater than they would be under a naïve buy-and-hold model.

Three additional conditions would be noted later by Fama (1970), such that: there are no transaction costs, all available information is available costlessly to all market participants, and that all actors agree on the future implications of such information for the current price, and future price probability distributions. Again, Fama (*ibid.*: 387) is conscious of the lack of realism this presents as he notes "[b]ut a frictionless market in which all information is freely available and investors agree on its implications is, of course, not descriptive of markets met in practice." He (*ibid.*: 388) later would soften his stance, noting these conditions need not apply as long as "'sufficient numbers' of investors have access to available information."

In actual fact he states that price dependence does exist to some degree, but of an arbitrary amount, insufficiently large to affect his model. Furthermore, this idea of price-independence is viewed as person-independent as well. A practitioner may have a different definition of what it means to have independent prices than an academic does (Fama 1965a: 6). To further the idea of price-independence, Fama (*ibid.*: 36) makes the note that "stock prices may be just the accumulation of many bits of randomly generated noise." By this he means that prices may not be related to any real, intrinsic value in the economic sense, but merely the subjective valuations of each individual actor. That these valuations may not agree with one another,[13] or over time may not agree with one another, will keep prices from affecting subsequent prices. Bachelier (1900) and Osbourne (1959) had both formulated these arguments earlier, as they noted that information arises independently over time, and as this will not follow a consistent pattern, the effect on prices (manifested price changes) must be random as well.[14]

The second assumption, that price changes conform to a distribution, was viewed by Fama as being of secondary in importance to the previous point. In fact, it did not matter what type of

[13] Indeed, short of being in an imaginary world where static equilibrium is the case, or in Mises' evenly rotating economy (ERE), valuations, and hence prices, will always be in a constant flux. He would coin this series of independent valuations as "noise."

[14] Interestingly, Fama would in the same article note empirically the preponderance of results that showed large one day moves, followed by large one day moves, but of opposite direction. That by his own empirical approach this fact was ignored speaks of the scrutiny he attended to his own methodology. It was justified as a market 'overreaction' to information in the first period, corrected in the second. But this explanation does not erase the fact that some sort of dependence exists between successive prices. Also, he (1965b) notes that managed mutual funds statistically earn the same as unmanaged portfolios *after expenses are factored for*. This implicitly means that fund managers earn a higher than random return on their portfolios.

distribution these changes followed, as long as it could be known that they followed *some* definable distribution. However, Fama (1965b: 41) posits that this specific type of distribution was important for future empirical researcher. Later, he (1970: 386) would clarify the position and note the importance of this distribution remaining the same throughout successive price changes.

The conclusion of Fama's thesis was that prices reflected in financial markets must fully reflect[15] all known information, and therefore should reflect the collective belief the relevant actors hold for a given asset. From this it naturally followed that it is impossible to outperform the market using information that is already widely known to the market.[16] Fama (1970:413) would also note that:

> [T]he theory of efficient markets is concerned with whether prices at any point in time “fully reflect” available information. The theory only has empirical content, however, within the context of a more specific model of market equilibrium, that is, a model that specifies the nature of market equilibrium when prices “fully reflect” available information. We have seen that all of the available empirical literature is implicitly or explicitly based on the assumption that the conditions of market equilibrium can be stated in terms of expected returns.

A further refinement of EMH was given by Fama (*ibid.*) as he noted three distinct types of efficiency that a market could embody:

[1] Strong-form efficiency
[2] Semi-strong-form efficiency
[3] Weak-form efficiency

Strong-form tests would focus on whether individuals, or groups, had monopolistic access to the relevant information for price determination. Semi-strong-form tests looks at the possibility that all relevant and obvious public information is available. Finally, the weak-form test would concern itself only with historical prices or returns for future price determination. As he (*ibid.*: 414) would also note, the strong-form is not realistic, but should be thought of as the absolute standard, or benchmark, from

[15] Fama (1970) would later clarify the term 'fully reflect' in purely static terms by using equilibrium prices as developed by Sharpe (1964).
[16] Or in other, more general terms, it is impossible to buy or sell assets at a more optimal price than the market is already valuing them at. This conclusion ignores the chance possibility of luck.

which to view the other two forms.[17] Hence, although strong-form efficiency incorporates all possible information, the weak-form definition takes on a more realistic approach whereby market actors only use information until its benefits are outweighed by its costs.

Fama (*ibid.*: 416) would brilliantly state at the end of this seminal paper:

> [T]he evidence in support of the efficient markets model is extensive, and (somewhat uniquely in economics) contradictory evidence is sparse.[18] Nevertheless, we certainly do not want to leave the impression that all issues are closed. The old saw, "much remains to be done," is relevant here as elsewhere.

In fact, for Fama (1965a: 7) contradictory evidence would source from one of two methods. The first would be to directly test serial returns to see if there was any evidence of correlation between the past on the future. The second would be to try to refute technical methods, through back-testing to see if technical theories held any weight in reality. If neither of these approaches could yield a statistically higher return than a passive "buy and hold" technique, EMH would be empirically proven to hold true. The seemingly endless barrage of supportive evidence, in the face of relatively little contradictory data, suggested Fama's theory to be true. EMH would occupy an untouchable pedestal for many years before any rigorous challenges were forwarded questioning its validity.

---

[17] It is further often assumed that barriers such as insider trader laws limit the possibility of strong-form efficiency. That these laws fully work in practice is debatable in itself.

[18] Empirical evidence, not to be confused with theoretical justification which by Fama's own admission earlier in the same paper, was lacking. Samuelson (as quoted in Bernstein 1992: 123) would be somewhat confused with this methodology as he hailed Fama's work as "a purely deductive, theoretical piece." It is not uprising then given this basic methodological misunderstanding by Samuelson that his students would also be trained similar point of view.

3. EMH As Used Today

Despite previous strong academic support, mainstream investors have been hesitant to grasp and utilize fully the conclusions drawn from EMH. Some books expounding the concepts of the EMH framework, such as Burton Malkiel's *A Random Walk Down Wall Street* have achieved best-seller status, but few investors are willing to accept a passive role in the investment environment. In fact, the existence of such consistently market beating investors, such as Warren Buffett or George Soros, have reinforced in the public's minds the idea that markets are not efficient, and that there must exist a method to achieve above average returns. Fama (1965a: 1) would himself note that EMH failed to grasp the non-academic world's attention, although mainly due to excessive formalism in presentation.

However, the success of some investors, even over long-periods of time, cannot on its own be proof that EMH is flawed. Although the reasoning may look valid *prima facie*, given a world with millions of investors, there is still a random chance that some would enjoy long sequences of success. As Fama (*ibid.*: 12) was apt to remind us, there is a 50% chance that you will be better than the average. Warren Buffett argues otherwise, noting that people who consistently beat the market are not random, but share common approaches, suggesting there are some ways an individual can achieve superior returns. In particular he (1984) points out that his colleagues who practice value-investing consistently produce above average returns inconsistent with EMH. Even Buffett (1996), however, notes that the average investor would earn a better return passively investing in the market, than actively chasing returns. Soros (1994: 47) offers a more scathing retort to EMH followers, stating they owe their belief to the fact that they personally cannot beat the market, and hence, mistakenly believe that above average returns are impossible due to “market efficiency.”

Much of this has to do with a general misconception of what EMH really means for asset prices. Many people take EMH to mean that the current price of an asset is a correct forecast of the future returns of a business. But EMH rather implies that an asset price is the aggregation of the actors' probabilities assigned to the future returns of an asset. Future returns (or prices) may vary widely from what was assigned at one time, but this on its own does not invalidate EMH.

Four Common Myths Surrounding EMH

As Clarke, Jandik and Mandelker (2001) point out, there are four crucial myths concerning EMH. The misunderstanding that has occurred between what EMH really says and what individuals think it says has created a great divide between practitioners and academics.

Myth 1: *EMH claims that individuals cannot outperform the market*.

What EMH really means is that as new information is randomly created, investors cannot *knowingly* outperform the market. New information can be released which is extremely beneficial for an individual's profit rate, however, this will be a chance event. One cannot be expected to outperform the market consistently, or in a predictable manner.

Myth 2: *EMH claims that individual analysis is pointless. A random selection would provide a return equivalent to professional selection*.

EMH does not deny that investors have individual preferences, and that there is a need for professionals to construct portfolios aligned with these personal preferences. Also, it recognizes that the competition between professionals is what disseminates the information throughout the market as prices; *even if it cannot explain the impetus for a professional to do so*. There is evidence that professionals can gain from their analysis, but this will not compensate for the added costs they must incur for their toil, nor when gauged in risk-adjusted terms.[19]

Myth 3:*EMH claims that new information is fully reflected in market prices*.

This is true, although many people think that the constant fluctuation in prices occurs in the absence of new information. In reality, new information is constantly created, requiring constant fluctuation in prices to account for it. If prices remained constant, markets would have to be viewed as inefficient, as they would not be utilizing this new information.

Myth 4: *EMH claims that all investors must be informed or skilled to be able to analyze the new flow of information*.

[19] Cornell and Roll (1981), for example, demonstrate that efficient markets and security analysis simultaneously existing is not inconsistent. An individual who uses information for analysis can achieve a higher return than an individual not utilizing information, *but only in gross terms* – the net returns will be identical for both investors.

This is false, market efficiency can be achieved if only a small portion of the investors are well informed.[20]

[20] However, as Grossman and Stiglitz (1976: 248) point out, equilibrium is assumed to be reached as " [p]erfect arbitrage has one important implication – not all traders need to be informed. The informed traders make prices reflect true values, and the uninformed can simply take advantages of the services provided by the informed." But, in their analysis, as prices are not perfect transmitters of information, disequilibria exist to offer compensation (through profits) to those who seek the information better than others.

3. Mainstream *criticisms* of the EMH

Although enjoying a lengthy period of unquestionable academic acceptance in general, there have always been two aspects of EMH that market actors have found unlikely according to Kortian (1995: 5):

[1] The implication that a price will change only if the market receives new information relevant to the asset's economic fundamentals.
[2] That speculative action must be a stabilizing factor, moving prices towards their intrinsic value instead of away from them.

Furthermore, there are at least three common practices today that seem to explicitly reject the existence of EMH, as each bases its strategy on using past prices as harbingers of future performance.

The first involves the widespread use of technical analysis, which Fama (1965b) originally tried to refute. Were EMH to be fact, one would assume that technical analysts would decrease in relevance and prominence. Although this has occurred in some markets, others have seen a distinct rise in the predominance of technical analysts.[21] One would think that the market test would hold true for EMH, and that if it were reality, market participants would utilize its conclusions more fully.

Second, the actual use of stop-loss orders seems to explicitly falsify the EMH. Stop-loss orders are market buy/sell orders that trigger automatically at a pre-defined level. For instance, if a stop-loss order is set at $10, than once an asset falls to, or through, that predefined level, it will be automatically sold/bought for the investor. This would seem to suggest that past movements of prices do in fact, by definition, influence future prices as well.

Finally, portfolio insurance is an example of a dynamic hedging strategy that also seems to negate EMH. Typically, this involves selling into a falling market, or buying into a rising one.

Several events have also occurred in which stock prices drastically corrected, with no external influence to cause this to occur.

21 Many investment firms have made drastic cut-backs to their technical departments that cover stock markets, but other market segments have seen growth. For example, Frankel and Froot (1990) show the increased usage of technical analysis in the foreign exchange markets. In particular, they note that one survey conducted in 1978 showed only 2 out of 23 analysts using technical analysis, while the same survey given in 1989 showed an increase to 18 out of 31. Allen and Taylor (1990) have noted that in a similar type survey, finding over 90% of foreign exchange dealers at the Bank of England to be using technical analysis.

For instance, on October 19, 1987, American stock prices fell by an average of 22% in a single day. As no piece of important, relevant economic news can be found to explain this event, it could be assumed that the market was not priced efficiently given the existing news at the close on October 17th. For example, Shiller (1989) concludes that the crash is not attributable to a change in a fundamental economic factor, but in a general sway in public opinion. Conversely, Malkiel (2003: 27) tries to support EMH by noting that given a simple objective pricing model such as the dividend discount model, if the rate of return increases only 2%, then the price will decrease by one-third. This sudden decrease would occur in the absence of any subjective reaction to the new information.

Additionally, the Japanese 'Bubble' of the late 1980s seems to suggest a feverish investment mania grasped the world as prices spiraled in speculation. Some people have characterized this as one of the most extreme investment manias to grip the world in that century (Wood 1992). In America, the Dow Jones was pushed up to the 1000 level in early 1966, but would fail to cross that important barrier for another 17 years. How could EMH hold that investors had efficiently priced the information known at that time, for the Dow to fail to better this level by even one point in 17 years? Adding to this conundrum is the fact that relevant news continued to be somewhat positive at intervening times during this span. For instance nominal variables, and even some real variables, such as output, producers' prices and profits continued to increase despite the stagnant stock-market. As Shleifer and Summers (1990: 19) noted following Black Monday, "[t]he stock in the efficient market hypothesis—at least, as it has traditionally been formulated—crashed along with the rest of the market on October 19, 1987."

Beechey, Guen, and Vickery (2000: 6) list three main groups of distinct anomalies that seem to empirically argue against the existence of EMH, and its random-walk associate:

[1] Value effects – portfolios constructed of "value" stocks (those with high earnings or cash flow to price ratios) appear to exhibit above average returns over time. Portfolios with below-average past returns tend to return higher rates over subsequent periods.
[2] Momentum effects – the above noted value effects seem to hold over longer time horizons, but over shorter periods an opposite effect holds. High returns in the immediate past tend to continue producing these same higher returns over a 3-12 month period.
[3] Size anomalies – small stocks exhibit higher than average returns.

Furthermore, as information is gathered, profit opportunities disappear. This implies that well known trading strategies would trend towards a zero return over time. French (1980) discovered what he coined "The Weekend Effect",[22] whereby returns analyzed from 1953-1977 showed a tendency for US stock returns to be negative on Monday and positive for other days of the week. Steely (2001) finds that this effect has diminished, at least in the UK, since the article was published. This result is fully coherent with EMH, and the dissemination of new information. However, in Banz (1981) we saw that small firms provide excess stock returns to large firms. This effect is still in force today. Were EMH, as it was originally stated, true, why have these pieces of information been acted upon differently?

Rational Bubbles

The first serious attempt at academically determining an alternative explanation for the existence of bubbles, while still keeping the core of the EMH intact, was the "rational bubble" theory created in the late 1970s.[23] The goal of this theory was to explain why asset prices would diverge from their underlying values under specific circumstances. Even given rational expectations and actions, a bubble could still form, albeit a 'rational' one, bringing asset prices out of line with their intrinsic values.[24] This type of bubble would be defined as one whose price trend deviates ever more quickly from the value justified by its underlying market fundamentals. It is reinforced by the belief among otherwise rational actors that an asset's price somehow contains more value than the underlying fundamentals dictate. As EMH dictates that there cannot be any profit opportunities as they would already have been exploited, hence, it naturally follows that in the existence of a rational bubble, profits would be equal to those that would exist in the absence of such a bubble (Ohanian 1996).

Rational bubbles are thus justified when an actor purchases an asset with the intention of selling it later at a higher price. This in turn relies on another investor being willing to purchase at this higher price, and hence, keep the process in motion. This reliance on ever more investors willing to purchase assets at ever higher prices is the core of the rational bubble model. However, in this type of model, the reliance is also on a dynamic pricing model, where the "equilibrium price in the current period depends on expectations about future changes in the asset's price" (Kortian 1995: 10). Adding this dynamic component to the EMH provides a huge step forward in making the model more realistic, and hence, accurate.

---

22 Or better known as The Monday Effect in some circles.

23 Blanchard (1979), Flood and Garber (1980) and Tirole (1982) provide important contributions to the theory of rational bubbles.

24 Intrinsic value here is generally given as the present value of the future expected cash flows of the asset.

This model has not been without its own mainstream critiques, some of who originally developed it. Tirole (1985) for instance, shows that as new investors enter the market all the time, and as their individual time horizons (and hence expectations) all differ a rational bubble can exist only in a dynamic framework. Tirole defines dynamically efficient markets as those growing at a higher rate than the rate of interest. Further, if this condition is violated, there will exist a definite point in time when the existing wealth of the investors will be insufficient to purchase the asset at the inflated bubble price.[25] Previously, Tirole (1982) had demonstrated that bubbles could not occur in assets that have a finite life, and hence, a finite, determinable value.

As originally developed, the bubble component of the rational bubble was separate of the fundamental component. As a result, it was driven by purely exogenous factors. Academic work from the early 1990s started to question this relationship.

Intrinsic bubbles were a distinct class of rational bubbles developed by Froot and Obstfeld (1991). They noted that departure from fundamental values can persist for extended periods of time due to the fact that underlying economic fundamentals begin to be associated with this extended over/under valuation. For example, the initial mis-pricing may reinforce future mispricings. The implicit conclusion is that prices somehow initially over/under react to the new relevant information, and this pricing error carries throughout time to create a bubble.

Dornbusch (1976) used stickiness of real variables to explain bubbles. In his “overshooting model”, he assumes that financial variables have low to no stickiness; they adjust almost instantaneously to new information. However, real assets exhibit high degrees of stickiness, hence, they exhibit a lag in adjusting to new information. In the short-run financial assets will adjust to new information instantly, while the real assets lag. Over time, as real assets move towards their equilibrium values, the financial prices will move to reflect this also. The two values will eventually reach agreement at equilibrium. As Dow and Gorton (1997: 1115) point out, there is a significant difference between financial and physical goods' prices. Physical goods' prices are “Hayek prices”; we need not know how they were determined, only how they affect our allocation of resources. Financial asset prices serve to direct resources in an indirect way, hence, their very formulation is important to understanding their implications. They conclude that there is only a tenuous link between the efficiency of financial prices, and the efficiency of physical goods' prices; one need not rely on the other.

[25] Kortian (1995: 16) notes that this view-point looks at investing as a zero-sum game. As the people who gain from the sale of assets at inflated prices equal the losses that are incurred by the purchasers, this would leave the sum of these purchasers in a “negative-sum game”. Tirole (1985) argues that rational purchasers would recognize this, and refrain from the further purchases that would reinforce the bubble.

Along this line of thought, there is also the admission that at least some people will have a destabilizing effect on prices, generating excess volatility by reacting to fads, fashions or rumors.[26] This viewpoint, that asset markets may experience periods of time marked by irrational investor behavior was made popular due to the empirical work of Shiller (1984; 1989). Shiller brought to light the effect that mass psychology played on investors' susceptibility to fads. These fads would in turn cause investor sentiment to shift, at least in the short-term, and cause asset mispricing.[27]

Fads Model of Asset Prices

Following the lead of Shiller's note on the influence of mass psychology on asset price trends, a new model for asset prices emerged – *The Fads Model*. A fad may been defined as a divergence of an asset's price from its fundamental value due to a socially induced change in market sentiment or opinion. Unlike the previous rational model however, the price divergence here manifests not at an explosive rate, but at a slow orderly one. This is due to the diffusion of the fad throughout the actors, and the rate at which it exits. These two factors give rise to the inflating and deflating of the bubble, and have been referred to as the "infection rate" or "removal rate." Shiller (1989: 56) sums up this viewpoint thusly:

> A fad is a bubble if the contagion of the fad occurs through price; people are attracted by observed price increases. Observing past price increases means observing other people becoming wealthy who invested heavily in the asset, and this observation might interest or excite other potential investors.

Proponents of EMH have criticized the idea of fads by noting that arbitrageurs would exploit these apparent mis-pricings due to their profit opportunities. Shiller counters that this argument overlooks the main principle underlying EMH. As EMH implies that the future path of assets is unknown, even a rational investor would be unable to exploit mispricings that occur due to irrational behavior The inability of an investor to correctly forecast a price does not prove the existence of EMH, at least in the sense that asset prices will reflect their fundamental value (Kortian 1995: 19). It only shows that the

[26] Kindleberger (1989) provides a historical account of speculative frenzies that have occurred in this manner.

[27] However Shiller was not the first to note these irrational actions. Nurske (1944) noted that the excessive volatility experienced in the inter-war exchange rates were the result of irrational speculation that amplified otherwise small fluctuations into larger movements. Nurske thus viewed currency speculators as having a destabilizing influence on the market. Also, Keynes (1936) also made note of the irrationality of the markets, stating "the market can stay irrational longer than you can stay solvent."

future price trend is unpredictable, and hence, not exploitable by even rational investors.

Noise Traders

The idea that rational speculation will stabilize prices, and reduce irrational or destabalizing behavior, was first forwarded by Friedman (1953a). However the view of the speculator as a stabilizing factor is being challenged by the noise trader approach pioneered by de Long, Shleifer, Summers and Waldman (1990) and Shleifer and Summers (1990). This approach has recently been gaining wider acceptance among the academic community.

The core of the noise trader framework is based on two propositions:

[1] The co-existence of heterogeneous investors and traders
[2]Arbitrage limitations

The first proposition implies that there exist two distinct classes of actors have two distinct classes of behavior. The first are the rational speculators (arbitrageurs, smart money investors, fundamentalists) who base their actions on economic fundamentals. This group is assumed risk averse, with a relatively short time horizon. The second group are the noise traders, who display at least some degree of irrationality. They are assumed to be less-sophisticated and more susceptible to fads, rumors, or other exogenous information. In short, their actions cannot be fully explained by the market's fundamental data. The existence of this class of investors can cause a demand shift that is unexplainable by fundamental methods.

The second proposition assumes that a limited capacity to arbitrage exists, which is unable to fully counteract the demand shift noted above. Two effects work to limit arbitrage opportunities: (1) fundamental risk and (2) the uncertainty of future resale prices.

Noise traders may cause rational investors to rethink their own analysis in light of the new information provided through the market (price or volume data generated by the noise traders). In fact, this loss of confidence in their own objective analysis of the fundamental data may lead these 'more sophisticated' investors to be unable to discern even large asset mispricings.[28] The fact that even the

[28] This issue seems most pronounced in the foreign exchange market, in which the economics profession is not yet able to formulate a successful method to value rates. Meese (1990) states that "the proportion of (monthly or quarterly) exchange rate exchanges that current models can explain is essentially zero." This conclusion was similar to the earlier Meese and Rogoff (1983) conclusion, although has been contested lately by Mark (1995) who views the issues of non-

rational investors have a relatively short time horizon also complicates matters as it contributes to a general lack of foresight. This time horizon issue is attributable to the fact that most professional money managers are evaluated over a relatively short period of time. Hence, their goal will also be the maximization of profits over this same limited time horizon.

Strict advocates of EMH have criticized noise traders on grounds that they must incur large loses, at least in the longer-term, and will be forced to exit the market eventually. Thus, their overall effect on prices must be limited and marginal. Noise traders have countered that there are several reasons for their continued existence.

As noise traders consistently over/underestimate returns or risk, they are, on average, more bullish or bearish than fundamental investors. As upward trending markets reward those with a more bullish stance with more profits, noise traders should realize greater profits than their fundamentalist colleagues (de Long, Shleifer, Summers and Waldman 1990). This profit incentive will attract new actors to the noise trading world. Conversely, there will also be noise traders who suffer losses, but even those individuals may remain attracted to the approach, given the gains of others. New 'recruits' could erroneously attribute the high returns to skill, rather then the greater risk undertaken by the more bullish/bearish approach. The long-term survival of the noise-trader approach is not as easily dismissed as some EMH advocates think.

Information Aggregation

The third broad front that EMH is being criticized is on grounds of how information actually is spread, and utilized, by market participants. Information is not always complete, or relevant to the pricing of assets. Furthermore, some participants will ignore or miss information which is highly relevant to the pricing process. This can lead to substantial departures of asset prices from their intrinsic values.

The focal point of this strand of thinking is on the interaction between asymmetrically informed traders.[29] This views informational flows as being a process, with variable speed at which new, relevant information can enter, and disperse throughout the market. Additionally, information under this viewpoint is heterogeneous, which poses distribution problems of its own.

One of the central tenets of this new body of theory is what is known as the “paradox of fully

predictability of formal models as only a short-term phenomenon. This in turn is challenged by Faust, Rogers, and Wright (2001) who concur with Meese that models have little predictive value.

[29] The idea of asymmetric information flows was pioneered by Stigler (1961). Grossman (1976; 1981) and Grossman and Stiglitz (1980) have forwarded the relevant literature for asset pricing.

revealing rational expectations equilibrium." A fully revealing rational expectations equilibrium is typically where all diverse information is aggregated and reflected by the equilibrium price. Grossman and Stiglitz (1980) demonstrate that market efficiency and competitive equilibrium are incompatible ideas given information costs. Moreover, if information has a cost associated with it, there can be no competitive equilibrium price that fully reflects all such information. They argue that if markets were informationally efficient, there would be no incentive for an actor to exert resources to collect new information, it would already be fully reflected in the asset's price. Information could be collected at no cost by merely observing the market. However, were this the case, why would any actor at any time have an incentive to create, transfer, or use information? Malkiel (2003: 33) agrees with the study's conclusion be stating, "the market cannot be perfectly efficient or there would be no incentive for professionals to uncover the information that gets so quickly reflected in market prices." Grossman and Stiglitz (1980: 393) propose the following, somewhat contradictory, solution in regards to informational efficiency in the market, that "there is an equilibrium degree of disequilibrium."

According to Gennotte and Leland (1990), many market participants do in fact price assets in this manner. By only observing the current market price, and ignoring other relevant data, many participants make pricing decisions that further drive an asset's price. Crashes, or other drastic price adjustments, arise from the inability of actors to distinguish between "information-based trades" and "informationless trades". Put another way, actors cannot distinguish between prices that arise from the arrival of new information, and those that arise due to reactions, such as dynamic hedging strategies.

Romer (1993) states that price changes that are not a response to new external information may still be considered rational. This is due to two facts. First, that new information is dispersed among multiple actors, and the market is unable to immediately 'digest' all relevant information, instead requiring some time period to do so. Second, the pricing process itself can reveal information that has been previously released, but not yet fully comprehended by the market. This viewpoint can help explain large price movements that occur without corresponding new information (i.e., the stock market crash of 1987).

Bikhchandani, Hirshleifer and Welch (1992) contribute to the viewpoint when they discuss "informational cascades". Under this model, it is rational for actors to change their assessment after viewing the actions of previous actors. This helps to explain why uniform action seems to manifest by a large number of actors, who could possess different, and even conflicting, pieces of information. The conclusion is that it is not always irrational to ignore personal, or private, information, and to participate with the 'herd'. Prices based on these cascades are quite fragile in nature, as they can potentially be based on very little 'real' information. As Kortian (1995: 34) notes however, “individuals become part of a cascade when it is *optimal* for them to *ignore their own private information* and follow the actions of others.”

Huberman and Regev (2001) show that how the information is provided can effect prices. They found that a stock price for a firm searching for a cure for cancer more than quadrupled the day this information was published in the *New York Times*. The information had been publicly available for several months prior through other media sources, but the effect the *New York Times'* piece had was distinct to the effects provided on the stock's price through other exposures.

Kurz (1992: 1) notes that it is impossible to possess complete knowledge under a dynamic system: “the assumption that agents possess complete structural knowledge has no empirical support... [T]here is much in the dynamic structure of the economy which cannot be learned with certainty.” It logically follows that two actors receiving the same information will price it differently. This implies that over time, these inherent mispricings can build on one another, creating larger mispricings.

4. Today's Mainstream Perception of the EMH

Some people assert that given these academic fronts waged against the core theoretical framework of EMH, it has lost its role as the premier model for pricing assets. This is correct to some degree, but much of the work done is implicitly justifying the framework, although trying to add new caveats to it. The acceptance of the hypothesis, especially within academia, runs deep. Miller (1991) surmises that, although it may be flawed, the concept of efficiency outlined in EMH is still the best we have, and hence, should not be ignored. The three groups of individuals vying to correct the EMH have taken separate approaches to increase its applicable validity.

The first group of mainstream opponents, the “Rational Bubble” group, try only to determine that bubbles, or asset mis-pricings, can exist and have a rational explanation in line with the EMH. In fact, as EMH states that prices are only priced accurately based on the aggregate of known information, the existence of a bubble *per se* does not invalidate this, its existence could merely be the result of the erroneous interpretation of the known information.

Likewise, “Noise Traders” could also be utilizing their given information correctly. As the correct intrinsic price cannot be known before hand, Noise Traders are only pricing, and acting on those prices in the way they deem most efficient. The existence of a Noise Trader is fully compatible with EMH, even though it is a new piece of information to consider when pricing an asset.

Only the third group of mainstream critics, the “Aggregated Information” group, has made a serious attempt at undermining the efforts of the EMH crowd. By attacking the core tenets of the EMH, that information is costless and instantly dispersed throughout the actors, and that actors rationally act on that information, this group has formed a basis for future critiques of the EMH.

Even Fama (1991: 1575) believes that the EMH, in its extreme strong-form existence, is unlikely to hold in reality; “[Strong-form EMH] is surely false...[but] it is a clean benchmark that allows me to sidestep the messy problem of deciding what are reasonable information and trading costs.” If the strong-form is to be viewed as a benchmark of sorts, is it suitable for this role? Mises' ERE also provides an important benchmark from which we can add changes to the economy and see their effects in isolation. However, as will be argued, the EMH has internal inconsistencies that make it ill-suited for even this task. Another important point if EMH is used as a benchmark, is that when viewing deviations from it to assess ensuing changes, the deviations must be realistic representations of our world.

One common link joins the three groups of mainstream criticisms, and whether you agree with their conclusions or not, places them on theoretical shaky ground. This link is the methodological basis, which Fama (1970: 383) saw as being a problem with his original development of EMH. This purely empirical approach proves nothing, and in fact, detracts from the general usability of the aforementioned theories. The lack of any pure theoretical approach, that would create a true asset pricing model irrespective of future events, is yet to be developed. Additionally, all these new theories are simple "amendments" of the original EMH (Shostak 1997: 38). The general thesis of informational efficiency and equilibrium conditions remains.

IV. THE CAPITAL ASSET PRICING MODEL

1. The Pre-History of CAPM

If the origins of EMH coming from an obscure French mathematician seem strange, the early origins of the CAPM may be viewed as downright bizarre. The idea of pricing a security as a product of its risk has also enjoyed a colorful history. The model, which was created to price a single security, has its roots in an older, more complicated portfolio pricing theory theory that was pioneered by Markowitz (1952). As Markowitz (1999: 5) notes however, one of the earliest allusions to the model can be seen in Shakespeare's *The Merchant of Venice* as Antonio would note:

> My ventures are not in one bottom trusted,
> Nor to one place; nor is my whole estate
> Upon the fortune of this present year;
> Therefore, my merchandise makes me not sad.
>
> (Act I, Scene I)

Clearly, Shakespeare understood the trade-off between risk and return.[30] In fact, Antonio can be seen to practice the epitome of diversification, across geographic locale, industry, and time. All that was needed to complete this idea was, in Markowitz's eyes, a theory. Although he has primarily been bestowed the title of “Father of Modern Portfolio Theory”, it should also be noted that the same year as Markowitz developed his idea, Roy (1952) had done much the same. In fact, Roy not only was able to effectively demonstrate the return of a portfolio was the weighted average of its components, but also that the risk involved not only that which was firm-specific, but also intra-firm as well.[31]

[30] Rubinstein (2002: 1041) shows that a closer predecessor may be found in Bernoulli (1954): “it is advisable to divide goods which are exposed to some small danger into several portions rather than to risk them all together.” Of course, we are familiar with Sancho Panza advising Don Quixote that the wise man does not venture all his eggs in one basket.

[31] It is interesting to note that, despite publishing their respective articles in the same year, and that they essentially forwarded the same theory, Markowitz was awarded the Novel Prize in Economics for his contributions to portfolio theory in 1990, while Roy was not. Markowitz could only surmise the reason was that Roy had only written one article on finance, albeit a very important one, and then disappeared from the field, while he, himself, was still quite active in 1990, having written two textbooks and an assortment of articles. The author in no way implies the opinion that Markowitz and Roy should have shared the Prize, nor that Markowitz should have won it at all. See Markowitz (1999: 6) for more on his view of Roy. In contrast, Friedman (as quoted in Newsweek 1990) would later say that he did not think that either of Markowitz, Miller, nor Sharpe were deserving of their shared 1990 Nobel prize: “If you were to ask me to list 100 major contributions to economics, these [contributions of the three winners] wouldn't have made my list.”

If Fisher (1906) was the first economist to use variance of return as a measure of risk, Hicks (1935) was the first of the modern economists to realize the advantages of diversification. By pointing out that an investor's choice to hold cash would balance their propensity for low risk, as well as the want for higher return. Additionally, Hicks would demonstrate that holding multiple assets simultaneously could reduce total risk. However, no systemic study would be made of the risk/return trade-off that would later distinguish the CAPM. As a result, no distinction was made between what would later be coined “efficient vs inefficient portfolios.”[32]

As Constantinides and Malliaris (1995) bring to light, neoclassical economists overlooked the question of asset allocation. Portfolio selection was something outside of their realm. Marshak (1938) was, perhaps, the first to deal effectively with risk and uncertainty, viewing money as the hedge against the uncertain future. However, as Markowitz notes, Marshak himself did not feel his contribution to portfolio theory was that significant. In fact, as he (1999: 12) recounts a story from his graduate days at the University of Chicago:

> My thesis supervisor was Marschak himself, and he never mentioned Marschak (1938)... Marschak kept track of my work, read my dissertation, but never mentioned his 1938 article.

The main contribution of this work would be the addition of human preferences to investment decisions. Actors would not just passively choose assets, but would take into account their aversion to waiting, desire for safety, and other factors not present in static worlds. Hence, Marshak (1938: 323) would demonstrate that cash holdings are held to guard against future uncertainty. This uncertainty, which manifested as a wider dispersal of returns, would incite an investor to hold more cash. The article would fail to specifically mention that actors were acting within a portfolio framework, instead treating all holdings as part of a balance sheet of assets. This same premise would be laid-out by Williams (1938) in the same year. Williams would employ the law of large numbers to show that within a framework of definite assets, given the inclusion of a sufficient number of securities, gains and losses will offset on another, and a form of insurance will occur whereby expected yield will be quite close to the actual yield. Williams would imply that variance, or risk, can be completely eliminated from a portfolio, a conclusion that Markowitz would later deny. The greater contribution of Williams, and also foreshadowed by Graham and Dodd (1934) is that risk cannot be viewed in isolation, instead, the

[32] Although lost to the annals of history, the person that coined these two terms (efficient and inefficient portfolios) did the financial world a great service, saving us from Markowitz's original term, “efficient mean-variance combinations.”

marginal risk added to a portfolio is what concerns the investor.

Leavons (1945) would meaningfully contribute by pointing out that there are two separate types of risk involved in holding a security. Sometimes, it will be acted upon by specific, independent causes that cannot be fully managed. However, diversification may not protect a portfolio against a second type of risk, that which would affect the whole industry. In the event of this possibility, the suggestion was diversification into separate industries, however this may also not protect against fluctuations that could affect all industries simultaneously.

2. Development of CAPM

With this precedent in place for viewing return as a price of risk, much groundwork had already been established by the time Markowitz published his article *Portfolio Selection* in 1952. We will see that the development of the model has progressed through four distinct phases. The first commenced with Markowitz (1952) and ended in the early 1960s. The focus of this stage was on finding efficient portfolios with minimal variance for a given return. Phase two really took off with Sharpe (1963) and would continue through the decade. The emphasis was on formalizing the model, with the introduction of beta, and the further development of efficient portfolios. The CAPM would be formally defined, with its ability to assess risk and return profiles for individual securities, instead of portfolios. Black (1972) would officially kick-off phase three, as he relaxed the assumptions inherent in the risk-free asset utilized before. Various developments over the remainder of the 1970s would continue refining CAPM, and introducing more realistic assumptions. Lastly, a final, continually ongoing, era from the late 1970s to today has seen a movement of empirical testing of CAPM, with mixed results. The debate as to whether the model has validity, or if so the magnitude of its importance, continues to this day.

We will take a look at each phase, before examining in more detail the basic CAPM, which is defined today as based upon Sharpe (1964), Lintner (1965) and Black (1972).

Phase I

Markowitz quickly established the assumption that would forever alter the way asset pricing was achieved in his first paragraph, “we consider the rule that the investor does (or should) consider expected return a desirable thing *and* variance of the return an undesirable thing” (1952: 77). Hence, investors would not be viewed as merely maximizing discounted returns on assets any longer, but would be doing so within the context of minimizing risk as well. The crux of Markowitz's article was that risk (as measured by standard deviation of a return) was not additive, and that, as a result, when assets are held together, their total risk is less than the additive sum of the individual holdings' risk measures. He was keen to add that the goal was not to continue adding securities to a portfolio to minimize risk, but to add securities that had little correlation with the existing risk. The example provided of a portfolio of 60 railway stocks, which would offer little diversification or mitigation against risk, as they would all be subject to the same industry-wide risk. Hence, for Markowitz, risk was seen as inter-industry, with firms of the same industry generally being affected the same in regards to risk or the variance of returns.

Markowitz's (*ibid.*: 77) first lines of his article read:

> [The process of selecting] a portfolio may be divided into two stages. The first stage starts with observation and experience and ends with beliefs about the future performances of available securities. The second stage starts with the relevant beliefs about future performances and ends with the choice of a portfolio.

He (*ibid.*: 91) fittingly perhaps ends with the following disclaimer:

> In this paper we have considered the second stage in the process of selecting a portfolio. This stage starts with the relevant beliefs about the securities involved and ends with the selection of a portfolio. We have not considered the first stage: the formation of the relevant beliefs on the basis of observation.

Although able to determine a measure of portfolio risk, and expected return, Markowitz was unable to see the risk-return trade-off except in terms of multiple securities held simultaneously in a portfolio. It would take some years before new academic ground would be broken in this direction.

In the lead up to Sharpe's (1964) basis for asset pricing, Tobin (1958) would provide one additional crucial contribution. By introducing expectations into the picture, he was able to show that the want to balance a risky asset against an un-risky one was not due solely to the investor's expectation of future risk and return trade-offs, but also the elasticity of these expectations. In particular, he sought to demonstrate that the less elastic an investor's expectations were, the demand to hold a less-risky asset would increase. In other words, if a person was unable, or unwilling, to change their perception of future expectations, they would demonstrate this through increased cash holding in their portfolio. Most importantly perhaps, Tobin (1958: 65) would question the wisdom of holding cash as a hedge for future uncertainty, instead of an safe interest-bearing alternative:

> Why should anyone hold the non-interest bearing obligations of the government instead of its interest bearing obligations? The apparent irrationality of holding cash is the same, moreover,

> whether the interest rate is 6%, 3% or ½ of 1%. What needs to be explained is not only the existence of a demand for cash when its yield is less than the yield on alternative assets but an inverse relationship between the aggregate demand for cash and the size of the differential in yields.

Hence, Tobin would also demonstrate that holding a "risk-free" cash equivalent, such as a short-term government bond, would provide similar liquidity to cash, but at a positive return.[33] Additionally however, he would point out that as relative yields on cash equivalents increased, the demand for cash would also falter. This would hearken back to Keynes' theory of liquidity preference, which was more clearly stated by Kaldor (1939a: 15), "[i]t is ... not so much the *uncertainty* concerning future interest rates as the inelasticity of interest expectations which is responsible for Mr. Keynes' 'liquidity preference function'."[34] Tobin would modify this view somewhat however. Now cash, or an equivalent risk free-asset, would be held in a portfolio not due to expectations of uncertainty, but the assumption of the expected capital gain or loss on cash as being zero.

Previously, all assets were viewed as being risky. This changed that viewpoint and resulted in Tobin's "separation theorem." Portfolios could be split into two distinct parts. The first would be the optimization of the risky assets to be held, in order to maximize the return for a given risk level. The second phase would be the addition of a risk-free asset, and the allocation of funds between these two parts. Hence the risky and the risk-free would be balanced out for the risk-return profile the individual requires.[35] Furthermore, the optimal portfolio of risky assets held would be identical for all investors. The *ratio* between two risky assets would be the same for all investors, although more or less risk averse individuals may choose to buy these assets in greater *absolute* quantities. The significance of this insight is that every individual faces an identical set of opportunities, regardless of their wealth. The fine-tuning of an individual's portfolio would therefore be placed in the hands of the risk-free asset. Hence, the separation theorem's conclusion that each individual shares an identical opportunity has an important corollary – each individual has an identical price of risk reduction.

Markowitz (1958) would bring this all together in a monograph published by the Cowles

[33] Tobin would attribute this insight to Samuelson (1947: 123) where he concluded that, in situations of certainty, stocks and money would both yield the same return. Samuelson would be incorrect in his assumption that, in situations of certainty, money would even exist. See Mises (1949: 417) for this contrasting viewpoint.

[34] Or, in Keynes' (1936: 201) own words, the desire to hold cash is based not on "the absolute level of *r* but the divergence from what is considered a fairly *safe* level of *r*."

[35] Several years later, Hicks (1962) would demonstrate the same dichotomy investors face, explicitly stating the trade-off between the risky and the risk-free portions of the portfolio.

Foundation. The problem that faced investors was outlined in precise terms. There existed a dichotomy of wants among investing agents: (1) the want for high return, and (2) the want for a stable return (*ibid.*: 6). Hence, the result of this monograph was the clarification of four requisites for Sharpe to later define the CAPM. The first was the separation of efficient from inefficient portfolios. The goal of an efficient portfolio was to obtain the highest return for a given amount of variance, or risk. Second was the definition of combinations of return and risk available in a portfolio. Investors were assumed risk averse, and chose a portfolio at a given time $t - 1$ that produced a return (assumed stochastic) at time t. In other words, market portfolios were established to show the possible profiles available to be constructed. Third was that an investor needed to determine what risk/return profile suited their needs optimally, which lent itself to the fourth point, to design the portfolio that best matches their risk profile.

For Markowitz, the two most pressing concerns involved uncertainty and correlation of returns. He (*ibid.*: 5) would liken the correlation of returns to the probability of a coin toss, "[i]f security returns were not correlated, diversification could eliminate risk. It would be like flipping a large number of coins." The confusion surrounding the probabilistic nature of returns will be returned to later in part II. He would also note that investors may have a multitude of criteria when selecting a security, but that the risk-return trade-off was the one common criteria all would share. Also, uncertainty was defined as a random variable. In fact, later (*ibid.*: 39) he would illustrate the concept with a "wheel of fortune", with returns having distinct, known probabilities of occurrence. Strangely, in a smaller "Generalizations" section near the end of his monograph, Markowitz would note that returns are not random variables, but rather uncertain events. The distinction implies that investors must base their decisions on beliefs instead of objective probabilities.

Markowitz (*ibid.*: 303) closed his monograph with the statement:

> In designing a portfolio analysis, two objectives should be kept in mind. First, an attempt should be made to keep the analysis simple; second, an attempt should be made to understand the salient implications of major simplifications.

It is difficult to say whether either of these objectives have been heeded. The calculation of efficient portfolios giving minimal variance for a given return was the hallmark of this first phase. A defining point at this stage of development was that investors only made a binary choice with portfolios:

variance and return. However, some issues would work against the early success of these contributions. Computational difficulty was the main failing point. The calculations necessary to determine these combinations were quite complex for the technology of the day. In themselves, the calculations were quite numerous as well. What was needed was a simplification, and a refinement to the contributions of this stage.

Risk and the separation theorem at the end of phase one

Markowitz's (1952) trade-off between risk and return marked a hallmark for security analysis. Tobin's (1958) "separation theorem" completed this final phase, theoretically at least, as it showed that a risk-free asset could be used in conjunction with a risky asset to temper the trade-off to the individual's risk aversion. The following section shows a brief derivation of these risk measures that marked this stage of the theory's progression.

Determining risk

The determination of risk is fairly straightforward. First, the investor must establish a list of the possible returns that are possible. This list specifies expected scenarios regarding two metrics: (1) expected return, and (2) the probability that that expectation is met. Hence, we can see a probability distribution for expected returns will be created. The total expected return for security – E(r) – is then the sum of its expected returns – r(s) – multiplied by the probability belief held for each – p(s).

(3.1) $$E(r) = \Sigma\, p(s)r(s)$$

We can view risk as being the surprise that an actual return yielded is different than what was initially expected. Hence the "uncertainty" that surrounds an investment can be expressed as a function of the magnitude of possible surprises occurring in the future. Note that risk is based on the *ex post* return, compared to the *ex ante* expected return.

In order to summarize risk as a single number, we can use the variance between the expected return r(s), and the average of the expected return E(r).

(3.2) $$Var(r) \equiv \sigma^2 = \sum p(s)\,[r(s) - E(r)]^2$$

The deviations are squared so that the negative deviations do not offset the positive ones. Hence, any variance from the mean is risky. Bodie, Kane and Marcus (2001: 157) refer to this measure of variance as the “expected value of squared 'surprises' across scenarios.” It must be noted, however, that squaring the deviations is a nonlinear transformation. It will exaggerate large deviations, and under-emphasize smaller ones. Also, note that the measure of risk, σ^2, is a squared term while return, r(s), is not. Hence, to find a comparable measure of risk, the root is taken of the variance:

$$\text{(3.3)} \qquad \text{standard deviation (r)} \equiv \sigma = \sqrt{\text{Var(r)}}$$

One drawback of this measurement of risk is that it treats negative and positive deviations from the norm symmetrically. We see that in reality, we are only primarily concerned with lower returns than our expectations dictate. However, if the distribution of returns is symmetric (i.e., follows a normal distribution, or bell curve), that is to say, the probability of negative and positive surprises are equivalent, σ will approximate the risk that measures solely the negative deviations (Bodie, Kane and Marcus 2001: 158). Additionally, we could use a measure such as semivariance, which concerns only deviations on one side of the mean. Hence, instead of using a mean absolute deviation measurement for risk (as we have previously), we could utilize a semivariance measure which only concerns itself with the negative deviations from the mean that the investor would not prefer. For returns which are symmetrically given around the mean, or normally distributed, the absolute mean deviation will be equivalent to the semivariance. For this reason, and the inherent computational simplifications consequently involved, absolute mean deviation is the commonly used measure of risk concerning random security returns.

The risk premium

The risk premium that we are concerned with when viewing a security is gauged in terms of return in excess of what can be earned on a risk-free asset – typically a short-term government bond. The risk-free rate is assumed to be a certainty in receipt.

Hence, a risk averse investor is one who takes on more risk than usual; they are a speculator. The speculator takes on the added risk as they deem the risk-return trade-off to compensate them accordingly. This is different than the gambler who takes on added risk with no, or little, expectation of being adequately compensated for it. If we view risk aversion as the amount of risk taken, $E(r_p)$, in excess of the risk free-rate (r_f), it can be measured as;

(3.4) $$E(r_p) - r_f = F * A * \sigma^2_p$$

The term F is a scaling factor, of no consequence but to make the result more manageable, and the term A is the measure of an investor's risk aversion. We may like to rewrite the equation to isolate the A term as:

(3.4a) $$A = [E(r_p) - r_f] / [F * \sigma^2_p]$$

It is important to note that there is a large difference between the risk-free rate, and the expected risk; one is observable and the other not. At least, it is not observable in an *ex ante* sense. The risk-free rate can be taken to exist at any time. However, the true risk-free rate will not be revealed until after an investment is complete. This expectation is always available in the investor's mind. However, it is important to keep in mind that the two are never measurable *at the same time*.

Portfolio risk

If we now have a measure to distinguish between risk and return, what occurs when we exercise Tobin's (1958) “separation theorem”? It becomes apparent that any portfolio, which by definition holds more than two investments, may be spread between risky and risk-free assets.

If we spread our investments between two options, risky and risk-free, the risk premium now becomes a fraction of the risky asset held. If y is the percent of the portfolio held in a risky asset, R_c becomes the expected return on the portfolio.

(3.5) $$E(R_c) - r_f = y[E(r_p - r_f)]$$

In figure 1, we can see the options available with the inclusion of a risk-free asset. Combinations that include both risky and risk-free assets are represented on the 'Capital Allocation Line.' Point A represents an investment 100% ($y = 0$) invested into a risk-free asset. Point B represents 100% ($y = 1$) invested in the risky asset. Hence, at point A the return is the risk-free rate, and the standard deviation, or measure of risk, of return is hence 0. At point B the expected return is based on equation (1) and the standard deviation is given by equation (3).

Points midway between A and B are the possible combinations of risk and return from holding the risky and risk-free asset. The return at any point on this line is given by equation (5) with the risk of these points being proportional to the risk of the risky invested portion:

(3.6) $$\sigma_c = y\ \sigma_p$$

The Capital Allocation Line

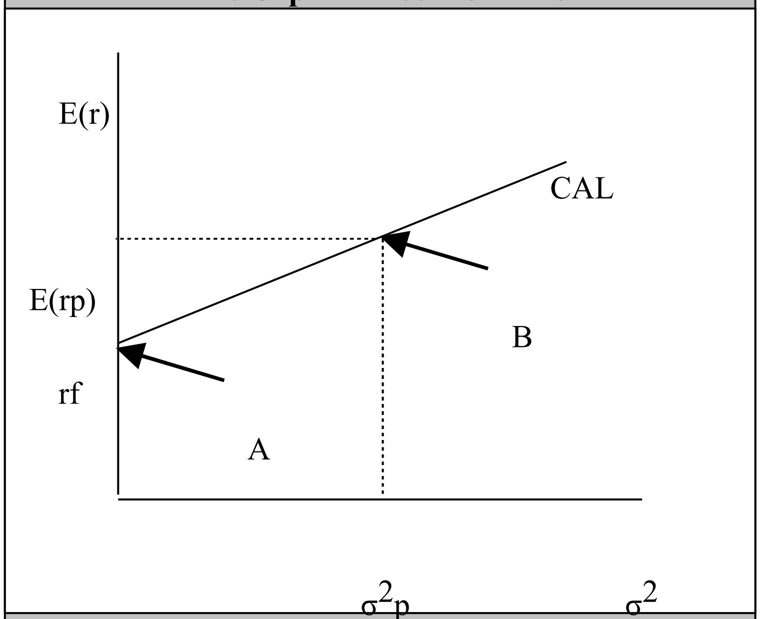

Figure 1

As the risk-free asset is assumed to have no variance of return, we see that the only contributor to risk will be that sourcing from the risky portion.

Diversifiable and nondiversifiable risk

We see that some risk can be eliminated through diversification. As a portfolio increases the securities contained, total risk (variance from the mean) should decrease as well. Risks can be attributed to specific influences. However, as not all of these influences will create variance in the same way for all securities, the addition of multiple securities can cause some degree of offsetting, and hence, stabilize returns.

Risk that is inherent in a specific firm is referred to as unique or firm-specific risk. It is nonsystematic risk; diversifiable through the addition of securities that offset its risk. However, we find that even if a portfolio is held of numerous securities, an element of risk remains.[36] This portion is known as market, or systematic risk. It is the portion that cannot be eliminated through diversification.

Covariance and correlation[37]

In phase I of CAPM development, it was assumed that portfolios are split between one risk-free asset, and one risky asset. With the distinction between diversifiable and nondiversifiable risk, we see that multiple risky assets can be held together, and that doing so can decrease risk simultaneously. The key to this concept of risk reduction is through the correlation of assets.

In table 1 we can compare the expected returns of two funds. The fist, a stock fund exhibits positively correlated returns with the strength of the economy, the second, a bond fund, exhibits a negative correlation.

		Rate of Return (%)	
Scenario	**Probability**	**Stock Fund**	**Bond Fund**
Recession	1/3	-17	17
Normal	1/3	12	7
Boom	1/3	28	-3

Table 1

In table 2, we have determined the applicable expected returns, variances and standard deviations..

[36] Elton and Gruber (1987) demonstrated the effects of risk and diversification on NYSE stocks. They found that there is little difference between a portfolio of 20 stocks, and one of 1000 regarding risk. In either case, each portfolio contained about 40% of the return variance (risk) that holding a single security would have. Hence, we can see through diversification that diversifiable risk is quickly eliminated, but that it ends quickly as well, leaving market risk remaining.

[37] The example of correlation and covariance computation is taken from Bodie, Kane and Marcus (2001: chapter 7).

	Stock Fund			Bond Fund		
Scenario	Rate of Return	Deviation from Expected Return	Squared Deviation	Rate of Return	Deviation from Expected Return	Squared Deviation
Recession	-7.00%	-18	324	17.00%	10	100
Normal	12.00%	1	7	7.00%	0	0
Boom	28.00%	17	-3	-3.00%	-10	100
Expected Return	11.00%			7.00%		
Variance	204.7			66.7		
Standard Deviation	14.30%			8.20%		
Table 2						

We find that not only do the funds' general returns vary in opposite direction based upon our expectations, but that also their general risk levels are dissimilar as well. The bond fund, as might be expected, is much more stable in return but a much lower return as well. This trade-off is the one expected by Markowitz (1952).

Now, given the same data, if we hold a portfolio comprised of the two funds (divided 50-50 equally), we can see what the expected portfolio results might look like in table 3.

		Rate of Return		
Scenario	Probability	Stock Fund	Bond Fund	Portfolio Return
Recession	1/3	-7.00%	17.00%	5.00%
Normal	1/3	12.00%	7.00%	9.50%
Boom	1/3	28.00%	-3.00%	12.50%
Expected Return		11.00%	7.00%	9.00%
Variance		204.7	66.7	9.5
Standard Deviation		14.30%	8.20%	3.10%
Table 3				

The most important thing to note is that while expected return is the simple average of the two assets, the expected standard deviation of the whole portfolio is lower than either fund could achieve alone. This is due to the inverse relationship that exists between these two funds, or, the negative correlation that they share. The important question that needs to be answered is how to measure this correlation that exists between different assets.

The first thing that needs to be determined is the degree of covariance between the two assets. This is the same as the variance looked at previously, but involves the deviation from each others'' returns instead of deviation from the risk-free rate. For this example we can see the computation in table 4.

	Stock Fund		**Bond Fund**		
Scenario	**Rate of Return**	**Deviation from Expected Return**	**Rate of Return**	**Deviation from Expected Return**	**Product of Deviations**
Recession	-7.00%	-18.00%	17.00%	10.00%	-180
Normal	12.00%	1.00%	7.00%	0	0
Boom	28.00%	17.00%	-3.00%	-10.00%	170
	Covariance = Average of product of deviations = 1/3 * (-180 + 0 – 170) = -116.67				
Table 4					

Notice that the deviation will be negative if one fund performs relatively poorly while the other performs well, and the deviation will be positive only if both assets perform, *well or poorly*, in the same direction. The average of these deviations is computed as the covariance, as shown in the bottom line of table 4.

However, covariance as a measure is difficult to understand in regards to our previously established risk or return measures – risk and return are percentages, covariance is a general number. This metric can be transformed into the correlation coefficient, through the following formula:

(3.7) $$\text{Correlation coefficient} = \rho = [\text{Covariance}] / [\sigma_{stock} * \sigma_{bond}]$$

Hence, in our example, with a covariance of -116.67, a correlation coefficient of -0.99 is yielded. This measure tells us how much the assets move together, and is always between -1 and +1. Two assets with perfect correlation would have a covariance of 1; two assets not correlated in any way would have a covariance of 0; and two assets perfectly negatively correlated would have a covariance of -1. In our example then, we see that the bond and stock funds are almost perfectly negatively correlated with a covariance of -0.99.

The data to determine these metrics was previously assumed to be subjectively derived probabilities. However, it is assumed possible to use historical data, and project this into the future. The idea is that as averages are used, these will change slowly over time – especially for correlation and covariance.

Three rules of portfolios with two risky assets.

If we assume that a proportion w_B is invested in the bond fund, and the remainder is invested in the stock fund ($1 - w_B$ or w_S), there are three ancillary rules which can be formulated about the covariances.

Rule 1 – portfolio return is a weighted average of the component returns:

(3.8) $$r_p = w_B r_B + w_S r_S$$

Rule 2 – the expected portfolio return is the weighted average of the expected returns of the components:

(3.9) $$E(r_p) = w_B E(r_B) + w_S E(r_S)$$

Rule 3 – the variance of the rate of return on the two-asset portfolio is given by:

(3.10) $$\sigma_p^2 = (w_B \sigma_B)^2 + (w_S \sigma_S)^2 + 2(w_B \sigma_B)(w_S \sigma_S)\rho_{BS}$$

The final term ρ_{BS} is the correlation coefficient between the stock and bond fund returns. The total variance of the portfolio is the sum of the individual components' variances plus a term that involves the correlation between the two funds. This final term is where we will see an increase, or reduction, in total portfolio risk if the component assets have a positive or negative correlation. Conversely, we see that only in the event where two assets are not correlated in any way ($\rho_{BS} = 0$) that the portfolio's overall variance will be the mere sum of its component parts. Furthermore, gains from diversification are essentially cost-free, as they allow higher returns, without an increase in any offsetting risk.

We can further see that return is a linear transformation – the result of the two weightings of the components. However, risk is must more complicated. The result is that there will be a continuum of risk-return trade-offs for any given portfolio, not all of which will be efficient.

The Efficient Frontier

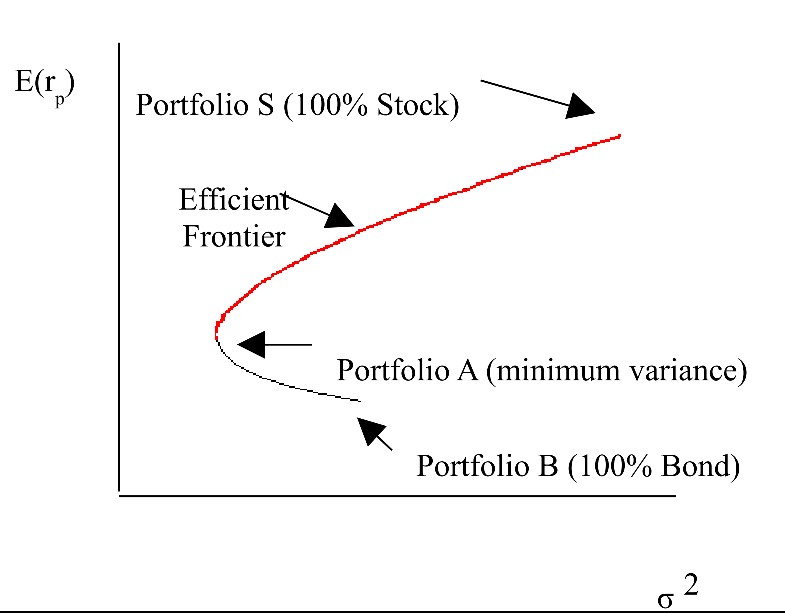

Figure 2

In figure 2, we can see that the risk-return trade-off will exist along the curve, however, this will not be linear. The advantages of diversification can also be noted. An efficient frontier exists above the minimum variance portfolio. An investment in a portfolio containing only bonds, which independently offer the lowest return, will continue to offer the lowest return, but at a risk level higher than could be achieved through diversification. Hence, we can see that assets can be combined to lower their diversifiable risk, and hence, reduce portfolio risk for a given return. The amount of risk an investor will take on will still be an individual matter, determined through their personal preferences and risk aversion.

Also note that an efficient frontier exists. All allocative mixes on the curve A - S will offer higher returns than those located on curve A – B at a lower risk, for every level of risk. Therefore, an investor would always choose to invest on the efficient frontier, and maximize their return for any risk level.

The optimal risky portfolio with a risk-free asset

The previous section looked at the trade-off that existed solely between two risky assets. However, following Tobin (1958), we see that we may split the investments between risky and risk-free assets.

The Efficient Frontier and the CAL

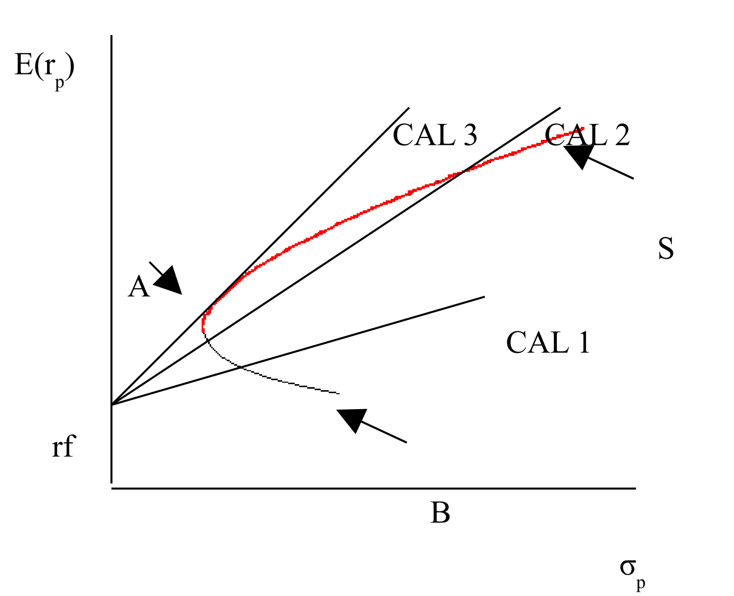

Figure 3

We will return to the capital allocation line from earlier, and see how this mixes with the portfolio opportunities in figure 3 above.

CAL 1 shows combinations of a risk-free asset, yielding r_f, and risky portfolios at a given level of S (reward-to-variability) ratio. CAL 2 represents a portfolio that invests more into stocks than bonds, which were more risky, resulting in a more steeply slopped S. However, we see that for any portfolio on CAL 2, there exists a more optimal combination, that is, at every level of risk we can create a portfolio that yields a higher return. This is given by portfolios on CAL 3. Therefore, we see that an optimal portfolio can be created where there is tangency between the efficient frontier, and the CAL. This is the mix of risky and risk-free assets that offers the highest reward-to-variability ratio, with the highest yield at the lowest risk.

The preferred portfolio will always be the one that maximizes the risk-reward ratio, given an individual's aversion to risk. We see than that all investors will invest along the CAL 3, as this represents the returns that maximize this ratio. The only difference will be the amount invested in the risk-free asset. This will tailor the risk-return trade-off to suit the individual's risk aversion needs. The same analysis can be offered for more than two assets, the only difference being that multiple efficient frontiers are created. The question remains as in our two-asset example, which efficient frontier and

CAL maximizes the expected risk-return trade-off. Conceptually, this may seem difficult, however, the advent of the computer age in the early 1960s did much to ease this process, and increase development of the concept.

We now see the full importance of Tobin's (1958) "separation theorem." Portfolio choice can be separated into two tasks: (1) determining the optimal risky portfolio (and the efficient frontier), and (2) designing a portfolio between these risky and risk-free assets to satisfy the individual's risk-reward preference. *In all cases, the optimal risky portfolio will be identical for all investors at any given time, the only differentiating factor will be the amount invested in a risk-free asset.*

Conclusion

Return is positively correlated to risk. Therefore, both the risk premium, and the standard deviation of the portfolio will increase as a higher proportion is invested in the risky asset. However, this increase in return will be constant.

An efficient frontier exists whose risk-return profile dominates all others that can be created through the allocation of two or more risky assets. This is due to the fact that returns are additive, but risk is non-linear. Hence, combining assets *may* yield a reduced return as the assets' covariance reduces overall risk. Some combinations will produce higher returns at all levels of risk.

The slope of the CAL is what defines an increase in portfolio risk and return. The slope is commonly referred to as the "reward-to-variability" ratio (S). As we see that the slope is given by rise over run, then it is defined as the risk premium divided by the standard deviation. Or, as shown below:

$$\text{Slope (S)} = [E(r_p) - r_f] / \sigma_p \tag{3.11}$$

Risk to reward combinations may differ, but the ratio will always remain a constant. This can be seen graphically in figure 1 as the linear relationship between the two variables. The CAL is the way we express the trade-offs between risk and return that can be achieved through two or more risky assets are mixed with a risk-free asset. Once an optimal efficient frontier has been established, all investors will share this preference to maximize their risk-return trade-off based upon it. The only differentiating factor at this point is, what percentage of assets will be invested in a risk-free asset to satisfy their risk aversion. This is an individual factor that will vary for each individual. Markowitz was thus able to provide two important insights into the risk-return world:

[1] diversification need not rely on perfectly uncorrelated assets, imperfectly correlated assets also reduce risk through diversification.
[2] Risk reduction is limited to the extent that returns are imperfectly correlated.

The results from phase I would teach us that variance is an adequate measure of risk for whole portfolios of assets. However, it is not particularly useful in describing individual assets that do not lie on the efficient frontier. As a result, at this point it was not possible to compare single risky assets with diversified portfolios. Therefore, there was a need to distinguish between portfolio risk and the contribution of the risk from a single risky asset to an already diversified portfolio (i.e., the market portfolio). It would take Sharpe (1963) to move the stage to the next level and show these results on a security by security basis, instead of solely in terms of portfolios.

Phase II

Sharpe (1963) would usher in the next stage of CAPM development. The simplification of Markowitz's search to identify efficient portfolios was offered. Instead of viewing risk as a factor of variance between all securities, risk was now reduced only to the variance a security exhibited against the market portfolio. The reduction of the computational complexity was viewed to be so important as to warrant its own section (section 7) in his original article. Sharpe also developed a computer program specifically to analyze his new model, which was able to calculate at about 2% the cost of comparable programs (*ibid.*: 277). The emphasis placed on computational efficiency for research is apparent, with a growing base of empirical results to derive theory from. The simplification, however, greatly expanded the scope of possibilities in both testing, and expanding upon the developments of phase I.

Sharpe (1964) would critically define the formal introduction of the capital asset pricing model to the world. In his opinion (*ibid.:* 427), the missing piece of the puzzle in previous models was that "none has yet attempted to extend it to construct a market equilibrium theory of asset prices under conditions of risk." He proceeded by starting with the same binary choice facing investors. Figure 4 shows the result of these contributions graphically, which although seemingly quite similar to those achieved by the end of phase I, were arrived at in an importantly different manner.

The Security Market Line

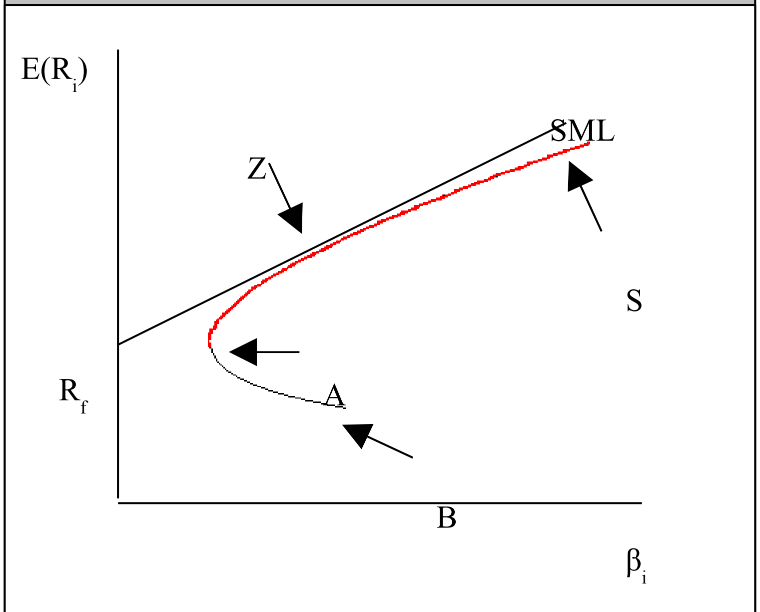

Figure 4

Curve *AS* is the minimum variance frontier, previously our efficient frontier, and demonstrates the different combinations of return and risk for a portfolio comprised solely of risky assets. Investors holding only risky asset would prefer combinations on curve *BAS* only above point *A*, thus maximizing return for a given amount of risk. These portfolios are known as mean-variance-efficient. Adding a risk-free asset results in the creation of the security market line (Lintner 1965b). If all assets are invested in a risk-free asset, the variance of return will be zero, and the return will be the risk-free rate. Combinations of risky assets and the risk-free portion will result in risk-return profiles along this linear line. To obtain the mean-variance-efficient portfolio available with the risk-free asset, the tangency at point *Z* demonstrates the resultant combination. Hence, all risk-free portfolios are a combination of a risk-free asset, and the single risky portfolio. This was previously the contribution of Tobin's (1958) "separation theorem."

The security market line rests on three main assumptions (Sinclair 1987: 27). The first is that all investor's utility functions are either quadratic or normally distributed. This would be an essential component for the mathematical construct of the curve, and follows from the previous assumption that investor expectations are homogeneous. Second, any portfolio on this line will have all diversifiable risks eliminated. Only non-diversifiable, or market-risk, will remain. Lastly, the market portfolio and risk-free asset will dominate the risky assets.

Hence, it now becomes clear that as all investors hold homogeneous expectations, they will all hold the same portfolio of risky assets. This gives rise to the now famous CAPM formula:

(3.12) $$E(R_i) = E(R_{ZM}) + [E(R_M) – E(R_{ZM})]\beta_{iM}, \text{ where } i = 1, ... , N.$$

And also that:

(3.13) $$\text{Market } \beta_{iM} = [\text{cov } (R_i, R_M)] / [\sigma^2(R_M)]$$

Given that:

$E(R_i)$ is the expected return on asset i.
$E(R_{ZM})$ is the expected return on assets that have market betas equal to zero.
$E(R_M)$ is the expected market return.
β_{iM} is the risk premium on asset i regarding the market. This also represents the slope of our linear function in figure 4, or more simply, the risk that each additional unit of asset i contributes to the market portfolio.

The last step in development of the Sharpe-Lintner CAPM, was the addition of the risk-free asset. As a risk-free asset is uncorrelated to the market return, its beta will be zero. Hence, the addition of such an asset would contribute nothing to the variance of the market return. When the assumption of risk-free borrowing and lending is utilized, we see the return on assets uncorrelated to the market, or $E(R_{ZM})$, must be equivalent to the risk-free rate, $R_{f.}$ The expected return on asset *i* is now given by the risk-free rate, plus a risk-premium, which is given by the beta (or additional unit of marginal risk) times the premium of the market portfolio over the risk-free rate. Hence, the equation is now altered to:

(3.14) $$E(R_i) = R_f + [E(R_M) – R_f]\beta_{iM}, \text{ where } i = 1, ... , N.$$

This formula defines the classic Sharpe-Lintner CAPM, critically assuming unlimited risk-free borrowing, or lending.[38] There were additional critical contributions during this period however. Lintner (1970) would later argue that the price of market risk is inversely proportional to the size of the market, under conditions of perfect competition. This was to be challenged as a general proposition by Budd and Litzenberger (1972). Additionally, all models of this period assumed homogeneous expectations on the part of investors. However, the addition of the beta concept as a formal price of risk, with the separation theorem earlier developed by Tobin (1958) would prove a combination crucial to the further development of asset pricing.[39]

Despite some flaws in the logic utilized, the general crux of the model would remain essentially untouched until Black (1972) ushered in phase three of the movement. At first glance, the results of phase II seem very similar to those of phase I. A closer look will reveal the differences that lay hidden, that contributed to the success of the CAPM as a pricing model.

Single-factor asset model

The computational difficulty inherent in phase I of CAPM development was in need of simplification. For instance, in a universe of 100 securities, there would be 100 expected returns to be established. Then, the covariances would all require calculation, which, for 100 securities would entail 4,950 separate covariances.[40] Sharpe's (1963) main contribution was to develop the "factor model" as a method to simplify the computational process and identify the specific sources of risk. Prior to this phase, risk was measured separately for each security regarding its variance with all other securities. Sharpe established that it is only the variance that a security adds to a market portfolio that is of concern to the investor; variance against other securities is diversifiable, and hence, return is non-dependent on it.

Hence, we may wish to separate the risk between firm-specific factors, which are diversifiable, and market factors, which are nondiversifiable. Sharpe's assumption was that one common factor is responsible for all the covariance of returns, with all other factors being attributable to firm-specific

---

[38] Jack Treynor would forward a manuscript in 1961 to Lintner outlining essentially the same basis for CAPM which he had independently formulated. Lintner would offer no positive feedback, but continued to forward the same conclusions several years later. The unpublished manuscript did receive notice from Sharpe (1964: 427) who noted in a footnote the existence of the ideas, and the paper, but that it was unfortunately unpublished. This would remain so until Korajczyk (1999) when Lintner was finally given the opportunity to have this paper formally released as a chapter in the book. See also Treynor (1962).

[39] Beta has come under increasing attack with a burgeoning wealth of studies demonstrating its lack of relevance in explaining returns. See Banz (1981), Basu (1983), Rosenberg (1985) and Chan, Hamao, and Lakonishok (1991) for empirical studies questioning the metric's predictive relevance.

[40] In reality, a universe of only 100 securities is quite small. Even in an unrealistically small universe of 1,000 securities, there would be 499,500 separate covariances in need of calculation!

factors. That is to say, we can reduce the market risk, or non-diversifiable risk, to one common measure – beta – denoted β_i.

$$(3.15) \qquad R_i = E(R_i) + \beta_i M + e_i$$

The term R_i represents the excess return that a stock exhibits over the risk-free rate, hence:

$$(3.16) \qquad R_i = r_i - r_f$$

The term $\beta_i M$ represents the effect of macroeconomic surprises. We can say that β_i is the responsiveness of a security to a shock, and that M is a measure of the economic surprise (M = 0 being no surprise). Finally, e_i is the result of firm-specific factors, or events. Both M and e_i are assumed to have *expected* values of 0, as they are the impact of unanticipated events. Therefore β_i becomes the responsiveness of a security *i* to macroeconomic, nondiversifiable events or risks.

Specification of a single-factor model

This model is of little use if we cannot specify the factor that specifically affects security returns. As a result, a proxy is used for the market return, and this is assumed to be the common factor causing nondiversifiable risk. A market portfolio would theoretically be an impossible construct (Roll 1977), due to the complex nature of assets involved that may not have prices available. However, for simplicity a market index proxy is used, such as the DJIA, or S&P 500.

With the use of this market proxy, returns are specified between macro (systematic, or nondiversifiable) and micro (firm-specific, or diversifiable) components. Thus the excess return of a security, R_i, is a function of three parts:

α_i – the excess return of a stock if the market's return is neutral ($R_M = 0$)
$\beta_i R_M$ – the component of return due to general market movements, hence, β_i is a stock's responsiveness to the market.
e_i – a measure of firm-specific risks, or, unexpected events.

We may now write the excess return of the stock over the market as:

(3.17) $$R_i = \alpha_i + \beta_i R_M + e_i$$

Now a security's return has been divided between its two specific sources of risk. Market, or nondiversifiable, risk is contained in the $\beta_i R_M$ component (with β_i representative of the form's sensitivity to this risk), and e_i representing the firm-specific, or diversifiable, portion of risk.

$\beta_i R_M$ and e_i are assumed to be independent of each other. The variance of a security's return would normally be given by:

(3.18) $$\text{Variance } (R_i) = \text{Variance } (\alpha_i + \beta_i R_M + e_i)$$

However, as the two sources of risk are assumed independent, we can break the variance into its component parts:

(3.19) $$\text{Variance } (R_i) = \text{Variance } (\beta_i R_M) + \text{Variance } (e_i)$$

(3.19a) $$= \text{Variance } (\beta_i^2 \sigma_M^2) + \sigma^2(e_i)$$

$$= \text{Systematic risk} + \text{firm-specific risk}$$

The α_i term is assumed to be constant, hence, it has no bearing on the risk of a security. We find then that variance in return is attributable to two parts. The uncertainty of the whole market represents a nondiversifiable risk. Hence, the security's variance is dependent on the market's volatility (σ_M^2), and its sensitivity to the volatility (β_i). Also, the security's variance is attributable to firm specific factor's, represented by (e_i). This portion is independent of the market's general performance.

Assumptions of the capital asset pricing model

The CAPM, in a nutshell, is a model that represents the rate of return required on a security to compensate for risk, as measured by beta. It basically rests upon six simplifying assumptions, although some of these are relaxed and altered in more complex versions. Ergo:

[1] Investors do no affect prices by their individual trades.

[2] All investors plan to hold a security for one identical duration holding period.

[3] Investors build their portfolio from a set of publicly available assets, i.e., stocks and bonds, and have unlimited risk-free borrowing or lending privileges.
[4] No taxes nor transaction costs.
[5] All investors construct efficient portfolios, that is, they are all rational mean-variance optimizers.
[6] All investors analyze securities the same way, by the same measures, and hold a homogeneous expectation of the world. Hence, all probability distributions concerning future events are identical. This generates a unique and optimal risky portfolio.

Given these six assumptions, there are four ancillary conclusions that are drawn.

[1] All investors choose to hold the market portfolio (M) that includes all assets from the security universe.[41]
[2] The market portfolio lies on the efficient frontier. It will also be the optimal risky portfolio, resulting in a location where the tangency of the capital allocation line (CAL) touches the efficient frontier. Hence, the capital market line (CML), that line that connects the risk-free rate with a market portfolio, is also the optimal CAL. All investors therefore hold M as their optimal portfolio, differing only in the portion held between it and the risk-free asset to satisfy their risk preference.[42]
[3] The risk premium on the market portfolio is proportional to the variance of the market portfolio and an investor's degree of risk-aversion, given as (A*).

(3.20) $$E(r_M) - r_f = A^* \sigma_M^2$$

[4] The risk premium on individual assets is proportional to the risk premium of the market portfolio (M), and the beta of the security on the market portfolio. The implication is that the market's return is the single factor of the security's market. Beta is the sensitivity of the security to general market movements.

If all investors now use the same market portfolio as the optimal risky portfolio, the CAL now becomes the CML. The market portfolio that is held, and the CML representative of it, is now the same

[41] As all investors are mean variance optimizers, we find this market portfolio mean variance efficient by definition.
[42] All investors would hold M as the optimal portfolio necessarily through the assumptions. As they all are mean-variance maximizers (assumption 5), looking at the same universe of securities (3), have identical time horizons (2), use the same analysis methods (4), it follows that they must arrive at the same optimal portfolio.

as the allocative options available, or the CAL. As all investors will hold this identical portfolio, the only question is at which price. Arbitrage guarantees that this will be an identical price for all, and hence, if all investors are willing to hold an identical risky portfolio, is must also be the market portfolio.

CAPM and individual securities

The insight of CAPM is that risk of an individual security is that which adds to the risk of the portfolio. Hence, risk is what concerns investors, and constrains their investment decisions. Nonsystematic risk (diversifiable) can be reduced, and almost eliminated, through diversification. Therefore, investors cannot demand a premium for this risk, as it can be removed. The result is that investors only demand risk for systematic (nondiversifiable) risk. As a result, a security's risk premium is proportional to its beta. As we know the market has a beta of one, we find that:

(3.21) $$[E(r_M) – r_{f]}/ 1 = [E(r_i) – r_f] / \beta_i$$

Or, rearranging equation (21), we find that the expected return for a stock is given by:

(3.21a) $$E(r_i) = r_f + \beta_i [E(r_M) - r_f]$$

Hence, a security's rate of return will be equivalent to the risk-free rate, plus its systematic risk (beta) times the risk premium of the portfolio. Equation (21a) represents the most familiar expression of CAPM.

The security market line

As beta represents the expected risk-return trade-off for a security, is can now be graphed as in figure 5. As the market is the measure that beta is relative to, the market beta will always be one. Therefore we see that the return on the SML given a beta equal to one will be the expected market return. Higher beta values represent returns above the market, and vice versa. The slope of the beta represents the sensitivity of a security to the market's risk.

The capital market line graphed the risk premiums of efficient portfolios, those portfolios that are comprised of the optimal risky market portfolio and the risk-free asset. The SML shown here graphs an *individual security's* risk profile. The appropriate measure of risk for a single security is not its overall standard deviation, but its beta. That is to say, only a security's contribution of risk to a portfolio will be the correct way of viewing this risk. The SML is valid for demonstrating the risk contribution of both portfolios and individual securities.

The Capital Asset Pricing Model

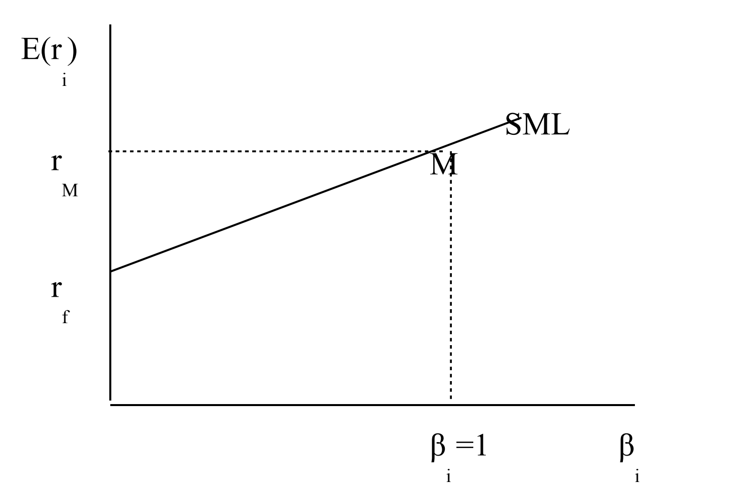

Figure 5

The SML therefore also provides a benchmark. As it demonstrates the return required for a given level of risk, it can show if a security of a given risk level, as measured by beta, is priced appropriately or not. Assets that are trading at their equilibrium "fair" prices are located on the SML, underpriced securities below it, and overpriced securities above it.[43] Therefore, any asset not trading on the SML will be in disequilibrium and will be repriced accordingly by arbitragers. Hence, the risk adjusted return of each asset will always lie on the SML in equilibrium.

The hallmark of this phase is Sharpe's insight that although standard deviation of returns is an appropriate measure of risk for an efficient portfolio, it is not an appropriate measure of risk for individual assets or for comparing the general riskiness of portfolios with the general riskiness of individual assets. Hence, the covariance a single risky asset has against a portfolio is the sole appropriate contribution of risk for a single asset. As a result, this covariance with a portfolio (or the market) is the only type of risk that investors will pay to avoid.

---

43 Remember from equation (17) that α_i is the measure of return differing from the neutral market return. Therefore, if a security is priced above or below the SML, the difference will be this alpha.

Phase III

Black (1970) would commence the third development phase by listing the four primary assumptions the previous Sharpe-Lintner CAPM were based on:

[1] All investors hold homogeneous expectations about end-of-period values for all assets.
[2] The probability distribution of all asset returns is assumed normal.
[3] Investors choose portfolios that maximize their expected wealth, and are risk averse
[4] An investor can be long or short any given asset, *including the risk-free asset.*

The assumption that an investor could borrow or lend any amount of money at the risk-free rate was the primary focus of Black's disdain. He would also take issue with the lack of time definition in the model. Empirically, he would posit that distributions varied wildly over time, hence this lack of a defined temporal element proved unsettling. Lintner (1969) had already demonstrated that giving heterogeneous expectations to agents did not materially affect the model. Black would also assume assumptions [2] and [3] to be reasonable approximations of reality.[44] However, he explicitly sought to correct the mistake he felt inherent in the fourth assumption, one he felt would substantially change the model.

Even at this early point of time in the model's life-cycle there was starting to be a growing body of empirical evidence pointing towards flaws implicit in it. Black would focus specifically on Pratt (1967), Friend and Blume (1970), Black, Jensen, and Scholes (1972) and Miller and Scholes (1972). Miller and Scholes in particular was troublesome as they set out to originally correct the model for the bias the two previous studies discovered whereby higher-risk portfolios were found not to exhibit above average returns over their low-risk counterparts. Even after Miller and Scholes' adjustment however, they were unable to rectify the seeming contradiction of a negative correlation between risk and performance. Given the growing body of evidence against the empirical validity of the CAPM, Black (1972) sought to rectify this through altering the fourth assumption.

Black would examine two cases. The first would assume that no risk-free asset was included in the portfolio, and hence no risk-free borrowing or lending. The second would have a risk-free asset

[44] In reality, returns cannot follow a strict, symmetric, normal distribution. Returns are always unlimited positively, but are limited to -100% negatively. Also, as Fama (1965a) discovered, empirical results based on New York Stock Exchange securities showed statistical fat-tails, with no finite variance. Fama (1965b) concluded that this these issues made no difference, provided the distribution is symmetric and stable over time.

available for lending purposes (long), but borrowing (or being short) would not be allowed. The assumption remained that an investor could take an unlimited long or short position in any of the risky assets. The result was that the risk-free asset made no difference to the portfolio. The same result can be obtained by allowing unrestricted short sales of risky assets. Hence, provided investors can sell short portfolios located on line *ab* in Figure 4, they can emulate the existence of the risk-free asset.

The relationship between the Black CAPM, and the Sharpe-Lintner CAPM, is found in the treatment of $E(R_{ZM})$. For Sharpe-Litner, this term must be a risk-free asset. For Black, it must only be less than the expected market return, or $E(R_M)$. However, as French and Fama (2004: 30) point out, the belief that unlimited short-selling of risky assets exists is as unrealistic as the belief of unlimited short-sales of risk-free assets. However, they also note that “all interesting models involve unrealistic simplifications, which is why they must be tested against data.”

During this period, attention was also drawn to the unrealistic assumption of homogeneous expectations. As Rabinovitch and Owen (1978: 585) point out, individuals possess partial information sets, and hence, also have heterogeneous expectations due to this incompleteness of information. Their main conclusion however would be that under equilibrium conditions, information has no value. Additional models were forwarded utilizing heterogeneous expectations, namely Lintner (1969), Fama (1971) and Gonedes (1976).

The CAPM was initially a static, single-period, model, although generally treated as though it applied inter-temporally as well. Merton (1973) attempted to create a temporal dimension. This aspect found that the equilibrium relationships found in the static model would also hold inter-temporally, but only under two special assumptions. These would be the expectation that agents maximize their utility, and that assets have limited liability.

Despite these additional modifications to the model, the essential CAPM to this day is still that based on Sharpe (1964), Lintner (1965b) and Black (1972). There are several key implications that can be derived from the model.

[1] It shows us what expected return is *not* based on. As firm-specific risk is diversifiable, return will be based on market risk, or that which is non-diversifiable. High return stocks will be so based on their sensitivity to the broader market.

[2] Beta is the measure of risk that *cannot* be diversified away. The market, and the average stock by definition, has a beta equal to one. On a graph with risk (as measured by beta) on the horizontal axis,

and return on the vertical axis, all securities lie on a single line, the securities market line (SML), as in figure 5. This is due to all risk being in relation to the market (non-diversifiable). In figures 1 through 3, individual securities could have lied scattered anywhere on the diagram, as risk was viewed as stand-alone, or diversifiable.

[3] Return varies only with beta. Other models, the Gordon-Shapiro growth model for example, rely on other metrics (i.e., future expected cash flow, dividend yield), which may be difficult to establish in advance. With CAPM, the only metric that needs to be forecast is the expected beta.

Despite these advantages to other methods, the model is increasingly facing attacks by empirical testing looking to concretely invalidate it. The following section will take a brief look at some of these challenges

3. Empirical evidence contra CAPM

Empirical tests of CAPM are based on three cornerstones of the risk-return trade-off: (1) that beta is non-zero, always implying that the market return is greater than the return of uncorrelated assets, (2) that expected returns are linearly related to betas, and (3) that assets uncorrelated to the market have expected returns equal to the risk-free rate, hence giving a beta premium equivalent to the expected market return less the risk-free rate. Douglas (1969) was the first individual to find empirical problems with the CAPM. He found that, based upon the model, nonsystematic risk did not statistically explain returns and also that the estimated SML is too shallow. That is to say, high beta stocks tend to exhibit negative alphas, and low beta stocks exhibit positive alphas.[45]

Two separate camps emerge when viewing the empirical failures of CAPM. The first are the behaviorists The focus here is that stocks with high book to market value ratios are stocks that have generally fallen into difficulties or are out-of-favor. Contrast these with low book to market value ratios that are typical of stocks for growth firms. Behaviorists argue that when viewed this way, it becomes evident that investors overreact to bad news, and results. Hence, betas are skewed accordingly by this general overreaction inherent in the treatment of negative news. Typical of this viewpoint is DeBondt and Thaler (1987) and Lakonishok, Shleifer and Vishny (1994).

The second camp represents those who believe that the CAPM is just not complicated enough to deal with the inherent complexities it sets out to. CAPM is restrained by many assumptions which may not be entirely realistic. For instance, the assumption that investors are concerned only with their risk-return trade-off may prove entirely unwarranted for most. If this is the case, beta may not be a completely descriptive measure of risk, and hence, would explain why it has performed so poorly under empirical testing.

Roll's (1977) famous argument that the fatal flaw in testing CAPM lies in its inherent testing impossibility continues to be a thorn in the model's promoters' sides. The problem is that the market portfolio that is the crux of the model is theoretically, and hence empirically, impossible to formulate. Which assets should be included in a market portfolio are legitimate, and the available data for these assets creates issues for the testability of the model. Roll's main argument is that as we can only use a proxy to test CAPM, we can never learn anything directly from it. Fama and French (1992) offered a scathing blow to CAPM by showing that once other factors (such as firm size, book to market value

[45] Miller and Scholes (1972) argued that there are statistical problems that hindered Douglas' test of CAPM. It becomes evident that anything can be rationalized, if the assumptions are set appropriately.

ratio, etc) are factored for, *beta explains none of a security's future returns*. Cochrane (1999) points out that most investors earn income not only from their investments, but also from employment or other sources. Their risk profile will therefore not be limited solely to their investment portfolios, but will also include other factors related to their other income sources.

Despite these failings, the CAPM continues to be a “theoretical tour de force” despite having no empirical support (Fama and French 2006). Academics continue teaching it, despite these acknowledged problems, as an introduction to portfolio pricing. Additionally, it is often built upon to create more complex models such as Merton (1973). The relatively simple risk-return concept continues to make this a popular teaching tool, despite misgivings held as to its true validity.

4. A Comparison of Models

CAPM as it is normally recognized today is the culmination of three models: Sharpe (1964), Lintner (1965b) and Black (1972). We have also seen that Treynor (1961) had independently developed near identical conclusions, although not having them published. A quick look into the similarities and differences will prove instructive.

Several key similarities come to light that make the "kernel" of the model. All four developers used the same assumptions concerning risk and utility. Risk was accepted by all as being best demonstrated through variance of returns, as Markowitz (1952) had formalized. Additionally, agents are seen as being risk averse. Furthermore, the Markowitz influence in the assumption of utility being a binary function, between risk and reward is evident. All developers viewed the choice for investors as being between greater risk, or higher returns.

In table 5 we can see a direct comparison as to where these similarities, and differences, lie.

Early CAPMs and Relative Assumptions

	Treynor (1962)	Sharpe (1964)	Lintner (1965b)	Black (1972)
No Taxes	Explicit	Implicit	Explicit	Implicit
No Transaction Costs	Explicit	Implicit	Explicit	Implicit
Agents as price takers	Explicit	Implicit	Explicit	Implicit
Agents are utility maximizers	Explicit	Explicit	Explicit	Explicit
Utility a function of risk/return	Explicit	Explicit	Explicit	Explicit
Variance is measure of risk	Explicit	Explicit	Explicit	Explicit
Agents are risk averse	Explicit	Explicit	Explicit	Explicit
Risk-free asset utilized	Explicit	Explicit	Explicit	None
Homogeneous expectations	Explicit	Explicit	Explicit	Explicit
Short sale allowed	Allowed	Disallowed	Allowed	Allowed
Time horizon constant	Explicit	Explicit	Implicit	Implicit

Table 5

Therefore, it followed naturally that investors would maximize the utility of this risk/return trade-off. The one additional simplifying assumption shared by all is that all agents have the same homogeneous expectation of the future.

The other key differences are largely of semantic nature. For instance, the treatment of taxes is either explicitly modeled, or implicitly assumed in all versions. This point may change a particular outcome of the CAPM, but the underlying logic behind it remains, and hence, the material conclusion will be essentially the same under any model with these assumptions.

Following Tobin (1958) we can see that the risk-free asset played a critical role in the CAPM. Black (1972) showed that a risk-free asset need not be assumed. This key difference is the final major theoretical progression in the model's development.

V. THE LASTING INFLUENCE ON FINANCE

Both EMH and CAPM form the cornerstones of modern finance theory. Indeed, their influence extends latently at least to most finance theories and security pricing models. While presenting Markowitz, Miller and Sharpe with their Nobel Prize in 1990, Assar Linbeck (as quoted in Mäler 1992) of the Royal Swedish Academy of Sciences commented:

> Before the 1950s, there was hardly any theory whatsoever for financial markets. A first pioneering contribution in the field was made by Harry Markowitz... The Capital Asset Pricing Model has become the backbone of modern price theory of financial markets.

If Fama had shown that past information was not a determinant of future performance through EMH, the developers of CAPM had shown that past information must be correct and can be used for pricing future values. These two points would become the stepping stones to the vast majority of future developments in the financial realm. In Rubinstein's (2002: 1044) own words, "the ideas in [Markowitz's] paper have become so interwoven into financial economics that they can no longer be disentangled."

These two concepts have been readily embraced by academics, despite some misgivings as to the finer points each has. However, practitioners have been slightly reluctant to fully utilize EMH, unlike the CAPM which enjoys wide industry success. Brunner, Eades, Harris and Higgins (1998) conducted a survey of leading financial advisors and investment companies and found that over 80% of them to be using CAPM. In the same survey, some firms commented on using other models, but these were in the small minority. Furthermore, the specific CAPM used was exclusively the original model (Sharpe, Lintner and Black), essentially untouched since 1972. EMH on the other hand has been a tougher sell for the investing public. The idea that no one can beat the market, as epitomized by its original conception, is not intuitive to practitioners who regularly see individuals earning above average returns, even over longer time horizons.

However, in utilizing CAPM as their primary pricing tool, practitioners are unwittingly using EMH. The idea of costless and perfect information, as espoused by traditional EMH theorizers, remains a carry-over assumption to most CAPM developers. This internal inconsistency will catch-up eventually with practitioners, who will then be in a position to adopt a new method.

In academic circles, Fama remains an icon. His work remains at the core of modern finance theory, and the spread of his empirical method is wide. He was awarded the *American Finance Association Award for Excellence in Finance in 2008*, and the association had this to say of his contributions:

> The empirical methods he has developed have become the starting points for every academic researcher, and many practitioners, working to bring coherence to financial data, and the substantive empirical findings in his papers have triggered tens of thousands of subsequent papers attempting to extend, quantify, or refute them. Indeed, in studies of citation counts, his name invariably leads the list, and by most metrics he is separated from the field by a large margin. In addition, his many doctoral students and junior coauthors have, in their turns, gone on to be intellectual leaders of our field.

Cochrane (1999: 36) sums the progression of financial economics nicely. He notes that in the mid-1980s, there were three "bedrocks" of the finance world: (1) CAPM was a good measure of risk, and explained why some securities returned a higher rate than others, (2) returns were unpredictable, like a coin flip, based on EMH, and (3) professional managers could not reliably outperform the market once returns were adjusted for beta. The climate has changed now and these point have been replaced with: (1) some asset returns cannot be explained by beta, (2) returns are somewhat predictable, and (3) some funds have consistently beaten a passive strategy and cannot be explained through increased risk exposure.

To combat these new "bedrocks", several adjustments have been made. Multifactor extensions of CAPM have been developed to create additional betas to explain returns (Ross 1976). These models are called arbitrage pricing models and rely on the identification of several separate betas to determine the exact causes of price changes. The Fama and French (1993) Three Factor Model would be a good example of this creation, where beta is broken down into smaller parts to help further identify the source of return. EMH is being modified slowly to allow for anomalies. Krueger and Rahbar (1995) have developed a variable beta CAPM, to manage the effects of changing preferences over time. The work done on asymmetric information has done much to bring to light the need for such changes. Academics are warming up to the empirical fact that there seem to be some reliable indicators for future performance, such as price to book value. Finally, as practitioners have long known, some people do consistently outperform the market.

All these “attacks” on the traditional theories have not been as poignant as one could hope for. They all lack in the same regard, questioning the crux of these theories. Fama's contribution based on perfect information still implicitly remains in most academic literature. Markowitz's trade-off between risk and return is still treated as scripture. These two ideas are assumed to be unassailable to modern finance theory.

However, we can also see these two ideas are, at their cores, mutually incompatible. CAPM, to be effective as anything but an historical tool, must be concerned solely with future values. EMH tells us that future values must be approximately correct for the given information of the present. Therefore, CAPM must always yield an expected future value, *ex ante*, approximately identical to the present. If one believes in EMH, they must also admit that it serves no value than confirming what is occurring in the present. If one believes in CAPM, they must not accept EMH's prime tenet.

The result of these two separate theories has been that three core elements engulf the present world of financial theorizing:

[1] Prices can be represented as random walks.
[2] Markets exhibit a risk-return trade-off, where risk is rewarded with an increased return.
[3] There is no possibility of earning more than the risk-adjusted rate of return by design.

Furthermore, the field has become awash in mathematical analysis, unable to comprehend the true subjective nature the financial realm exists in. Rubinstein (2002: 1044) may have summed up Markowitz's contribution most aptly:

> Near the end of his reign in 14AD, the Roman emperor Augustus could boast that he had found Rome a city of brick, and left it a city of marble. Markowitz can boast that he found the field of finance awash in the imprecision of English and left it with the scientific precision and insight made possible only by mathematics.

There is no doubt the financial realm is more precise now than it was 40 years past. We cannot state whether this has truly made it any more scientific or not. The Roman empire may have spanned its widest less than 100 years after Augustus remarked those famous words, before falling into centuries of decline. We cannot state now if the same fate will befall the dominance of these early ideas of finance. What follows will hopefully contribute to, what no doubt will be a slow progression, of the replacement of the existing paradigm with a more fundamentally sound foundation to grow upon.

Book II

Theoretical Problems with the Models

I. COMMON CONCERNS

The treatment of both these models should be separate. However, we find that four of their building blocks are found equally erroneous and can be treated as such.

The concept of time is mistreated in CAPM, and removed completely from EMH. The temporal element is what unites all actors throughout society; its presence is inescapable. Therefore, we see that the repercussions are important for the rest of the theories utilizing some conception of time. In fact, many of the theoretical flaws that are found to follow are a direct result of this temporal mistreatment. EMH assumes that time does not exist in a strict sense, instead, it sees all action as occurring in a static state. Information for example is spread effortlessly without a temporal element. CAPM, in distinction, treats time explicitly. However, its definition of time is inapplicable for the realm of human action. Newtonian time may be able to describe many physical phenomena, but not the actions of humans. Instead, we see that time for humans is a dynamic flux, with an emphasis on the past in creating it own future flow and importance. The Bergsonian viewpoint of time is forwarded with the results for CAPM offered.

Risk and uncertainty are both misplaced ideas in EMH and CAPM. The idea of risk is only suitable to a world without human action. It assumes a closed-end system, where all future eventualities are knowable, if not known, in advance. However, we see that the world that concerns us is more dynamic. An element of Knightian fog exists that makes the future fundamentally unknown in advance. Instead, the world of human action, and any finance theory that aims to describe realistically this world, must make use of the correct concept of future uncertainty. Future events exist only on a case by case basis, hence, the use of statistical measures to objectively describe and mitigate risk is denied.

Instead, the mitigation of this temporal risk is attributed to the entrepreneur, an element that receives zero treatment in either theory. This is due as well to the lack of the correct view of time. As the entrepreneur is the force that envisions the future, and moves us into that future state, we see the importance in looking at the formation of prices. Furthermore, we see that the Knightian uncertainty that rules the future is not wholly indeterminate. Instead, the alertness and foresight of the entrepreneur navigates through this fog and brings a coherence to the market.

The most devastating problem however, is found in method. A purely empirical and mathematical approach has been utilized to the detriment of solid theory. The liberal use of these approaches has meant that little room has been made for concepts that are not able to be incorporated into them – specifically, entrepreneurship and Bergsonian time. The concepts that are unable to be modeled or observed have been ignored. Also, in their replacement, wholly inapplicable concepts have been introduced. The use of statistics as a measure of this unknowable and unknown future is criticized accordingly. Only a solid theoretical foundation, based on deductive logic, can be utilized to create the foundation needed to build a science.

Finance always lacked the proper methodological foundation necessary for solid advancement. That is why today there is so much discontent with many of its ideas. However, financial economists lack the necessary tools to effectively remove these flawed ideas, and introduce more applicable concepts. A methodological shift towards a priori deductive logic would do much to rectify this problem and place finance on a complete foundation. To the extent that this one shift would also create the opportunity to correct the previously discussed three flaws as well, we see this one movement as the single most important step facing financial economists today.

1. Time

When we look at time, it is easy to think of it as being a homogeneous entity. One minute today is equivalent to a minute of tomorrow, or a minute of next year. As these are objectively measurable identical units, it follows that the utility one can benefit from a unit of time must be the same. Hence, all time must be objectively valued identically. The world of CAPM treats time in this manner. Returns are equated to risk levels over time periods that are deemed homogeneous. Market returns at one period of time can be compared to portfolio returns of a different period.

In the real world, we can see that all units of time are not created equally. Time is not some homogeneous entity, immune from human valuation. Instead, we see that it is subject to the same forces as any other entity in scarce supply. Hence, the valuation of time is specific for the individual, and subject to change throughout the passage of time. *Time alters our valuation on time*.

Time is the one concept that unites all actors together. Other inputs may vary in degree and duration, but the existence of time is objectively requisite for every action. However, the assumption that interpersonal comparisons of time are possible is fallacious, and yields troublesome implications for the rest of the CAPM. Once we see the true nature of time, its own time-variantness creates complications for formulae based on its supposed constancy.

Also, there is a significant difference between the conceptions of time *ex ante* and *ex post*. *Ex post* we can see how much an individual valued their time. *Ex ante*, this is an impossibility. Furthermore, if we take Shackle's (1994) viewpoint of imagination being the creator of the future, we can see a substantial difficulty in trying to determine the value that will be placed on future time. If our future is created through our imaginations, then the time that comprises the future, although a certainty in passing, will not be valued to anyone other than the individual actor.

CAPM uses time as a factor for reducing rates of return to a comparable standard. This is erroneous. To the extent that time is the result of independent, subjective valuation on the part of the actor, we can see that a comparison of it is never possible. We can draw conclusions from it historically, but this will always be clouded by the historical issue of judgment. A true asset pricing formula must recognize this aspect of time and account for it accordingly.

EMH fails to recognize time in any meaningful explicit manner. Instead, it confounds the idea, using three separate conceptions, none of which is applicable. By assuming that time fails to exist, exists for a finite time, and also that time exists distinct of itself. The results are processes that occur without a proper temporal passing, and the omission of the key events and eventualities that result from this element. This concept requires recognition, and repair, to place asset pricing on more solid ground.

The Nature of Time

Time exists continually. All action must take place in both space and time; these are the factors that unite all actors together in an abstract sense. The passing of time is, however, a constant that all must share equally in – the passage of time becomes the only constant in action. Time in a static sense may be similarly conceived by all. However, we can see that static time is not really time at all. Its mere existence implies the passing of time – it cannot exist any other way. When actors view time in the dynamic sense, we can see that its definition can vary. The concept of time therefore is not necessarily constant. We can see there exist two different methods we can view the passage of time.

Typically, time is viewed in a very *Newtonian* sense. The passage of time could be viewed as existing on a one-dimensional point, moving forward through time. The passage of time is therefore a very linear progression from one point to the next. Viewed this way, one unit of time is equivalent to any other. The minute that it took to boil a gallon of water last year is equivalent to the minute that it will take now, and the minute it will wake one year from now. We can see the conclusion to be drawn then is that what was true of time at a single point in time must be true for all points. *The value inherent in time is constant – it is time-invariant*.

The second definition of time stems from Bergson's "duration" - Bergsonian time. Time in this sense is mobile, or fluid. Time represents not a measurable entity; its continual flux implies that by the time an attempt to measure its existence is undertaken, it will already have passed. Hence, instead of existing on a one-dimensional line, we can think of this Bergsonian time as existing outside of a measurable area. Time exists, but not in a physical space. *Time exists only in the passing of time*. As Mises (1949: 100) would describe this process:

> The 'now' of the present is continually shifted to the past and is retained in the memory only. Reflecting about the past, say the philosophers, man becomes aware of time... Time as we measure it by various mechanical devices is always past, and time as the philosophers use this concept is always either past or future. The present is, from these aspects, nothing but an ideal boundary line separating the past from the future.

Time represents then a purely immeasurable process, its existence negates attempts at concretely defining and directly comparing its units. We know time exists however as everything we do occurs in time, its existence envelops us. What gives actors knowledge of the existence of Bergsonian time is our consciousness of its passing. This is what gives humans a time-variant aspect that differentiates us from the physical world. As Lachmann (1977a: 85) pointed out, "[t]ime and knowledge belong together." Furthermore, as Shackle (1972: 156) showed us, "[s]o far as men are concerned, *being* consists in continual and endless fresh *knowing*." We see therefore that for actors, time and being are inseparable.[46]

A curious point concerning time is the irreversible nature thereof. Time exists, but can only progress from one point temporally prior to it to another point more temporally distant. It is impossible to revert the passage of time. We see a distinct separation between physical entities, whose physical states can be reverted, and timely action, whose existence will forever be irreversible. What occurs in a moment of time remains in that moment forever.[47]

Additionally, like physical goods that actors economize, time is finite. Its existence can, in fact, never be increased in a strict sense. It may be argued that a person may be near death from a heart attack, and receives an operation that not only saves their life, it actually extends it a little. The nature of time remains – it is always a declining balance. Furthermore, we may think we can extend time's availability, but we can never be assured this to be true. When we purchase an additional good, a car perhaps, we know that its purchase will increase our present quantity by its objective quantity. The passage of time implies an uncertainty.[48] We can therefore never be sure that any action in the present will increase our time available in the future. Unlike the purchase of a physical good, the car, we can never purchase time with certainty that we will be able to derive utility from it.

Lastly, due to the finite amount of time available to actors, time is never viewed in an infinite way. Instead, actors always act for a definite (if ill-defined) amount of time. One cannot commence an action thinking it will last forever. The concept of time permeates every undertaking of action. For example, it may be said that a student freshly finished University, and looking for their first job will view this coming lifetime of employment as lasting *forever*. However, we can see that this may be an abstract manner of looking at the future events. Instead, the graduate will be knowingly searching for

---

[46] See also Mises (1949: 99) that time is a praxeologic category of action. Action necessarily implies the passing of time.

[47] Accordingly, Jevons' famous adage may have been "bygones are forever bygones," however, as we shall see shortly, bygones have a significant bearing on the present we experience, and the future we expect.

[48] Garrison (1984) points out that the opposite need not necessarily be true. Uncertain situations need not be caused solely by time, hence, although time breeds uncertainty, uncertainty is a poor proxy for time.

employment until retirement, or some other future *point* in time. The amount of time may be unknown in any manner of exactitude, and it will invariably be subject to change. However, that the action is initially undertaken with respect to the expected duration of it is undeniable.

The nature of time may have best been summed by Shackle (1958: 13) thus:

> It is natural for the economist to think of time as what the mathematicians call a space, or as one dimension of such a space, wherein distinct points do not, a priori, differ from each other in their general essential nature and properties: in such a space, what I will call a "pure" space, one point is, so to speak, as good as another. Yet plainly time is not of this character, when viewed from the human standpoint. For even if infinitely many distinct points of brief elements of time are thought of as co-existing, even if, that is to say, the whole stretch of history, past and to come, is looked upon as uniformly real, yet in the experience of human individuals each of these moments is in a certain sense, solitary. There is for us a moment-in-being, which is the locus of every actual sense-experience, every thought, feeling, decision and action.

Individualism and Time

We have seen then that time is not a constant progression, but is highly conditioned by our individual experiences of the past. At this point then, a simple example may serve to illustrate the point.

Assume that on the first of a month, a student is told they have a test on the last day of the month. They thus have 30 days in which to study. The student may choose to not study during the first 29 days. They have assigned a subjective valuation on those days which is relatively lower than the other acts they have chosen to perform instead. On the 30th day, the student awakens to the thought of the impending test that will take place in 24 hours. They will feel the pressure of the event, and place an increased emphasis on studying in this final available day. They choose to study at the expense of all other action, including eating. At the end of the day, the student is still not confident that they have studied enough. They wish for one more additional day to utilize for studying. They value an extra day of time highly.

Hence, we can see that the value actors place on time is not constant. It is highly conditioned by the circumstances, as well as the amount at their disposal. We can make two generalizations regarding the value of time.

The first is that actors will experience diminishing marginal utility from time as it concerns a given end. In our example above, the units of time experienced temporally distant from the expected end (the exam) were valued less highly than those located temporally closer. We can see that, although time is not homogeneous, it is still valued as a progression of a singular entity. This point may require further explanation. The actor in the present knows not what the future will bring, except that the end will arrive at an, exact or not, specific time. The time of the future until this point may be, thence, judged homogeneously. The only difference to the actor in the present is the sheer quantity of available time. One month must pass before the exam is written at the onset, and this is reduced as the exam date nears. Hence, early in the decision or action, there may be a large amount of homogeneous time expected to be available. This diminishes with the progression of time. The corollary of this is that the actor experiences increasing marginal utility with this passing. As time passes, the supply expected to be available decreases, and hence, the marginal utility of additional units of time increase. The value an actor places on these continually quantitatively reduced pieces of time increases accordingly.

The second point is that humans will experience a general change in their marginal utility from time. Each human will, in each specific case, have an expected total amount of time at their disposal.[49] This total amount will set an expected limit on the total amount of action. Take two points in one person's life: (1) on their thirtieth birthday, and (2) on their eightieth birthday. Further, assume that on each birthday, the individual assumes they will live to the age of eighty-eight. On their thirtieth birthday, they will see that they have an expected amount of time remaining (their remaining life) of fifty-eight years. On their later birthday, they may realize that they only have eight years remaining. Ceteris paribus, what will be the result? We can see that each moment in time will be treated less highly for the individual who has more time expected to elapse in the future.

The experiences of time will condition each individual to have a distinct viewpoint on their expected remaining time. In two broad ways, we can see that the value placed on time in the present is a function of the total expected remaining time an actor has at their disposal in its totality, and how much is expected to remain for a given end. That the second is forever a subset of the first is obvious. However, we can also see that as expected time diminishes, the spread between the value placed on the

[49] The total amount of expected time will forever be conditioned by the length of our own life we expect is remaining at any point in time. However, we also see that their can be smaller 'absolute' expected limits on time as well. For instance, if a person is only on vacation for two weeks, we see that this will set the origin for the inevitable declining balance of time available in the future during this period. No specific end will condition this value in time, but instead the fact that there is a set amount of time for this specific part of a person's life will condition the value they place on time specifically during this finite time period.

two will converge. The value an individual places on their total expected time remaining will be approached by the value placed on that time spent seeking individual ends.

Specific Determinants of Time-Value

Previously we looked at how, in a general sense, value is determined by the amount expected to remain. If everyone shared the exact same expected future duration, we may be inclined to think that all actors would value time in an identical manner. Those, for instance, that have thirty years of life left would place a value on those thirty years identical to others. This value would be identical if we held everything constant, including the expected future time remaining. We see however that the real world is not a product of this condition.

In the general sense we saw that the value placed on time is always a product of the total amount expected to remain. The general value is a meaningless, abstract concept however, with no direct bearing to individual actors. No actor takes an action in light of the total expected duration. Instead, they always act in regard to a specific goal. Simultaneously, we see that an individual seeks multiple needs in the same general time-frame, if not at the exact same moment.

Suppose now that the student from our previous example had not just one test to study for, but two (both at the same approximate future date, though they need not necessarily be so). Let us further assume the subject matter is mutually exclusive, so that the time dedicated to the attainment of one end is wholly separate of the time dedicated to the attainment of the other. The supply of time available for both ends is identical, however, we know that value is also determined through demand. Suppose that one exam is in a subject the student understands well, and the other concerns an obscure subject. The obscure subject will require relatively more time spent dedicated to it than the one they are well-versed in. We see that, regarding the attainment of these two *individual ends*, the value of expected time will be significantly different. The student's demand for time is wildly different in each of the two cases. *Hence, we see that for one specific end, the student may value the exact same Newtonian time wildly different than that regarding the other end*.

It is possible, indeed foreseeable, that distinct individuals will have the same available supply of time regarding similar ends. However, it is unlikely that any two individuals will ever have the same demand to use this time – why is this? Two specific factors mark this distinction.

The first is that people will value ends differently. If we use our previous example again, even with two students writing the same exam on the same day, and assuming the same level of knowledge of the subject, it may be that they will place a different emphasis on the actual exam. Hence, for one student that values it greater, it may be that they would prefer to have more time left available to study. The value of the ultimate end conditions the value on the specific portion of time. Individuals have a multitude of differing preferences, this fact will give rise to different values placed on means to achieve them, *which time must be considered part of*.

Secondly, we see that the passage of time itself implies differing values placed on future events. The passage of time will never be identical for any two individuals. As Lachmann already demonstrated, time is an essential ingredient in the knowledge process. As our expectation of the future will always be conditioned by our knowledge of the present, we see that the previous passage of time implies that no expectations will necessarily be identical. To return to our example, in the past, one of these students may have had the same professor for a different class. This will be the first exam administered by this professor for the second student. It could be that the professor has in the past given easy exams, and that the first student knows this through experience. We can see that as a result of this history, the two students may place differing values on the means (which time is one of) used to achieve a given end, *even if they value the end identically*.[50]

Time is therefore, always valued conditionally on the total amount the individual expects they have remaining. However, time to be used for specific ends is the problem that concerns us. We have seen that all action is directed towards a specific end, and that the value placed on this end, and the available means needed to achieve it, will determine the value we place on the time thereof. No two individuals will place the same value on time necessarily, although it is possible that they could. Instead, given the plethora of variables that determine value, ends, means and preferences, we can see that it is more likely that no two individuals will ever value the same moment in time identically. As Shackle (*ibid.*: 13) points out, "[t]he moment-in-being is always individual and distinct but always evolving as it rolls forward." Actors may share the general concept of the passage of time in common, but the sole similarity lies in the "rolling forward" of time that Shackle notes. The "evolution" is a factor that forever creates individual determinants of time's value.

A Note on the Temporal Perspective of Time

[50] Stigler (1961: 216) would note, while analyzing search costs of information, that income level would have a determinant level in the value placed on time by an individual. Individuals with disposable income for this end would value their time more highly than money, relative to an individual with less disposable income available. However, we see that this is only true in strict *nominal terms*.

It was previously noted that time, in an *ex ante* sense, *may* be valued homogeneously. Our student with an exam on the thirtieth of the month may value, on the first of the month, the time of each of the next twenty-nine days identically. Time implies an element of uncertainty. As such, there are variables that the actor cannot foresee at the time when a decision is made. The temporal spread on the value placed on times occurring at relative temporal points will be affected by this uncertainty.

For instance, our student in the present may view their mundane life as being quite constant. No significant changes occur at any specific period, and there is no reason to believe that any will happen in the coming month during which the exam approaches. As a result, *ex ante*, they may decide to assign each projected portion of time the same value regarding the same end. Hence, one hour will be set aside each day for study for the exam.

As time itself passes however, we see that the value placed on it will change. Suppose that on day two, the student has an emergency, and can't study the hour they expected they would on the initial day. The result may be that the value placed on the value of time in the remaining days will be affected accordingly. Each future hour may take on an increased significance, as they will be used to counter the missed hour from day two. The alteration in time could be evenly spread across the time horizon, thus keeping each individual unit of time expected to be utilized for the end identical.

Also, we see that regarding the given end of having much knowledge available for the exam, that time temporally closer to the actual exam will be valued more highly. Knowledge gained closer to the actual time the end is expected to be realized will be fresher in the memory, and more readily accessible. Hence, the student may place a greater value on an hour of study time the day before the exam, than they will a month before the exam. Time, like all means, is valued according to its expected serviceableness in attaining an end.

It follows that *ex ante*, it is entirely possible for time to be valued identically provided three conditions are met: (1) the time applies to the same person, (2) the end is identical for all periods of time, and (3) each period of time will share a common use for the individual. This value will be subject to change as time passes, but *ex ante*, it may be valued identically.

Ex post we can see how much an individually actually did value the time. This can be achieved through observation of their actions. The usage of time can dictate the relative value that was placed on a specific period of time. The implication of this is that true value can only be assessed after the passage of time, never prior to it. It is only through demonstrated preference that we can objectively establish the value an individual places on time.

Shackle (*ibid.*: 15) noted that, "[t]he moment-in-being rolls, as it were, along the calender-axis, and thus ever transports us willy-nilly to fresh temporal viewpoints." This I shall call the dynamic movement of time. All actual transformations of one situation into another are affected, we may say, by this dynamic of actual translation of the "present moment." Huerta de Soto (2005: 46) might describe this process the best, writing:

> According to this subjective notion of time, the actor perceives and experiences its passage as he acts; that is, as he creates, discovers, or simply becomes aware of new ends and means... In this way, the past experiences stored in the actor's memory continuously fuse in his mind with his simultaneous, creative view of the future in the form of mental images or *expectations*. The future is never determined, but instead the actor imagines and creates is step by step.

An important occurrence manifests continually in the temporal present. Our valuation of time changes, and we are allowed the opportunity to concretely value the fleeting moment of time relative to other past portions of time, and our expectations of the future moments of time we will experience. The *ex post* viewpoint opens up the possibility to compare ordinally the values we have placed on periods of time, a possibility that we cannot undertake in the temporal present regarding our yet unknown future.

Implications for Finance Theory

Both EMH and CAPM rely on distinct viewpoints of time. For CAPM, the usage of Newtonian time is apparent, with its applicable failings. EMH in distinction downplays the role of time in action. As a result, it too suffers a grave theoretical deficiency. The following two sections will look at each in turn.

Time and EMH

EMH exists in a static world. The passage of time is removed with one minor exception – information is created at a point in time – this is the only point at which a concept of time enters the theory. This neglect brings several grave deficiencies. First, the role of time in information dispersal is overlooked. Second, we find that under EMH, only new information has the ability to affect current prices (Fama 1970). However, the passage of time on its own can significantly vary actors' subjective valuations.

Third, EMH neglects the passage of time as an element in the creation of information *requiring no exogenous input*. Furthermore, EMH assumes that opportunities will all become fully exploited over time. This assumes that there is a finite amount of time, after which no new opportunities will become apparent. Bachelier (1900) and Osbourne (1959) both contributed to the pre-history of EMH by positing that information created at a time is independent of the influence of previous time. This lead to the central tenant of EMH that price movements in one time period of temporally unrelated to movements from separate time periods (Fama 1965b: 34). However, as we have seen, a significant value is placed on time that is wholly attributable to the past experiences of an individual. These five points all serve to defeat the dynamic use of the hypothesis.

Time and information dispersal

Information is assumed under EMH to be spread costlessly, and more importantly, timelessly to all relevant actors. This implies that any changes that will result from this new information will occur instantaneously on the market place. We see however that this is completely erroneous. Time is an element of all action. To make a separation would be to negate the process of action. As such, we see that time plays a significant role in the creation, and dispersal of information.

The creation of physical information occurs in time through the process of action. The dissipation of information is also a time-consuming process. As an action itself, it could not occur in any other way. The assumption that information is spread in a timeless manner masks some very important implications of this process.

The temporal aspect implies an element of uncertainty. When we view information as being dispersed throughout time, we see that there is no way to know *ex ante* what future events may occur or their bearing on the information in question. For instance, new information developed may make the current information irrelevant. Or, conversely, new information may develop making current information relevant. Under conditions of Knightian uncertainty, it is impossible to say what the future will hold for the relevance of a given piece of information. We can say, however, that this uncertainty implies that whatever importance a piece of information may have in the present, it will be altered over the passage of time. The fact that information must encounter a temporal element for dispersion implies that actors can never know the full importance of information in the present.

Time and endogenous changes in valuation

Fama (*ibid.*) concluded that only new information was able to influence prices. Two points become apparent here.

First, as was seen in section above, new information may not be that which was created in the temporal present. Much information exists that has not been accounted for by actors. As it is indeterminate as to when an existing piece of information will become "discovered", we find that there is no way to state that only newly produced information will be the sole determinant of prices. In fact, in light of this, there is no fundamental difference between one previously produced piece of information, and a fresh piece. What matters is not when the information is produced, but when it reaches the minds of the acting individual. This point in time is not known in any degree of exactitude in advance.[51]

Second, we see that the passage of time on its own endogenously changes valuations. As we looked at previously, the overall value a person places on time will be conditioned by the amount they have available. As a result, the means that an individual utilizes will also be valued accordingly. Take for example, a situation with no changes in variables, except the passing of time for an individual. As an individual ages, their income requirements alter. Normally, individuals toil their young lives so that they can save income to be utilized during their retirement. The passage of time alone changes the means individuals use to achieve their ends. An individual of a young age is thus a demander of investment vehicles. They require securities to place their savings into. As the passage of time brings them to their retirement years, they start removing demand from these securities as they stop contributing to their retirement savings accounts, and eventually withdrawal therefrom. *Time endogenously creates changes to the valuation of distinct ends*.

Time alone creates new information

Time on its own creates an additional informational input, regardless of the external environment. Information is never constant, but is constantly created throughout the mere passage of time. The idea that all the information could ever be collaborated in any meaningful way overestimates the abilities of actors to understand information, or underestimates the ability of time to create information.

EMH presupposes that it is possible in the static sense to know all the relevant information applicable to the *correct* pricing of an asset. Much like Hayek showed us however, the entirety of

[51] In an early empirical study of random-walk movements, Godfrey, Granger and Morgenstern (1964: 16) noted that a great deal of price changes observed occurred while financial markets were closed. What is important is not the point when the information occurs, nor when the financial markets receive it, but when the acting individual attains the knowledge. This temporal aspect may not conform to traditional trading hours.

information can never be known at any given time. Also, as Shackle (1972) previously demonstrated, the passage of time on its own results in additional "knowing" for an actor. The passage of time is therefore inseparable from the creation of new information. This constant flood of new information creates the impossibility of any one individual being able to fully attain, and properly utilize the information available, *in a dynamic sense*.

Time is infinite, not finite in duration

EMH also presumes that there may exist profit opportunities existing from the incorrect interpretation of information, but that a period of time exists in the future whereby these opportunities will be fully exploited, and hence the disappearance of any informational asymmetry.

However, the flaw with this way of thinking is in the viewpoint that time is finite. EMH, in this regard, acknowledges the passage of time to reduce informational asymmetries that exist. It does so by assuming that after a finite amount of time, all asymmetries will be removed, and hence, asset prices will fully reflect the available information. Time, it must be remembered, can be considered as infinite in a dynamic sense regarding the general concept of action. Although very much limited in a scarce way by individual actors' availabilities, the constant creation of humans creates a new supply of time in the market place. As Tirole (1985) cast light upon, the market place is constantly inhabited by individuals with fresh time scales. This continual influx implies that there is an endless continuum of time available. We can, in this way, view the market as a location that exists with limited physical resources, but unlimited general time. However, we must always keep in mind that the temporally-defined and limited individual is very different than the endless market. Individual's are forever constrained by both physical resources and their temporal expectation.

It thus becomes evident that the assumption that mis-pricings will be removed at some future date of time neglects the influence of new time on the present, and the expectation that the new future time will effect in some unknown way upon the unknown future. As a continual influx of new individuals, with new time durations, enters the market arena, time will cease to be a limiting factor *in the aggregate*. Instead, we will see that time will continue unfolding indefinitely, never reaching some finite ending point. The implications for the previous three points become clear with this in mind – their processes will continue forever in a general sense, but not necessarily by the same physical actor.

Time is determined through the past

Lastly, we see that EMH assumes that all time that occurs in the present, and will occur in the future, is completely separate of the past. This is a distinctly Newtonian viewpoint. As Bergson demonstrated, our futures are always colored by our pasts. Our interpretation of future events will forever be not in isolation, but the result of the past events that have occurred.

It is meaningless to speak solely in terms of time. For the actor, time never exists in solitude. Instead, time is always coupled with an action. It follows that when we speak of future points of time, we do not imply that the point is our focal point. What we actually are interested in is the action that will materialize during that the temporal advance to that future point. As a result of this, we see that the question is not whether the future time is separate and distinct of the past time. The correct question is: "is our *future* action separate and distinct of our *past* actions?"

We have seen that the past must always influence our future actions – our perceptions, expectations, tastes, preferences, etc. As such, our future actions will partly be influenced by our past actions, and by the uncertain events that will occur distinct of us. The assumption that price movements must be distinct when existing in separate time periods implies that prices exist outside the realm of action. However, prices are not entities that exist in solitude. Instead, prices are the sole result of human action striving towards ends. As such, prices cannot exist a temporally isolated life. They must always be influenced through the temporal past of the action that brings them into being.

Conclusion

We find the concept of time used in EMH erroneous. One aspect of it denies the existence of time (for the creation and dissemination of information). One aspect of it views time as existing as a finite entity (that time exists until all mispricings are rectified and then ends). Finally, a third temporal viewpoint contained in EMH is that prices exist in an isolated, Newtonian state (the future cannot be influenced by the past).

This grand temporal confusion is murky, and leads to internal inconsistencies within the hypothesis. How can these three, absolutely distinct, definitions of time coexist? The answer is they cannot – the definitions of time used are not applicable to the dynamic world of humans. EMH rests on a temporal bed that leads to false conclusions – conclusions not supported by the world humans must act in. Furthermore, it leads to conclusions that are inconsistent with parts of the world that the EMH posits to exist in.

Time does not exist as a separate entity absent action. Instead, *time has value to the acting* *human due to action*. Action exists in a world marked by time – the two cannot be separated. Only by adopting a Bergsonian, subjective viewpoint of time can we see how prices action is formulated in our temporal world.

Time and CAPM

CAPM is plagued by two specific temporal follies. The first is the assumption that preferences remain constant over time. A portfolio could be chosen at a given point in time, and the passage of time would not affect an actor's actions surrounding it. A movement to rectify this misgiving (through an inter-temporal capital asset pricing model (ICAPM)) has been expanded upon and developed over the past 35 years. A fundamental mistake still exists as the nature of future time, and the degree to which it can be determined. Both these points will be addressed.

Time and preferences

Markowitz (1958) assumed that an investor would select a portfolio at a point in time based on existing preferences. The passage of time would imply no fundamental shift in their preferences, and hence, the criteria for holding a portfolio would remain the same for the future.

As we have seen previously, time creates an uncertain future element existing in the present. Preferences cannot remain constant, as the dynamic essence of temporal passing creates disturbances that alter the setting that the decision was initially made in. In fact, as Shackle (1958: 13) mentioned, time is very individual, "[n]o moment has a brother. Between it and other moments there is an impassible barrier, it is imprisoned in solitariness." The conditions that exist at any point in time cannot remain identical given the temporal passage of time. This implies that as time passes, the preferences that an individual based their initial decision on must alter as well. A portfolio selected at a time in the past cannot persist to the future, as this future involves uncertain shifts that will alter preferences.

Dynamic time and preferences

Black (1972) tried to rectify this issue with his ICAPM. The issue that arises however, concerns the nature of time. Black, and followers, have confused the fundamental nature of time – uncertainty – with risk. The difference becomes manifest when we realize that there are events – uncertainties – which cannot be modeled in any concrete, determinant form.

Black would note that variables do indeed change over time. He would rectify this by assigning them stochastically. That is to say, variables would be assumed to move randomly. Future measures of risk, or return may be random variables, irrespective of market actors' actions. This part of CAPM would find no problem accepting the position that EMH offers.

However, the mere passage of time does not imply random action. In fact, we see that all actions undertaken must have a direction in the eyes of the actor, not just the random movements with no respect for a definite end. The passage of time, and its inherent uncertainty, are mitigated through the purposeful action of individuals, directed at reducing its uncertainty. A stochastic process cannot account for this directed action – it is purposeful in both the end it is directed at, as well as the time in which it occurs.

Time, when employed in the CAPM, creates a modeling problem. It does not pass randomly, nor does it pass with full certainty. It has an undefined degree of uncertainty inherent in it, but the way in which actors mitigate this uncertainty is purposeful.

Time and choice

The CAPM rests on the assumption that the choice investors make in portfolio/security selection is time invariant. Hence, the trade-off will continually be risk versus return. However, we have seen that the passage of time implies changing preferences for actors. The implication is that this choice cannot be assumed to continue indefinitely.

For instance, we will see in section shortly that an investor may not choose between the binary risk-return trade-off; they may prefer an additional variable to factor in to the decision. This decision itself cannot be seen as constant throughout time. As personal situations change, as well as market data, the criteria influencing choice change also. Cigarettes were at one point of time seen as having no negative health effects. An investor could possible have chosen to buy a cigarette company security based on criteria not considering this fact. We can assume that at that point in time, the criteria utilized were risk and return. However, let us move into the future where cigarettes are seen as having significantly negative effects on individuals, regardless of if they smoke or not. An investor conscious of this fact may now alter their criteria to include a new “healthy” criterion. Now the choice they make is based on three factors: risk, return, and how “healthy” the company's product it. It may even arise that the criteria now is based solely on “healthiness”, the original factors may not serve a direct purpose to the decision any longer.

As long as actors exist in time, they will face the invariable alteration of their preferences throughout time. This implies that a model cannot be built around the assumption that these decision criteria will remain constant throughout the future, unchanging from their initial essence.

Time and variables

The individual metrics that apply to the CAPM are also seen as time invariant. For instance, beta or the risk-free rate may be seen as static givens, existing outside the dynamic world of reality. However, just as was previously looked at, actor preferences will not be static, and the variables they decide upon cannot be static neither.

However, the problem, much like we saw previously, is not that these variables are dynamic, but that they are not randomly so. A stochastic process cannot be applied to these. For example, beta is a measure of specific risk a security has *vis à vis* the market. This is not a randomly determined variable, but is the product of human action. It will be purposefully altered, but in a way that may not be evident or clear to an investor.

The primary issue concerning the value of time is that the method that an actor uses to mitigate its uncertainty may be fully revealed to themselves, but not to outsiders. Hence, a company may be undertaking measures to alter its relative beta, but this may not be apparent to an outsider. In fact, it may not even be evident to a company insider. But neither of these negates the fact that the alteration of the metric occurs not as a random variable.

Time and the past

To mitigate the problem arising from the issue of the indeterminacy of future variables, an extrapolation from the past is undertaken. Hence, the future is projected from the previous behavior of the variable in the past. A confusion is made between the variable, and the cause of the variable's movement.

To use beta as an example again, it is not moved in isolation of other variables, but is the conscious action of an actor, or actors. Each will try to mitigate the uncertainty of the future based on their own personal experiences. This may be unknown to any actor but themselves. The implication is that the projection of an occurrence of the past into the future may have little consequence or predictive value. The reason is that the actual determinants, the reasons why an actor acts in a certain manner are distinctly conditioned by the time in which they exist, and act. As we have seen the individualism of time for the individual, we see the effect this will have on their action. An attempt for an outsider to try to understand this action, and hence extrapolate it into the future, may be mired with difficulties.

Conclusion

CAPM is affected adversely by ignoring the inter-temporal subjectivity of the decision making process. Time is explicitly acknowledged at every point of CAPM, a marked improvement over the inconsistencies previously looked at in EMH. However, we see that the time is not of a dynamic nature, but instead is viewed statically. *Time, when viewed this way, ceases to exist – time ends*. Rizzo (1995: 3) stresses the passage of time as the essential link that unites the remembered past with the unknown future via our real present:

> In the static conception of time it is virtual movement from past to future or, more precisely, from memory to expectation. The mnemic link to the past is responsible for the continuity of flow. But the flow arises out of the contrast between the remembered past and the expected future. Without the novelty of the future (seen only as "novel" only in contrast to the remembered past) there could be no sense of temporal passage.

The passage of time implies that variables will change. Actors' preferences will be altered as the temporal element changes their subjective valuations. What was important to them in the past may cease to be in the future. Similarly, even if these criteria were deemed time-invariant (i.e., the risk-return trade-off was absolute), the variables themselves would provide a complicating element in a temporal setting. Risk-free interest rates, betas, or market returns would all face the uncertainty inherent in the future. Mitigating this uncertainty by projecting from the past into the future may prove futile as the action that drives the changes in these variables may only be understood by the specific individual enacting the change. A vicious cycle emerges where we see that the passage of time implies that these same individuals may not act consistently towards these ends.

CAPM's explicit acknowledgment of the temporal element does little to fix the issue inherent in it. The basic treatment of time itself is not correct. The uncertainty and time-invariantness of time itself must be acknowledged for time to be effectively managed in any pricing model.

Conclusion

We have laid out some of the temporal problems that exist with the two building blocks of modern finance theory – EMH and CAPM. EMH proves to have a very muddled perception of time, simultaneously encompassing three distinct viewpoints. The biggest issue with its three perceptions are that they assume action exists without time. In fact, the two form an inseparable duality. Effects of the future can not be viewed solely as distinct moments in a time. Instead, we see that time periods of the future can, to the acting human, only be viewed in light of the expected actions that will occur during the temporal passing. *Time can only exist meaningfully for the human as a passage, never as a static point*. CAPM on the other hand formally acknowledges time. However, its conception of time is actually the very negation of time. A *static* time has been forwarded in this model, one that denies future changes to occur. However, we have seen that the fundamental crux of time is its future uncertainty; an uncertainty that implies change.

Time must not be viewed as an entity, or flow, existing independent of the actor's world. Instead, the very concept of time must be viewed in light of the action occurring in its midst.

Newtonian time has been compared with Bergsonian time. Time is not a static passing but a dynamic flux of new experiences. Newtonian time occurs in time, but the time actors experience is a continual flow of endless newness. In fact, this constant renewal of experience of time through time must occur, lest "real time will cease to be" (O'Driscoll and Rizzo 1996: 59). We find that, for the acting human's world, Bergsonian is the only approach we can take. This form of living time is influenced by its own passing. Time of one period is not equivalent to another period as it would be in a more linear viewpoint. As for the actor, time and action are inseparable, we find that an actor's future time will forever be colored by their physical past, and indeterminate future.

Furthermore, we have seen the indeterminateness of the value placed on time. Time exists not in a homogeneous flow, but each point in the flow is distinct from every other (Shackle 1958: 13). This implies that not only will value differ across individuals, it differs across time. The value of time can only be applied as the value towards a scarce means. As such, the end that is deemed to be reached through the passing of time influences and conditions the value we place thereof. As our sought after ends change, so does our value attached to the temporal element. Time and its value are therefore of a personal nature; impossible to be gauged wholly by an outsider.

As action is inseparable from a temporal element, any attempt at modeling action must base itself on a correct understanding of this temporal element. To the extent that only Bergsonian time is applicable to the realm of human action, this is the sole approach that can be taken to assess human action in an inter-temporal setting.

Appendix A: The Privatization of Time

It is common to view private property solely in terms of physical goods. Goods exist exclusively in both space and time. However, the attention of private property is solely directed towards the physical element that these goods occupy, while the temporal element is overlooked. Furthermore, time is the one element that all must share together. While money is the medium of exchange in our market economy, time is the medium that constrains our actions (Garrison 1984: 20). Time is an inescapable flow; a declining balance that unites all actors together. As Shackle (1970: 21) fated us, we are all "prisoners of time."

However, we have seen the fundamentally personal nature that time embodies. It is not an absolute element – it cannot be viewed solely in terms exclusive to itself. Time is inseparable from action for the acting human. Therefore, when we utilize time, we always utilize it directed at a specific end, using specific means. These means then exist in a concrete space (physically) and time (temporally).[52] Humans forever search for a method to privatize their scarce means. The incentive to capitalize and exploit these means has driven the creative expansion of methods to demonstrate, and hence trade, ownership over given objects.

The manner that humans have developed to privatize and economize time is through *interest*. We know that interest exists as we prefer goods for the fulfillment of ends temporally sooner than later (Mises 1949: 480). The reason is that means exist to satisfy wants. Wants exist to remove a felt uneasiness. It follows that we would value means that eliminate want-dissatisfaction temporally sooner to later. Originary interest results from the value spread between means and ends (Hülsmann 2002: 102). Hence, it also follows that pure time-preference, and the ensuing pure rate of interest, must always be positive (Mises 1949: 527; Rothbard 1962: 451).

Just as we use money in an economy to provide a common account with which to compare personal valuations of goods, interest is the method that provides a common medium of account to compare the valuation of time. The money economy is viewed as essential for an advanced society. The comparisons of value made possible through its use allow actors to exchange goods to those who value

[52] Our end, in distinction, must always exist in an expected space and time, until the point when we actually realize its attainment.

them more highly. Likewise, the existence of interest allows actors to compare valuations of time that exist in the economy. They allow temporal-trade between individuals who value time more than others.

Interest in this sense however must be seen as always existing on a physical object. It is impossible to trade time directly. No one individual can trade an hour of their time today for two hours of another's time tomorrow. However, an individual can trade an hour of their labor today for two hours of another's labor tomorrow. Likewise, an individual can trade the use of their car today, for a sum of fifty euros tomorrow. The spread between these two objects is the interest rate. The actors have not only traded the physical embodiments of the car and money, but the time at which the services of these means are to be used as well.

As the physical goods are traded, the time at which they can be enjoyed is traded as well. If a person trades their car today for fifty euros tomorrow, we commonly say that they do so as they prefer the use of the money tomorrow to the use of their car today. However, what we mean to say is that they prefer the use of the car during a given time period, to the use of the money in a different time period. The spread between the inter-temporal prices represents not only the value differential between the physical goods, but also the periods of time they are to be used during.

In a static sense, we can see the existence of money allows an economy to coordinate goods from those who value them less to those who value them more. This implies that the final want satisfaction available through the usage of these scarce physical goods is brought into a state of being that moves the actors into a state of greater final want satisfaction. In an inter-temporal sense however, we see the necessity of interest rates to coordinate what time period the use of a good is more desired in. Given this dynamic element, we now see that interest is a necessary coordinating factor to compare the inter-temporal wants of distinct actors.

A quick example may clarify the difference. Actor A may sell their car to actor B for 1,000£ today. All that we, as observers, can determine of this transaction is that Actor A prefers 1,000£ today to the use of their car, and vice versa. We may also say that Actor A prefers the use of 1,000£ for the temporal future to the use of their car for the same temporal future, and vice versa. However, now assume that it is not a pure exchange *per se*. Now suppose that the use of the car is exchanged for money. The use of A's car for one year is now exchanged , and in return B gives A 1,000£. At the end of the year, the car, and its use, will return to A who will also keep the sum of money. Now we can see a different conclusion emerge. Instead of telling us that A and B had their static preferences explicitly revealed, as was the case in the previous example, we see the story takes on an inter-temporal twist.

Actor A now prefers the use of 1,000£ over the course of the future to the use of their car over the coming year. Actor B now prefers the use of the car for a year to the use of 1,000£ forever. The temporal element has become apparent. They have exchanged not only physical goods, but the temporal elements inherent in their respective uses.

The implication is that an economy may be able to statically coordinate itself with the desires of its individual actors through the use of money. However, only the concept of interest will be able to perform this task in an inter-temporal environment. As all action takes place inter-temporally, we see Mises' (1976: 170) insight into the matter as he wrote:

> An account for "time" does not appear on the business man's books. No price is paid for it on the markets. That it is, nevertheless, taken into consideration in every exchange could not be seen from the standpoint of the objectivist theory of value, nor could one be led to this reflection on the popular precept contained in the saying, "time is money."

What Mises failed to realize was that a price for time does exist: interest. The overriding concept that creates coordination in a static sense if that of private property of goods. One of the issues we have discussed with time is that it is a fundamentally heterogeneous passing for the acting human. Interest, in distinction, is homogeneous (Mises 1949: 526). The source of interest is always the same; a temporal trade-off. A heterogeneous temporal element produces massive issues concerning the exchange thereof. By valuing our temporal opportunities with interest, the opportunity to compare values and exchange arises.

We can see that ownership of a good implies the exchange thereof. To engage in inter-temporal trade requires a concept of ownership of time; a way to trade the use inherent in a period of time. The method that actors have created to this end is the concept of interest. The benefit of allowing comparisons of temporal values allow actors to trade those periods of time that are valued ordinally opposite between actors – inter-temporal coordination is thus enabled. Contrary to Shackle's dreary prospect, we are not fated to be prisoners of time forever, there is a method we can use to liberate this temporal element from our existence (at least for a time) – interest.

2. Risk and Uncertainty

A great deal of confusion exists as to what the true nature of uncertainty as is applicable to the financial realm. Both EMH and CAPM make the assumption that all uncertainty is fundamentally knowable. That is to say, the only unknowns that exist and are relevant to the pricing of assets are known unknowns. However, under this viewpoint, that concept which finance academics refer to as uncertainty is really a wholly separate class of events: risk. The distinction may seem small, but has grave implications for the finance base.

As Knight (1921) demonstrated, there exists an uncertainty that acts as a fog encompassing the world humans act within. This uncertainty is, in his eyes, unimaginable, and impossible to mitigate. Our economic system is thus open-ended, real changes will occur that cannot be determined in advance. The implication is that the uncertainty regarding an individual's future is fundamentally unknowable.

Mises (1949) was able to correctly see the dichotomy that exists and classify it accordingly. Events that fall under the realm of physical sciences have unknowns that are fundamentally knowable. No real element of unknowable surprise exists to skew the future. This represents a situation where risk exists, as all possibilities are knowable in advance, and hence, can be planned for to some extent. However, in the realm of the natural sciences, of which finance falls into, there exists a wholly separate class that has future unknowns that are fundamentally unknowable. Hence, a situation of true uncertainty exists. There can be no way of predicting in advance all the possible eventualities that may arise.

EMH assumes that all current information that is relevant to a security's price is known and factored into this same price. It follows that future price movements will only be caused by presently unknown future information. However, inherent in this belief is that idea that the future information is the *only* thing that will effect prices. This view fundamentally overstates the amount of uncertainty inherent in the future. In fact, as will be seen later, it removes the entrepreneurial function from affecting and determining prices. A state of uncertainty does exist concerning the future, but this is managed through insightful entrepreneurial action.

CAPM, on the other hand, believes that all future uncertainty is fundamentally knowable, and hence, reducible to a statistical function. What CAPM advocates refer to as uncertainty is actually a separate class of risk. When viewed this way, we see that there exists a fundamentally unknowable portion of future information, unable to be explained, or forecast in concrete, mechanically predictable terms.

When viewed in light of the true dichotomy between risk and uncertainty, we see that both EMH and CAPM have similar flaws in their reasoning. We also find that the only true view of the future concerning acting humans is available – uncertainty. Our actions cannot be reduced to a set of known variables, and their expected or historical probabilities. Instead, we find that action is open to surprise, elements that cannot be forecast in advance with any degree of certainty.

Risk and Uncertainty

Knight (1921) identified the future as fundamentally unknowable concerning human actors. In his eyes, the future was an immeasurable element, unable to be calculated upon due to this uncertainty. In his (*ibid.*: 26) own words:

> Uncertainty must be taken in a sense radically distinct from the familiar notion of Risk, from which it has never been properly separated.... The essential fact is that 'risk' means in some cases a quantity susceptible of measurement, while at other times it is something distinctly not of this character; and there are far-reaching and crucial differences in the bearings of the phenomena depending on which of the two is really present and operating.... It will appear that a measurable uncertainty, or 'risk' proper, as we shall use the term, is so far different from an unmeasurable one that it is not in effect an uncertainty at all.

It becomes clear that there is an important distinction to be made regarding the future course of events. On the one hand, there exists a class of events that is predictable in a measurable way: risk. On the other hand, there is the fundamentally unknown portion of the future, that future where surprises occur: uncertainty. Under this Knightian view, the uncertainty of the future is inescapable for actors; there is no method to manage it. In Lachmann's (1976: 55) confirming viewpoint, “the future is to all of us unknowable.”

Mises (1949: 105) noted that the element of uncertainty is fundamentally related to the very concept of human action; “[t]hat man acts and that the future is uncertain are by no means two independent matters, they are only two different modes of establishing one thing.” Mises (*ibid.*) would also add, “[i]f man knew the future, he would not have to choose and would not act.” Hence, we see that the uncertainty of the future is not only an element that human action must manage – *it is a requisite for it to occur*. A world lacking uncertainty would be a world lacking action.

As will be looked at in the next section, there is a method that exists to deal with future uncertainty, contra what Knight thought: the entrepreneur. The entrepreneur provides an important service in mitigating this unknowable element, and creating a more cohesive whole for the rest of the market's actors. Mises (*ibid.*: 107) divided the world into to realms, with applicable probability options for each: class and case probability.[53] Under class probability, we assume to know, or actually do know perhaps, all there is to know about a future class of events surrounding its possible future outcomes concerning the whole group. We might not know anything exact about any individual object in a group, but we know that the whole group will behave in a defined way. In Mises' (*ibid*.) own words then, a good example would be where:

> [T]here are ninety tickets in a lottery and that five of them will be drawn. Thus we know all about the behavior of the whole class of tickets. But with regard to the singular tickets we do not know anything but that they are elements of this class of tickets.

Examples of class probability fall into the category of market risk. There are knowable future outcomes. These outcomes thus have a fundamentally insurable nature to them, when viewed as parts of the whole group.

In distinction, a class of probabilities exists that is independent and unknowable in advance. These are case probabilities. In these individual cases, we may know some things which will determine a particular outcome, but there are other influencing factors that are unknown in advance. Case probability shares nothing in common with its class probability counterpart, except the fact that we have a fundamental incomplete knowledge set. If we have perfect knowledge about either instance, we would not act towards them. Hence, the fact that we are taking action implies that there must be knowledge in either circumstance that is unknown to us. Mises (*ibid.*: 111) would provide the following example of a case probability:

---

[53] Others have followed Mises' lead with this dichotomy. Langlois and Robertson (1995: 18), for example, also divide the world into two types of uncertainty: (1) structural – decisions based on the future are as yet unknowable, and (2) parametric – those that arise from a range of market imperfections, known in advance. The distinction is the same, and they, like Mises view parametric uncertainty as the only insurable type of uncertainty, or risk.

Two football teams, the Blues and the Yellows, will play tomorrow. In the past the Blues have always defeated the Yellows. This knowledge is not knowledge about a class of events. If we were to consider it as such, we would have to conclude that the Blues are always victorious and the Yellows are always defeated. We would not be uncertain with regard to the outcome of the game. We would know for certain that the Blues will win again. The mere fact that we consider our forecast about tomorrow's game as only probable shows that we do not argue this way.

This particular case represents an example of uncertainty. We may know some of the factors that will serve in determining the outcome of tomorrow's game, but there are other factors fundamentally unknowable to us in advance.

In table 6, we can see the dichotomy as Huerta de Soto (2005: 47) delineates it.

Two Realms of Probability	
The Field of Natural Science	**The Field of Human Action**
Class Probability: The behavior of the class is known or knowable, while the behavior of its individual components is not.	Probability of a unique class or event: class does not exist, and while some of the factors which affect the unique event are known, others are not. Action itself brings about or creates the event.
A situation of insurable risk exists for the whole class.	Permanent uncertainty exists, given the creative nature of human action. Uncertainty is not insurable.
Probability can be expressed in mathematical terms.	Probability cannot be expressed in mathematical terms.
Probability is gauged through logic and empirical research. Bayes' theorem makes it possible to estimate the probability of class as new information appears.	Probability is discovered through insight and entrepreneurial estimation. Each new bit of information modifies *ex novo* the entire map of beleifs and expectations (concept of surprise).
An object of research to the natural scientist.	A concept typically used by the actor-entrepreneur and by the historian.

Table 6

As Huerta de Soto (2005: 46) tells us, “the future is *open* to all of man's *creative* possibilities, and thus each actor faces it with *permanent uncertainty*.” This future element of uncertainty is not known in any exactitude in advance; it represents options that we may not even realize are available yet. As Shackle (1972: 422) sums up, “[s]urprise is that dislocation and subversion of received thoughts, which springs from an actual experience outside of what has been judged fully possible, or else an experience of a character which has never been imagined and thus never assessed as either possible or impossible."

Unfortunately, this conception of two differing classes of uncertainty – risk and uncertainty for Knight – or – class and case probabilities for Mises – has largely been lost in the present world of economic analysis. It may prove instructive to quote Hirshleifer and Riley (1995: 7) from their influential text, *The Analytics of Uncertainty and Information*, as they outline the approach to be utilized, and the assumptions it is based upon:

> The approach here does not allow for the psychological sensations of vagueness or confusion that people often suffer in facing situation with uncertain (risky) outcomes. In our model the individual is neither vague nor confused. While recognizing that his knowledge is imperfect, so that he cannot be sure which state of the world will occur, he nevertheless can assign exact numerical probabilities representing his degree of belief as to the likelihood of each possible state. *Our excuse for not picturing vagueness or confusion I that we are trying to model economics, not psychology.*

In fact, the true excuse for not envisioning vagueness or a lack of complete knowledge as to future existing states of being is the problems this creates in statistical modeling. As a result, any idea or conception of Knightian uncertainty is removed from the actor's decision making process.[54]

We can see that the realm of finance, itself a part of the natural sciences, or human action, requires a definition of case uncertainty. Our individual actions are creative in nature, and hence, constantly provide a fresh, previously unknown, input to the future. We can also see that uncertainty, in this sense, cannot be reduced to a mathematical probability. It is impossible to state that there is a 50% chance an individual will partake in a specific action at a given date. In fact, case probability is the only situation where statistical measures can be used and the future can be reduced to a numerical expression (Mises 1981). We can only reduce our expectations in an uncertain world to a fundamentally subjective belief, dependent on our own current level of knowledge concerning the

---

[54] Of course, neither Hirshleifer nor Riley seems to exhibit a thorough understanding of what true Knightian uncertainty is, or its implications. They (1995: 10) continue by stating that, "[i]n this book we disregard Knight's distinction, which has proved to be a sterile one. For our purposes risk and uncertainty mean the same thing." However, at the end of the same paragraph they seem to support the Knightian distinction writing, "[d]ecision-makers are therefore never in Knight's world of risk, but instead always in his world of uncertainty." That the distinction is thus underemphasized cannot, at this point, be overemphasized.

event in question.[55]

EMH and Uncertainty

One conclusion of EMH is that as all information is currently factored into an asset's price, any new movement can only result from new information. This viewpoint drastically overestimates the amount of uncertainty in the future. In fact, it takes on the Knightian view that there is no method that humans can use to proactively account for future changes. As will be developed later, the entrepreneur's prime function is to look into the future and move us towards that state. As such, there is a method available to mitigate the uncertainty of the future.

As Lachmann (1977a: 90) would state, "[t]he impossibility of prediction in economics follows from the fact that economic change is linked to change in knowledge, and future knowledge cannot be gained before its time." A brief remark should be made at this point however as to the nature of prediction. Although the future may forever exist as a fundamentally unknown state, we can utilize praxeologic knowledge to grant us, *with pure certainty*, the outcome of various courses of action (Mises 1949: 117). However, a condition on this certainty is that it can only be gauged in purely qualitative terms. We could never state that if a certain event happens, an asset's price will increase by 4%. We can state that, *ceteris paribus*, if a specific event occurs, an asset's price will increase. The implication for EMH is that, although the future is fundamentally uncertain, there exists a way to mitigate this uncertainty *in a qualitative way*. It is the role of the entrepreneur to perform this task, which will require further explanation in the following section. In this way, we can see the entrepreneur creates the future he will act in, through the present acts they have performed (Kirzner 1985: 56).

CAPM and Risk

CAPM, on the other hand, views the uncertainty of the future erroneously. Although it is called uncertainty to academics developing CAPM, the word, and definition, of the future they search for is risk. By treating the future as fundamentally knowable, it is assumed that future risk can be reduced to statistical measures. For instance, the return on a risk-free asset is deemed known in advance. Or beta, as a measure of security specific volatility regarding the market, is assumed as being fundamentally

[55] In this way, uncertainty can be summarized as the dispersion of the sum of the individuals' subjective beliefs as to the future. See Hirshleifer (1973: 31).

knowable in advance. We see that these are metrics that result from human action, and as such fall into the realm of uncertainty, or case probability.

The definition of uncertainty that is assumed in CAPM can be summed up from this passage from Arrow (1974b: 33):

> Uncertainty means that we do not have a complete discretion of the world which we fully believe to be true. Instead, we consider the world to be in one or another of a range of states. Each state of the world is a description that is complete for all relevant purposes. Our uncertainty consists not in knowing which state is the true one.

Arrow assumes here that uncertainty exists in the future, but we know that it exists. The only problem with this viewpoint is to select which future state is the most likely, and act accordingly. This is a misguided viewpoint however, as Sautet (2000: 10) states, "a definition of uncertainty where there is only uncertainty over which future state of the world is true does not help us to understand the economic problem."[56]

Rothschild and Stiglitz (1970) demonstrate that the definition of risk as asset Y having a greater variance of returns than asset X may be misplaced. By their own reckoning there are three better, and equally applicable, measures of risk that should be used instead. The first is where asset Y has a return that is affected by a random "noise" which disturbs the flow of returns. The second is the case where every risk averse investor would prefer asset X to asset Y, we could surely conclude that asset X is less risky than asset Y. Finally, if asset Y has more weight in its return distribution's tails than asset X, it is reasonable to state that asset Y is more risky than asset X. In cases two and three we should note that the risk is still concerned with a knowable risk. It is only in the first case that we see a case where genuine uncertainty could exist. However, the fact we have no knowledge of this risky "noisy" term implies we cannot effectively conclude anything about it until it occurs. Using measures of variance to determine risk or uncertainty (if it could be measured) have generally taken prevalence due to the long historical use of variance as a measure of dispersion in statistical theory (*ibid.*: 226).

The true problem that presents itself is that the future options are fundamentally unknowable in advance. They cannot be reduced to a single statistical measure, as CAPM assumes.[57] This viewpoint of

56 Arrow (1994: 7) would change his definition of uncertainty later in his career, whereby the problem actors are faced with is new information in a system, not merely information that exists and can be acquired from someone.

57 Remember that earlier we saw that the measure of an unanticipated event occurring, e_i, was assumed to be zero, as it represented an event that was unexpected. Herein lies the exact issue of true uncertainty however.

risk wildly underestimates the possibilities that human actors are able to create in the future, *irrespective of the current available data or information*. We thus see the problem that arises in trying to price an asset according to an unknown future variable. After Markowitz (1952) we see that the trade-off underlying CAPM is that higher risk will, *ceteris paribus*, return a higher rate. However, to follow this to any useful end, we assume that the degree of future risk is known in advance. Huerta de Soto (2005) recalls that the unlimited nature of future uncertainty renders traditional notions of risk as unsuited for the realm of human action. In particular, he reckons this is due to two reasons: (1) actors are not conscious of all possibilities (we exhibit bounded rationality), and (2) actors only possess subjective, continually modified beliefs concerning the future (Mises' case probabilities). As future risk, by definition, lies in the future, it is not risk at all, but uncertainty. As such, it cannot be reduced to a probabilistic statistic, as a chance event within a class of similar class events.[58]

Conclusion

We have seen that there is a fundamental difference between risk and uncertainty on the market. Risk concerns objects of the physical world. It belongs to a class of similar assets, whereby their future behavioral possibilities are known in advance, and thus can be projected statistically. This type of risk, or class uncertainty, can be mitigated through insurance considering the class as a whole. All entities of the larger class behave according to their past, and thus, this is projectable into the future.

In distinction, our human realm is not marked by risk, but uncertainty, or "case probability" to use Mises' (1949: 107) term. The primary aspect of this is the individual nature of events. Independent events are not assigned probabilities dependent on a class that they can't belong to. Instead, each event must be judged on its own merits, and relevant information. In this way, we see this uncertainty is fundamentally uninsurable as no objective probability exists for its future occurrence. Instead, entrepreneurs must rely on their subject beliefs as to the probability of these future events occurring, and as to their ancillary effects.

Knight (1921) viewed the future as fundamentally unknowable. It is thus a fog that is impossible to manage with any sense of certainty. Knight, however, overlooked the human entrepreneurial ability to mitigate this fog of uncertainty (Kirzner 1985: 97).

---

[58] It becomes more than a little disconcerting that Markowitz (1991: 470) could not correctly identify the true definition of uncertainty, given that he viewed "the existence of uncertainty [as being] essential to the analysis of rational investor behavior." Dothan (1990: 3) defines uncertainty, as it concerns mathematical modeling in finance, as "a listing of all the basic events or states that could occur during the time period of the model and their probabilities."

EMH users imagine the world that exists as being Knightian in nature. The future is unknown in advance, and the only way to alter it is through new information becoming available. Hence, there is no way an entrepreneur could forecast in advance the actions, and results, of the future. However, we see this significantly underemphasizes the role entrepreneurs serve in moving the economy forward. There are methods they can utilize to mitigate this uncertainty. One, given by Mises (1949: 117), is that through praxeology actors can determine future events with certainty. O'Driscoll an Rizzo (1996: 38) had this to say about the entrepreneurial treatment of future uncertainty:

> [G]iven the overall context of a change in knowledge, we can show how the move from framework 1 (Fl) to framework 2 (F2) is intelligible, in the sense that a metatheory can be constructed in which a loose dependency on Fl is shown. F2 is more likely (though not necessarily highly likely or probable) given Fl than it would be given some other F1.[footnote omitted] On the other hand, we might say that, given Fl, many possible alternative frameworks can be ruled out and that only a class of subsequent frameworks (which includes F2) can be determined.

Hence, an entrepreneur can forecast a piece of information, to some extent, before it is actually developed. The removal of unlikely situations can reduce the infinite amount of future possibilities. The future may have an unknowable degree of uncertainty, but the subjective beliefs of the entrepreneur can mitigate this uncertainty to make the coming future manageable.

In distinction, CAPM assumes that the world that concerns us is not marked by uncertainty, but risk. Hence, future possibilities are known to exist, the only pressing matter is assigning a probabilistic aspect to the outcomes. Hence, all unknowns of the future are known to exist. This confuses the problem at hand. Human action always involves an uncertainty that is not known to exist in advance. There are future events that we cannot know to exist at any given point in time. The idea that we can know in advance, with probabilistic certainty, a factor such as risk obscures the true problem that we face – the fundamental unknowable nature of future data. As such, we cannot reduce the problem to one of assigning probabilities to the likelihood of future events, but must accept their unknowable aspect.[59]

One of the sharpest divides that separates the realms of Austrian economics from other, more

[59] Additionally, we see risk exists not in a general sense, but instead applies to the specific entrepreneur's insight and ability to correctly see into the future.

mainstream, schools of thought is the treatment of risk and uncertainty. Austrians have readily embraced the fog of the future, and the imitations that this places on our actions in the present. The mainstream, in distinction, have modified their viewpoint of uncertainty to meet their methodological needs. This has created problems, with statistical probabilities unable to cope, predict, or interpret the dynamic world of human action. As Arrow (1974a: 1) would expose in his 1973 Presidential address to the American Economics Association:

> [T]he uncertainties about economics are rooted in our need to a better understanding of the economics of uncertainty; our lack of economic knowledge is, in good part, our difficulty in modeling the ignorance of the economic agent.

Any finance theory must recognize the true natures of risk and uncertainty. Risk as a metric can only apply to a static field, such as is found in the natural sciences. Uncertainty, on the other hand, is how the physical scientist must consider the world. Human action always must face a future that has an unknowable element to it. However, as actors we are not without aid in dealing with this uncertainty. Economists such as Knight, Shackle, or Lachmann, were incorrect as to the true nature of the future uncertainty. Although there exists an uncertain future, "it does not follow... that everything concerning our future must be uncertain" (Hoppe 1997: 64).[60] As we will see in the next section, the entrepreneurial function continually moves towards mitigating these unknowable events. In this way, actors can add an element of cohesiveness to this uncertain aspect.

[60] Interestingly, Lachmann (1976: 59) would also note that "[t]he future is unknowable but not unimaginable." It becomes obvious that even Lachmann felt there was a method humans could access to mitigate the uncertainty of the future.

3. The Entrepreneur

A pre-history of the entrepreneur

One of the history of economic thought's most overlooked chapters concerns the entrepreneur. Cantillon, in perhaps his greatest contribution to economic science, was the first to identify and stress the importance of this role. The role that was required was one that would face the uncertainty of the future, invest and pay the expenses in the present, and the return in the future would be entrepreneurial profit. Profits, therefore, were seen as the reward for successful forecasting of the future uncertainty, and realized through the productive structure.

In fact, Cantillon would divide the world into two classes of individuals: hired people, and farmer-entrepreneurs. The first class were hired and received fixed wages, or rents. The farmer-entrepreneur is defined by the non-fixed, uncertain benefits they receive. As the investment they undertake is always undertaken in the present, it is viewed as fixed. Hence, any residual income that exceeds this fixed amount will represent the entrepreneurial profit.

The true source of the entrepreneur stems not only from future uncertainty, but also from the nature of the market itself. Cantillon posited that in a world with one monopoly owner, they could decide what to produce, and at what price to sell it at – the need for an entrepreneur would be eliminated. However, in the decentralized world of the market, there is much scope for the setting of production levels and determination of prices, and hence, a great need for the entrepreneur.

The Cantillonian entrepreneur focused on his function regarding bearing uncertainty, which contrasts sharply with the Schumpeterian entrepreneur, who focused more on personality aspects.[61] Also, contra the Schumpeterian role, Cantillon's entrepreneur provided a coordinative function in the market. That is, they are seen as balancing supply and demand in various markets inter-temporally.

Turgot, besides his brilliant theory on capital and interest, also formulated a theory of the entrepreneur. In his view, the capitalist accumulated wealth in the form of money, and then invested this wealth into capital goods. In Turgot's eyes, the capitalist-entrepreneur was the essence of this role. Hence, others' wealth depended on this individual as well. For instance, the capitalist-entrepreneur, through their savings, would pay the wages of their hired workers until an income stream could be produced by the investment.

[61] For example, "The entrepreneur uses his personality and nothing but his personality" (Schumpeter 1911: 417).

Cantillon's entrepreneur had missed one vital concept in development – a lack of capital theory. This was an element Turgot was more than able to provide. The driving force was no longer viewed as only entrepreneurial insight, nor capital accumulation. The role instead is inseparable and drove the economy ahead to progress. Turgot was able to identify that capital is not instantly productive, but suffers from a time-lag in realization. In his own distinction, contra to Cantillon's dichotomy, there were two separate classes in society: the capitalist-entrepreneurs who invested in advance for future needs, and the simple artisans, who held no property but their bodies and could not advance anything for the future except through their daily labor. Also pointed out was the workers paid in the present for goods valued in the future would take a discount on their wage in light of this. Hence, a Wicksellian "natural rate of interest" theory was derived by the time preference workers exhibited on receiving wages today over in the future.[62]

The French philosopher Condillac was highly influenced by Turgot's exposition of the entrepreneur. The view of the entrepreneur as being an uncertainty bearer was gaining acceptance. The inter-temporal role of investment was also gaining acceptance, complete with the additional uncertainty this temporal element caused.[63]

The Smithian/Ricardian trap of perfect knowledge had devastating effects on the entrepreneurial role. Adding to was this the coming acceptance of Walrasian equilibrium. There was no need for a co-coordinating factor, as the economy had already reached full coordination by the time equilibrium was reached. This purging of uncertainty from the system eliminated any role for an uncertainty bearer, and the entrepreneur was dropped and forgotten from economics.

Many place the loss of the entrepreneur on Smith's shoulders. However, as Holcombe (1998: 45) points out, Smith would hold a view of production, that would include an entrepreneurial element which would lead to increased innovation, and hence, in Smith's eyes, division of labor for the impetus for growth in the economy. Ricardo (like his contemporary Malthus), in distinction, would view growth as limited by resources, with Smith arguing for unlimited potential. Additionally, Ricardo held the view of supply and demand forever being in perfect harmony on the markets, without paying heed to the

62 Hence, Turgot was also able to dismiss the physiocrat fallacy that savings were "leaked" from the economy and thus, wasted. By demonstrating that savings are always used for something eventually, the entrepreneur does not save money in waste, but for a better opportunity in the future. This predates Garrison's (2001, 62) concept of "saving up for something" or SUFS.

63 Despite falling into neglect in Britain, there were pockets that still emphasized the entrepreneur's role. For example, the Irish-Australian economist William Edward Hearn followed in Turgot's footsteps, identifying the entrepreneur as contracting between labor and capital to maximize upon a future expected return. Hence, the entrepreneur pays a fixed price in return for the expected larger, but currently unknown, gains. Also integral was the view of capital being inseparable from the entrepreneur (Rothbard 2006: 464).

process that would achieve this occurrence. Hence, by Holcombe's reckoning, it is Ricardo who should properly be bestowed with the title of the “entrepreneur-eliminator” in modern economic thought.

Whoever eliminated this role from economic theorizing, Say resurrected it from a static grave, and saved economics from this equilibrium fate. It was not with the same emphasis as Cantillon or Turgot had emphasized, but the inter-temporal role was there none the less.[64] Say's entrepreneur was seen as the primary mover of the economy as the head of the firm. As such, the entrepreneur was also a capital owner, as owners of firms in his time were usually directly involved in the funding process. However, it was also seen that capitalists would rent their goods out to entrepreneurs for a fixed price, thus receiving a guaranteed rent, in exchange for the uncertainty that the entrepreneur would now bear. Hence, the speculative role was transferred.

Furthermore, Says' entrepreneurs act as brokers of sorts, collaborating with buyers and sellers on the market. Hence, productive factors are allocated throughout the economy directly through the entrepreneurs' actions. In this way, sellers need not be directly concerned with the actual demand curve for their product. This information was transmitted to them through the entrepreneurs acting as intermediaries. Likewise, buyers are made aware of the supply of goods through this same process. Prices are constantly moved towards equilibrium as the entrepreneur compares and adjusts them accordingly.

Foreshadowing Schumpeter's entrepreneur, Say would note that the entrepreneur had personal qualities that made them well suited for this role. Specifically, knowledge, judgment, and perseverance all aided their ability to satisfy the wants of consumers and producers. Furthermore, a necessary quality was to compare the prices as they appeared in geographical, and inter-temporal, form. Hence, they acted as physical and temporal arbitragers to keep prices in-line with market needs. Schumpeter, however, was quite critical of this viewpoint. He felt that Say's treatment of the entrepreneur was more along the lines of a static manager and organizer, not a dynamic risk bearer managing uncertainty.

Say was quite critical of Smith and the Smithians for eliminating the entrepreneurial role in the economy. The failure to distinguish between entrepreneurial profit and profit on capital was to neglect an important, distinct role in the economy. In fact, the issue of this division remains today. Kirzner (1973: 54) notes the difficulty in distinguishing between the two types of profit (purely entrepreneurial and capitalist), and hence, we can see how it is easy to mistakenly view areas with high profits as being capitalist in nature, to also be the result of growth that is capitalist and not entrepreneurial in nature.

[64] On the continent this was true, at least in underground circles. In Britain, the dominant thought was still geared towards equilibrium and the role of the entrepreneur lost (Rothbard 2006: 25).

Lachmann (1956: 98) also notes the difficulty that arises from trying to disentangle pure profits and capital gains.

Early Modern Views of the Entrepreneur

As we saw in the previous section, the Knightian view of uncertainty excluded the entrepreneur from having any sort of coordinating effect. The "fog" of the future is so thick that it is purely unmanageable. The implication is that entrepreneurial profits are as likely to develop as are losses (Knight 1921: 63). The only way for an "entrepreneur" to create a coordinating factor in the market place in this Knightian world is through sheer luck.

Schumpeter's (1942) entrepreneur was able to act knowingly looking to the future. The temporal element is manifest in this entrepreneurial function. However, in Schumpeter's eyes, the entrepreneur was not an individual who added a co-ordinating effect to the economy. Instead, entrepreneurial actions were just as likely, if not wholly certain, to cause a disco-ordinating effect on the economy as a whole.[65] Hence, this process would give rise to Schumpeter's, now famous, idea of "creative destruction." Just like Knight viewed the entrepreneur, there is no real coordinative effect bringing together the wants of consumers into a common directed goal.

In fact, by introducing new innovations, technologies, techniques, or distribution methods, the emphasis was placed on a *distruption* of the previous market process (see Schumpeter 1991). A disruption of the existing equilibrium appears, through a new innovative process that changes the course of events. As a pre-existing equilibrium is disrupted, a new "higher" equilibrium will take its place as a creative innovation will alter economic conditions. In doing so, however, Schumpeter (1949: 128) would also define entrepreneurial profit as "a surplus over costs... the difference between outlay in a business." Hence, entrepreneurial profit becomes a standard factor payment – a type of wage paid for innovative activity which is susceptible to eroding away.

Furthermore, Schumpeter would have significant difficulty identifying who the entrepreneur actually was in the economy. As he (1989: 226) states the problem facing firms in the market, "the question that is never quite absent arises with vengeance, namely, who should be considered as the entrepreneur." This inability to pinpoint who the entrepreneur physically was would create difficulties

[65] Kirzner (1973: 127) would describe this essence of Schumpeterian entrepreneurship as "the ability to break away from routine, to destroy existing structures, to move the system away from the even, circular flow of equilibrium... For Schumpeter the entrepreneur is the disruptive, disequilibrating force that dislodges the market from the somnolence of equilibrium."

as he tried to identify what the entrepreneur actually did.[66]

However, we can see *prima facie* that the entrepreneur must have a co-ordinating effect. The world is not mired in chaos, with each individual acting independently of others. In distinction, we see that the world is formed by individuals working with the actions of others in mind to move towards and create an as yet unknown future. It would take the presence Mises to establish this role of the entrepreneur.

The Misesian Entrepreneur

The explanation of the entrepreneur would be one of Mises' greatest contribution to economic science. The market was not only a process, but this process was the creation of entrepreneurial activity moving into the unknown future. Processes, and their inherent temporal element, are impossible to disentangle from the entrepreneurial function. The uncertainty inherent in every action (due to the future as has previously been seen) creates an entrepreneurial aspect in every action. The inter-temporal element that humans all must share and experience together creates entrepreneurs in all of them as well.

Driven by the profit motive, whether monetary or psychic, Mises (1951: 8) would sum up the entrepreneur thusly:

> What makes profit emerge is the fact that the entrepreneur who judges the future prices of the products more correctly than other people do buys some or all of the factors of production, which, seen from the point of view of the future state of the market, are too low... the difference [between these inter-temporal prices] is the entrepreneurial profit.

Hence, we see that a Misesian *pure* entrepreneur is an arbitrager.[67] Looking at expected disparities that exist between the present and the future, and moving to correct these through their coordinating

[66] A complicating factor was that Schumpeter was never able to shake the Walrasian paradigm that would dominate his discourse. In fact, the logical conclusion he would arrive at, at least according to Rothbard (1987: 102), was that in a static equilibrium, innovation, and hence change, could never occur absent an exogenous shock (i.e., a credit inflation, for example). As we have seen, the entrepreneurial creation of innovation defies any source of funding, which is what brought Schumpeter to the ill-conceived notion of the entrepreneur as being someone moving the market away from equilibrium. See also Rothbard (1985: 285).

[67] This immaterial input that the entrepreneur provides is what ensures that the most needed processes are undertaken by the economy, the ones that will add the most value to the consumers. See Hayek (1936a: 174), Hülsmann (1999: 64), or Mises (1998: 295) where he states "profit and loss [and hence growth] are entirely determined by the success or failure of the entrepreneur to adjust production to the demand of consumers."

process. Hence, we see the importance of Mises' (1957, 320) statement where he refers to entrepreneurs as “historians of the future.”[68] The future is a direct result of the entrepreneur's actions in the present.

The pure entrepreneur is loaned resources from capitalists to achieve their goals – hence the theoretical detachment of pure entrepreneurs from capitalists. Mises (1949: 254) that while pure entrepreneurs never *needed* to be resource owners, capitalists were quite often entrepreneurs. This occurred through their two-fold role as speculator and entrepreneur. In reality, however, there is a necessity for property and entrepreneurship to be intertwined; the possibility of loss is what separates the function from that of a mere hired employee. If we assume for a moment that entrepreneurs are loaned resources to realize their plans, we see they remain property-less as their assets will be completely offset by their liabilities. If their plans prove to be successful, the profit will become theirs – hence no significant issue arises provided the plan is successful. However, if their plan goes awry and losses befall them, then the loss will fall upon the capitalists who originally loaned them the resources. It follows that such an entrepreneur would really be a paid-employee of the capitalist, earning 100% of the profits, while not having to be concerned with loses. We see the logical necessity of having the two roles intertwined.[69] Mises would distinguish this type of entrepreneur from the *pure* type by referring to it as a “promoter-entrepreneur.”[70]

If the logical necessity of having both entrepreneurial ability *and* resources was known to Mises, why the need for a duality of entrepreneurs? As Mises himself lamented, the promoter-entrepreneur could not be defined with any sort of “praxeological rigor” (1949: 256). This becomes clear as the promoter-entrepreneur has qualities which are given by human character, and hence, do not affect all humans alike. As Salerno (2008: 195) reckons, Mises' true depiction of entrepreneurship depended on individual qualities that breed learned differences. Mises would refer to these qualitative differences in human character with words like “agile,” “restlessness,” “eagerness to make profits as large as possible, or “ingenious newcomers” (Mises 1949: 256; 1951: 17).

68 In light of Mises' (1957: 203) view of historians as being extremely limited in their outlook to the future: “The honest historian would have to say: Nothing can be asserted about the future.” We see that the Misesian entrepreneur is the antithesis of the historian.

69 Of course, Hayek (1931: 277) identified this necessity as much as 18 years prior:

> [T]hese two functions [ownership of capital and the entrepreneurial function] cannot be absolutely separated even in theory, because the essential function of the entrepreneurs, that of assuming risks, necessarily implies the ownership of capital. Moreover, any new chance to make entrepreneurs’ profits is identical with a change in the opportunities to invest capital, and will always be reflected in the earnings (and value) of capital invested.”

70 Later Misesian followers have largely dropped the term, for reasons owing the connotations associated with the term “promoter” today. See, for example, Rothbard (1985).

Using a thought experiment, we can see just how it is that entrepreneurship is an individual, learned quality. Mises (1949: 570) gives an example where all prices are "forgotten" one day, and then must be rediscovered over time. If entrepreneurship did not have some degree of learning required, prices could be rediscovered almost instantly (after the first trade of each item). However, instead we could expect that a significant time-lag would exist, as prices are searched for and demand and supply conditions are re-learnt.

Mises' promoter-entrepreneur than takes on a dual-role in the market which he sums up as follows: "The task of the entrepreneur is to select from the multitude of technologically feasible projects those which will satisfy the most urgent of the not yet satisfied needs of the public. Those projects for the execution of which the capital supply does not suffice must not be carried out" (Mises 1951: 17).

We see the first role is to look to the uncertain future and ascertain what demands will prevail. Second, the promoter-entrepreneur must look to the present to see what resources are available so that they can properly judge what resource constraints will allow them to produce an expected future. It is not solely enough that an entrepreneur may see a demand in the future, it is also a necessary condition that they have the present ability to move towards that future state. For instance, it would be easy today for an entrepreneur to know, almost with full-certainty, that a demand for a 10 carat diamond will exist in the future. However, lacking the available resources, the entrepreneur may see this as not a decision that will move the market toward a fleeting equilibrium – they can dedicate all the resources they want to the expected end, but if they don't think the end will be reached, the effort will be futile.

A second set of conclusions we may draw is that Mises' entrepreneur is gravely concerned with the uncertainty of a situation, but decidedly less so concerning risk. In fact, as he (1949: 810) points out, entrepreneurs do not invest in the projects expected to lose them the least yeild, but the projects they expect to earn the most. When they undertake a plan, they are concerned with the outcomes which the think are possible in the future. However, no heed is paid to probability – the risk-based assessment that concerns the expected losses that are possible, as well as gains. Hence, we see that the Misesian entrepreneur fully expects that their planned investment will net positive results in the future when it is realized – they pay no heed to the expectation that a loss may obtain instead.

The Kirznerian Entrepreneur

Kirzner's (1973) entrepreneur is an individual who is alert to the possibilities presented before them. However, this viewpoint is not concerned so much with searching for new information in general. Instead, the Kirznerian entrepreneur searches for “something someone has overlooked previously” (Kirzner 1997a: 71). Thus, the entrepreneurial element can best be summed up as:

> The entrepreneurial element in the economic behavior of market participants consists... in their alertness to previously unnoticed changes in circumstances which may make it possible to get far more in exchange for whatever they have to offer than was hitherto possible. (Kirzner 1973: 15-16)

The source of these profit opportunities is the pricing system, which acts as a signaling mechanism to inform the entrepreneur of disequilbira. As a result, the Kirznerian entrepreneur is embedded in catallactic situations, unlike the more praxeologically defined Misesian counterpart. Profit opportunities are then revealed in three ways. First is the recognition of previously made errors. These errors already exist and have yet to be acted upon to exploit profit opportunities. Second are new opportunities introduced through exogenously imposed changes. As alterations in the current state of affairs are altered, disequilbira are introduced requiring attention. These are different from the first type only in the temporal element of occurrence. Last, the uncertainty of the future leads to perceived opportunities that will await exploitation at a later temporal date.[71] Note that these three profit sources are differentiated by the temporal realm in which they occur – past, present, and future. Entrepreneurship is not a resource when viewed this way.

The Kirznerian entrepreneur contrasts the ever-present speculative nature of the Misesian entrepreneur, with the over-emphasis on economization that Robbins stressed. At the same time, however, Kirzner seeks to *de*emphasis the speculative nature of action. This is done mainly for reasons concerning the connotations attached to the phrase “speculative.” Kirzner recognizes that all action aimed at an uncertain future must be speculative in nature, the difference which should be remembered is that the entrepreneur actually *believes* that a profit opportunity exists (Kirzner 1973: 86-87). We see this actually mirroring Mises' true thoughts, if concealed, as he also reckoned the entrepreneur undertook an aspect of certainty when they made decisions concerning the uncertain future. Mises' entrepreneur was, it should be remembered, not concerned with the amount of money they would lose,

[71] See Kirzner's later works for this point of view, especially Kirzner (1982b and 1992).

but rather the amount they stood to gain. For the entrepreneur, they not only perceive a profit opportunity to exploit, they also *believe* it to be reasonable that they will attain this goal.

The future is created through the discoveries made by the entrepreneurs. Kirzner would even go so far as to state, “I have used the terms 'discovery' and 'creation' interchangeably” (1989: 40). One implication, which has been maintained by Kirzner throughout his career, is that the entrepreneur is able to discover opportunities that are “obtainable for nothing at all” (1973: 78). As ownership and entrepreneurship are completely separate functions, the role is now served by a *pure* entrepreneur who automatically excludes asset owners from the definition. As Salerno (2008: 190) points out, Kirzner was the first economist to make this functional distinction between ownership and entreprenership. Upon publication, *Competition and Entrepreneurship*, was met with much praise, although three reviews all expressed reservations about this distinction between ownership and ability (see Hazlitt 1974; and Rothbard 1974).[72] Kirzner's (1985: 44-45) response to these criticisms suggest that he understood the main-point as concerning the fact that the pure entrepreneur did not incur a chance of uncertainty or loss.

However, the failure of the Kirznerian entrepreneur to be intertwined with risk or uncertainty was really a secondary group of responses to Kirzner. As there is no distinction between arbitrage and uncertainty-bearing, there is no link to a temporal element apparent. For example, arbitrage necessarily deals with price differentials that exist in the present. However, as we have seen, although uncertainty may exist in the intra-temporal cases, it is also more manifestly apparent in inter-temporal situations. Thus, by stressing the resourceless arbitrage opportunities that an entrepreneur may discover, inter-temporal uncertainty and risk factors are de-emphasized. As a result, Kirzner's entrepreneur may only explain entrepreneurial gains, never losses.

Kirzner's entrepreneurial ability is not embedded or directly employed in a decision making process. Instead, it is almost accidentally utilized without any regard for its actual self (Kirzner 1979: 181). Hence, entrepreneurial profit is not a type of rent, or return, on its own ability. Instead it is, like it was for Mises, a form of pure arbitrage. As Sautet (2000: 61) comments, “Kirzner's entrepreneur is being alert. This is not incompatible with the Misesian entrepreneur.” As a result, there are two roles the entrepreneur is seen as fulfilling on the market. The first is the management of the arbitrage opportunities that were previously unknown. These are made available through the second role, the

[72] Additionally, High (1982: 166) notes, the separation of ownership and ability fully eliminates the concept of entrepreneurial losses. The very phrase “entrepreneurial loss” is eliminated as it is not possible for true losses to fall on any other than the resource owner.

attainment of information that was not previously known to anyone on the market. We do see a difference developing as Mises (1949: 253) viewed every market actor as not just a passive observer, but an active entrepreneur. Kirzner (1973: 43) would later argue against this, taking the stance that there exist two separate types of actors, pure entrepreneurs, and pure Robbinsian maximizing producers. Shostak (1999: 68) insightfully points out that two separate the two functions would separate the actors into two realms; "mechanical and conscious." That the two cannot be separated is indisputable, humans are not profit maximizing automata, but conscious, purposefully acting beings.

The entrepreneur is inseparable from action, and, as Hülsmann (1997: 33) points out, the notion of action is inseparable from private property. Kirznerian entrepreneurs are, however, essentially resourceless actors. They do hold an important element privately, at the exclusion of all others in the economy – action-knowledge. The action that entrepreneurs undertake creates, and discovers action-knowledge over time. The inter-temporal element is what necessitates the entrepreneur in the market place. They are the actors who search the future, see the disco-ordination that they expect will exist, and act accordingly.[73] Hence, they mitigate some of the uncertainty inherent in the Knightian "fog" of the future.

Despite the knowledge component that the entrepreneur may have, the resourceless assumption still remains a lacking point. Additionally, alertness to new opportunities is an imperfect quality to endow in the entrpreneur, *if they are to be viewed as an equilibrating factor.* First, alertness to new opportunities says nothing about the profitability of these opportunities. It is only through the use of owned resource that the entrepreneur may be able to determine entrepreneurial profit and loss, and hence, see if their discovered opportunities are truly profitable. Second, alertness may not be a universal of human action – as Kirzner himself realizes as he divides individuals into two camps (pure

[73] As Loasby (1982: 224) notes, the difference between the Schumpeterian entrepreneur and the Kirznerian is that the former earns profits through disruption, the latter through cohesion. He has also noted (1989: 178) that "[w]hereas Kirzner's entrepreneurs respond to changing data, Schumpeter's cause the data to change." Kirzner (1998: 14) would later reconcile the two disparate viewpoints:

> Consider the invention and innovation of the automobile in the U.S. This innovation, we may be sure devastated the livelihoods of many who had built their entire careers around the horse-drawn carriage industry. Virtually overnight, we may be convinced, enormous loss of value occurred in capital investments that had been made in that industry; large numbers of skilled professional workers in that industry find that the market value of their skills has fallen catastrophically. Yet, while understanding how Schumpeter can focus on the creative destruction which this successful and dramatic entrepreneurial innovation has wrought, I maintain that we must, at the same time, recognize the coordinative quality of this innovation, *even in regard to the horse-drawn carriage industry.*" (emphasis in original)

That we can only see Schumpeter's "destruction" *ex post* should become apparent when comparing the two situations. It is only through the consumer's revealed preferences however that this can be established.

entrepreneurs and Robbinsian maximizers). Additionally, Kirzner commonly referred to entrepreneurial qualities as "propensities" which seems to signal that he did not believe this quality was endowed in all humans, or at least, was not evenly endowed.

One additional shortcoming is the reliance on market prices. As prices may or may not be correct as per the underlying information represented, they may or may not be good indicators as to profit opportunities. If entrepreneurs are to function in a resourceless manner, and hence escape the possibility of pure losses, they will have no way to equilibrate these disequilibrium prices – losses, one half of the signals that assists in this process – profit *and* loss – will be missing.

Some modern viewpoints on the entrepreneur

There have been some interesting recent developments and refinements to entrepreneurial theory that deserve to be touched upon. Huerta de Soto (2004: 27) takes the view that it is the entrepreneur who creates the future possibilities for action, and provides us with the potential for increased growth.

Accordingly, the entrepreneur can be seen as achieving this in six ways:

[1]The generation of new ideas. Constraints that define action in one period are erased or expanded by the new information created through the entrepreneurial process.
[2] The nature of the entrepreneur's contribution is fundamentally creative. The opportunities that they find were latent, or non-existent before. In this way, these opportunities may not need to utilize existing resources for fruition (i.e., they are created *ex nihilo*).
[3] Entrepreneurship transmits information. This information can expand our future possibilities significantly without using any physical resources, it helps instead to coordinate people more effectively.
[4] The information transmission creates a coordination process that creates opportunities that were never previously seen, creating more possibilities for future action.
[5] Entrepreneurship is competitive in nature. This drive will ensure the trend is towards bringing forward the best possibilities, the ones that will serve people the most.
[6] Finally, entrepreneurship never stops. The previous five steps will continue forever. Action will continue forever, with entrepreneurship breeding opportunities for more entrepreneurship.

Unlike Robbins' view of the market actor as one who is necessarily constrained by the existing means-ends framework, the entrepreneur, acting as *homo agens*, is capable of realizing a new framework, one in which new, unforeseen possibilities can exist. Hence, entrepreneurs are not mere Robbinsian maximizers to the extent that the framework that they maximize within is constantly changing Kirzner.

Holcombe (1998: 54) writes that "entrepreneurship leads to more entrepreneurship." However, Hülsmann (1999: 64) notes that this self-reinforcing action arises as some profit opportunities are ignored in favor of others that are more valued to the economy. Hence it "creates some new opportunities, but at the same time destroys others." Entrepreneurship does not breed more of the same in the passive sense. Instead, acting humans seek out these opportunities and take advantage of them as they arise. Holcombe (2003: 7) defends his original statement, citing the development of one innovation as creating opportunity for others. The entrepreneurial process forever creates new possibilities as it results in a change in the present state of affairs which may not be wholly cohesive with consumers' needs. In this sense, entrepreneurs create new opportunities for other entrepreneurs. However, this creation must be viewed relatively. As Hülsmann correctly points out, there are opportunities created, while simultaneously ending opportunities in other sectors. As an entrepreneur upsets the plans of another in one sector, they may simultaneously create a new opportunity in a different sphere.

In an echo back to Turgot, Strigl (2000: 28) and Rothbard (1985: 283) both note the necessity the developed economy to work with the entrepreneur for progress to occur. Economic advance cannot occur absent capital, or money. It is only with the entrepreneur directing these resources for productive ends that progress is achieved. The entrepreneurial function, without any resources available, does not provide progress. Although it is not necessary for the entrepreneur to physically own resources, at least not in the Misesian or Kirznerian conception, their existence is still necessary to provide the data the entrepreneur acts upon.

Huerta de Soto (2005: 68) conceives the entrepreneurial process as taking on two elements: arbitration and speculation. The difference is the exercise in the present versus the future. Arbitration utilizes the means already available purely in the present to realize the potential entrepreneurial profit, with speculation taking on the more conventional inter-temporal role. Huerta de Soto notes that this distinction is, however, arbitrary; the primary entrepreneurial role remains identical in either case. The exercise of social coordination and the push to move the economy towards equilibrium remains the essence of each process.

More recently, Salerno (2008) has attached three primary roles to the entrepreneur. The first stems from the fact that to be an uncertainty-bearer in a meaningful sense, the possibility of losses must be available. Hence, they must be a capitalist who invests resources owned by them, and give rise to the possibility of lose. Second, as resources represent a heterogeneous structure, being transformed over time to create valued future goods, this resource control will likely be too complex in many situations for one individual to undertake. As a result, the capitalist-entrepreneur will also operate as a property owner, making the crucial decisions about the use of their property. Last, as there exist limits to the individual's skills and aptitudes (think of Mises' character traits), the decisions concerning the use of their resources will also involve the division of labor and knowledge, through the hiring of competent managers and employees to effectively complement and enable the capitalist-entrepreneur's capacities to flourish.

In any case, the definitions all share the point that the entrepreneur is the individual who looks to the future and sees the disco-ordination that they perceive to exist. This will necessarily be a subjective interpretation. In fact, the entrepreneur will likely be the only individual to notice this disco-ordination. If another had already noticed it, they would have moved to take advantage of it, if the benefits from doing so were deemed sufficiently adequate. The entrepreneur is an element inseparable from a process that involves a temporal element. Change never happens independent of action, in fact, change is a direct result of the entrepreneur's insight. As Rothbard (1962: 510) reminds us:

> The difference in the dynamic, real world is this. None of these future values or events is known; all must be *estimated*, guessed at, by the [capitalist-entrepreneur]. They must advance present money in a speculation upon the unknown future in the expectation that the future product will be sold at a remunerative price. In the real world, then, quality of judgment and accuracy of forecast play an enormous role in the incomes acquired by [capitalist-entrepreneurs]. As a result of the arbitrage of the entrepreneurs, the *tendency* is always toward the ERE; in consequence of ever-changing reality, changes in value scales and resources, the ERE never arrives.

Hence, as entrepreneurs continually exploit profit opportunities – both intra and inter-temporally – the market is moved towards a state of equilibrium. This final state never arrives as continual disequilibria are introduced to the system – endogenously, by shifting preferences, and developed opportunities – and exogenously, through situations which are by their nature fundamentally uncertain regardless of the entrepreneurial attributes aiming to bring certainty to them. The points we have just looked at represent modifications or additions to an underlying framework. Each author has implicitly started with either a Misesian or Kirznerian entrepreneur, and altered their function in some manner as a response to criticisms. However, it should become clear by now that the underlying basis of each entrepreneur has flaws which make their use as a starting point not fully suitable. Some of these fundamental failings will be addressed now.

The Present State of Entrepreneurial Theory – Some Issues

It is not enough to just say that an entrepreneur looks to the future with the eyes of an historian, as does Mises. Likewise, for Kirzner, judgment of profit opportunities in the present are not enough to move the economy forward towards any final resting place or equilibrium. Instead, what is needed are these profit seeking perceptions, coupled with the correct judgment as to their attainability. Kirzner tries to emphasis this aspect by removing the speculative role of the entrepreneur from his underlying Misesian framework. However, in doing so he also removes the inter-temporal aspect of the action, one which is essential for any concept of a fruitful entrepreneur. Mises tried to emphasize this by stressing the dual role served by entrepreneurs: that of looking for future disequilibria, and of determining if there existed sufficient resources to obtain this future goal. Hence, perception on its own cannot move the market to equilibrium, but *correct* perception must be emphasized.

As a result of the stress on correct perception, the possibility of the repercussions of *in*correct perceptions becomes evident. When dealing with a resourceless entrepreneur it is easy to see that the possibility of losses is, itself, lost. However, when we view the entrepreneur as having a qualitative aspect to their judgment (i.e., that they could err) and also that they personally have resources at stake in the process, we see that losses become a natural part of the entrepreneurial function, much like the more commonly stressed profits.

The distinction between entrepreneurship as a process, and the entrepreneur as a person becomes important. As an understanding of the *process* of entrepreneurship is stressed, we find that something more than a timeless arbitrage is necessary. As the end goal of entrepreneurship, as a process, is to remove felt uneasinesses and move the market towards an equilibrium (which never obtains), we see that the essential prognosticator is the entrepreneur (the person). The entrepreneur is however, in reality, not one individual with absolute qualities. Just like the division of labor exploits the heterogeneous productive capabilities of individuals, so to will entrepreneurs have heterogeneous aspects that they utilize in the entrepreneurial process. In fact, this difference – between an individual and a process – remained a source of bifurcation in Mises' own underlying theory of the entrepreneur. At one moment in time he (1949: 252-253) implies the process: "Economics, in speaking of entrepreneurs, has in view not men, but a definite function. This function...is inherent in every action...Action is always uncertain." At others he (1949: 333) stresses the individual qualities: "[Entrepreneurs] are the first to understand that there is a discrepancy between what is done and what could be done." It must be stressed that entrepreneur is building a future; the future they are building is uncertain in its source, but is certainly known to the entrepreneur building it; and that the entrepreneur hence sets in motion the action of entrepreneurship.

Mises and Kirzner's entrepreneurs, despite *prima facie* similarities, actually represent two polar opposite viewpoints. The Misesian entrepreneur looks to the future to see what opportunities they may exploit in the present. In contrast, Kirzner's prognosticator is alert to information that exists in the present, and gives rise to future opportunities. The temporal vantage point is reversed. Although syntheses between these two approaches have borne fruitful results, there is a fundamental unclarity which requires attending to. Both conceptions of the entrepreneur have advantages, but require a more solid basis from which they can be applied. It is with this in mind that we move towards creating a new framework for the process of entrepreneurship to progress within.

EMH and the Entrepreneur

EMH makes no explicit mention of the entrepreneur. As information is seen as not requiring a process to be spread throughout the economy, the entrepreneurial role is not necessary. The exclusion removes some important factors from the process that utilizes, and discovers, information.

Casson (1982: 14) sums up the entrepreneurial role thus:

> The entrepreneur believes that he is right, while everyone else is wrong. Thus the essence of entrepreneurship is being different – being different because one has a different perception of the situation. It is this that makes the entrepreneur so important. Where he not present, things

would have been very different.

EMH assumes that actors use information in a homogeneous manner. However, we see that the entrepreneurial function is not one of accepting what everyone else is doing, but in acting independently and differently. Information is not only heterogeneous in form, relevancy, and importance, but the manner that the entrepreneur deals with this information is also heterogeneous. Looking into the future, the entrepreneur only acts if they see a situation arising where there is a discrepancy between what they expect the market to require, and what the market is able to provide. Given the information that is available, and that which is serendipitously discovered, we see that the entrepreneur will move the market forward in a way that is not dependent directly on what others expect. In fact, it is because the entrepreneur has a different expectation of the future that they act – *if they shared the same beliefs, there would be no reason to believe that another entrepreneur had not already rectified the perceived problem*.

Prices are a result of this entrepreneurial action. They are not the cause of it, at least, not in a *direct* way. Always aware of the profit opportunities expected from their action, the entrepreneur uses prices in this way to give them an indication as to what profit they can expect from their action. However, the action they perform is spurred on by the appearance of new, heretofore unknown information, and how this information shapes their future expectations of profit.

EMH also operates under the pretense that all prices at any given moment are correct, *given the information that is available*. The view that prices must be "correct" in light of known information assumes that actors have an omniscience to them. However, we can see that the concept of "correct" prices at any given time is false. As Mises (1949: 337) demonstrated, all prices are "false" to some degree. This is so due to the constantly changing information, and preferences of consumers. A price can never fully reflect all the information and preferences as these will change before a price is established in *a dynamic setting*. It is only the entrepreneurial function that moves from more-false prices to less-false prices. The disequilibria created by false prices are corrected by the reference now provided through this same disco-ordination. As Hayek (2002) pointed out, false-prices arise as a result of entrepreneurial decisions that fail to grasp all the future implications of a present action. As the future is shrouded in uncertainty, and no individual could ever know the future with certainty, it follows that all prices must at all time be false to some degree. It is the entrepreneurial process that moves these false-prices to become more correct.[74]

Hence, by removing the entrepreneurial function from the pricing process, EMH has precluded any possibility that prices can be incorrect at any given point in time. They could not be so as there would be no way for them to become correct without the introduction of new information. However, as we see, the idea that information will be 100% correctly understood regarding presently unknown future conditions is erroneous. Bernstein (1999: 2) notes:

> You and I can disagree about the future, and your guess may well be as good as mine. None of us can ever be certain that we are right, but we can develop some confidence that we understand the situation better than others understand it.

An entrepreneur may not be correct at any given time. However, the primary function of the entrepreneurial process is to adjust prices that are incorrect, which they forever will be to some degree, based on newly discovered information. This information may be already available, and has only been previously misinterpreted. Alternatively, this information can be the product of the entrepreneurial function itself. Prices cannot be established absent human action. The entrepreneur, looking towards an expected future state, places in motion the process that establishes a price.

CAPM and the Entrepreneur

The CAPM view on time has been previously assessed. Time was viewed, in light of the model, as being static or Newtonian. What it uses as time is actually the antithesis of time. Time implies action, and action implies the entrepreneurial function. CAPM explicitly makes use of time, but the time in question is static, it lacks real-change. This lacking dynamism is where the entrepreneur becomes so apparent.

The relation that should exist between risk and return in the CAPM is not mechanical. Instead, it is the result of human knowledge of a piece of information that moves these two variables in a

---

[74] As Mises (1949: 337) points out, "The essential fact is that it is the competition of profit-seeking entrepreneurs that does not tolerate the preservation of *false* prices of the factors of production. The activities of the entrepreneurs are the element that would bring about the unrealizable state of the evenly rotating economy if no further changes were to occur." The elimination of false prices is through future-looking entrepreneurs operating in the present (Mises 1980). The *false* price could only exist under two conditions: the lack of the entrepreneurial correction process, and the temporal element necessary for this process. That the second is a precondition for the first should go without saying. See also Hayek's (2002), *Competition as a Discovery Procedure.*

cohesive manner. As such, we can see that an entrepreneur will not take on more risk than the expected return can justify to them *personally*. As a result, entrepreneurs buying and selling assets in the present is what moves the return that a security yields. CAPM assumes that actors are passive price takers, which is in stark contrast to the Kirznerian entrepreneur previously looked at who was a price setter.[75]

The removal of the entrepreneurial element removes all subjective interpretation of market data. Much of the empirical failure of the CAPM can possibly be explained this way. For instance, any piece of information becomes a person's knowledge through a subjective process that brings it from the external world and into their mind. This process ensures that, although similar, no two pieces of knowledge that two separate people have can ever be identical, *even if based on the same piece of physical information*. Likewise, the entrepreneur's expectations of the future will never be the same as any other's.[76] This is due to the subjectivity of expectations, coupled with the fundamentally uncertain future.

If the market return is *w*%, beta is *x* and the risk-free rate is *y*%, the expected return is not a mechanical *z*%. The expected return will instead be a number of different *z*%s, all conditioned by two things.

The first is the information that the entrepreneur currently has at their disposal. Not all will have access to the same information. This implies that not every entrepreneur will have the same expectation as to the present information. The variables that enter into the CAPM must be derived, they cannot be given.[77] One entrepreneur may derive beta from pieces of information *a* and *b*. Another may derive the exact same beta from sources *c* and *d*. There is no way to state that any individual will use the same market data to derive the variables necessary to compute the CAPM.

Additionally, we see that expectations of the future condition every entrepreneurial action. The arbitrage inherent in the inter-temporal aspect of life are what spur entrepreneurial actions. However, due to the uncertainty of the future, we see that expectations are forever bound to be individually based. There can be no homogeneity where subjective interpretations of heretofore unknown market

75 As Lavoie (1985: 83) states the nature of the problem, "market participants are not and could not be price takers any more than scientists could be theory takers... Entrepreneurs (or scientists) actively disagree with existing prices (or theories) and commit themselves to their own projects (or ideas) by bidding prices up or down (or by criticizing or elaborating existing theories."

76 This point is one reason why entrepreneurs can mitigate future uncertainty. As Hoppe (2007: 14) points out, each entrepreneur will have distinct knowledge better than anyone else as to their own knowledge, and hence, be able to better understand and predict their own personal actions.

77 The exception may be the risk-free rate which can be seen as the least risky rate that the individual can add to their portfolio. However, as will be shown later, several considerations preclude this from being given, and actors must derive this based on their own knowledge as well.

information are involved.

These two separate factors serve a fundamental role in the entrepreneurial process that drives prices: the present data held, and the expectations of the future. We see that neither of these factors can been identical for any two individuals, let along homogeneous for all market participants.

Hayek believed that prices served their informational role only when the market was in a "proximal-equilibrium" situation (Salerno 1993: 128). Prices do contain information, but it is only a portion of the information that is required of entrepreneurs. Hayek's stress placed on prices as carriers of information applies mostly, as Salerno points out, in an equilibrium setting. In the dynamic world we live and act within, we see that equilibrium is a fairy tale. As a result, prices may not be fully accurate determinants of future movements. CAPM, to be an effective tool, relies on current market prices as being faithful predictors of future conditions. We see that this is an unrealistic assumption in a dynamic setting.

Prices are not mechanically set absent of time, but result from the process whereby entrepreneurs move towards satisfying their subjective beliefs as to future market possibilities. These are not homogeneously shared throughout the world of actors. Nor is the body of information currently available purely homogeneous to all participants of the pricing process. We see that any theory of pricing that ignores these points will erase any attempt at subjective valuation. To the degree that CAPM denies the existence of the entrepreneur in the pricing process, the subjective element is removed, and the result is unrealistically suited for the true dynamic world.

Conclusion

Contra to Knight and Schumpeter, the entrepreneur exists and serves an equilibrating role. Mises stressed the importance of the entrepreneur as a planner for the future. They subjectively look into the future, see a disco-ordination, and act accordingly to remove it if it is deemed sufficiently profitable to do so. Kirzner's entrepreneur exhibits alertness to their surroundings. They see an opportunity and take advantage of it.

However, both EMH and CAPM preclude the role of the entrepreneur based on their very assumptions. The result is a pricing process that is unrealistic, and excludes any form of subjectivism necessary to describe our world.

EMH makes no explicit mention of the entrepreneur in the pricing process. This results from the removal of the concept of time from much of its theoretical foundation. The result is that information is seen to automatically, costlessly and timelessly be assessed, and prices adjusted accordingly. However, this assumption overlooks all the important processes that actually occur on the market and affect prices. It masks the true problem that concerns the economist, the collaboration of knowledge. Entrepreneurs must seek, and discover, knowledge – it is never given to them automatically. As a result, no two individuals' minds ever contain the same knowledge. The result is that the same interpretation of knowledge will never result. Prices do not arise, and change, due to the homogeneous interpretation of information automatically endowed in every actor. Instead, they change only by the changing expectations that entrepreneurs hold as to the future conditions, coupled by the way they assess the current situation. The removal of the subjective entrepreneurial element masks the true nature of the pricing process. It is not the price that is important, but the process that achieves that end.

CAPM on the other hand recognizes time, but the type of time recognized is not the correct one. The entrepreneur requires a subjective, dynamic time to operate. As a result, the role of the entrepreneur is masked in this theory as well. Prices are not set automatically in a mechanistic fashion such as the model supposes. Instead, an entrepreneur will subjectively decide that the risk they are shouldering is not offset by the expected return of a security. The result will be an action that will affect the current price. The process is fundamentally conditioned by two aspects: (1) the interpretation of the current situation that exists, and (2) the subjective expectation of the condition that will exist in the future. Both of these aspects disappear in CAPM. It is only by recognizing the subjective, continual role that the entrepreneur serves in pricing assets that a true grasp of the process can be discovered.

Sautet (2000: 14) reckons that the entrepreneur is excluded from much of the mainstream's analysis as a result of a failure to recognize the "Hayekian knowledge problem." The focus on all the information being objectively knowable in the present implies that there is no need for the entrepreneur to discover it. We see this to be more than apparent in both the EMH and CAPM. Also, the removal of a temporal element from both theories has eliminated the role of the entrepreneur to arbitrage between the present and the future. The entrepreneur understands action and can look into the future. They understand other actors' reactions and take this information to move into a future state (Mises 1978: 49). When this fundamental entrepreneurial role is eliminated, prices can be assumed to adjust correctly, instantly, and mechanically to any new information that the market creates. The problem that the pricing process faces, however, is that prices are never mechanically achieved. Instead, they are subjectively developed. Some may be inclined to state that there are periods lacking entrepreneurial activity. Eras with aggregate losses exceeding the aggregate profits may be viewed as lacking the entrepreneurial element necessary to drive the economy forward. However, even in a "retrogressing economy" we see entrepreneurial activity, only at a reduced scale (Mises 1949: 251). Any pricing theory must include the subjectivity inherent in entrepreneurial decision making if it wants to achieve a realistic, and faithfully true, result. To the extent that EMH and CAPM both exclude this possibility, their results are not representative of the world humans live within. As Pasour (1989: 102) so aptly puts it, "the driving force of the market system is eliminated."

4. Methodological Concerns

An interesting distinction occurred during the respective development periods of both CAPM and EMH. Early EMH developers seemed to be concerned that their employed methodology was not as optimal as it could have been. They had empirical results searching for theory. In contrast, we can find no such apologists during the same development period for CAPM. In particular, there are two broad groups of methodological concerns regarding the development of both concepts.

The first is that both ideas have utilized a positivist approach. This continues not just in their respective formulations, but in the continual testing they undergo to establish validity. The result has been, in both cases, ideas that worked for specific data sets in the past, but are now experiencing considerable problems retaining their validity. This concerns the nature of human action, and the constant change inherent in it. It is not possible to use one data set, or many data sets even, to try to extrapolate a theory for the ages. The result in the social sciences will always be theory applicable to only one specific episode of time.

The second methodological issue surrounds the use of mathematical reasoning. CAPM has always rested on a bedrock of mathematical concepts. These are found to be wholly unsuitable for the realm of human action. They have excluded concepts that are unable to be formalized such as subjectivity of interpretation, entrepreneurship, or a non-static conception of time. The result is that we have a mechanical concept, detached from the fluid world that humans act within. As Keynes (1936: 297) criticized this approach:

> It is a great fault of symbolic pseudo-mathematical methods of formalizing a system of economic analysis, that they expressly assume strict independence between the factors involved and lose all their cogency and authority if this hypothesis is disallowed... Too large a proportion of recent "mathematical" economics are mere concoctions, as imprecise as the initial assumptions they rest on, which allow the author to lose sight of the complexities and interdependencies of the real world in a maze of pretentious and unhelpful symbols.

The formulation of a pure theory of asset prices must lie on a solid bedrock of deductive logic. Time invariant concepts can be formulated that will apply to all situations equally. Furthermore, by removing the strict formalism of mathematics, we can proceed to open our viewpoints to more dynamic concepts.

Induction/Empiricism

The mainstream economics profession today hold that the prime methodological tenet be proof by induction. Induction is the process whereby it is reasoned that the premise of an argument is believed to be correct and support a conclusion, but cannot ensure it. Economists are able to infer conclusions based on a sample that they believe to be valid, but cannot be absolutely proven to be. Nature and the natural sciences serve the model for these economists. Seen as inferring data from the past to form a conclusion in the future, Austrians have eschewed this approach in favor of our deductive analysis.

Two related propositions characterize empiricists (Hoppe 1995: 9):

[1] Knowledge concerning reality must be verifiable, or at least falsifiable.
[2] Validity can never be established with certainty, but will always be hypothetically based on a contingent future.

The most influential type of induction used today is Bayesian. This approach uses probability theory to evaluate conclusions. An objective value for the probability that a model is correct is sought, with this factor providing and insight as to how likely a conclusion is.

As induction focuses on past events, and tries to extrapolate their results into the future, it can be seen that it is objectionable in the realm of the physical sciences. Mises (1949) would make this distinction, noting that something that occurred in the past to a human may not continue into the future, or, it may continue but the human's reaction could change unpredictably. Thus methodological monism is unacceptable. Sciences of the social realm must adopt a method distinct of that used by natural scientists. Bayesians are criticized as they try to achieve some objective probability regarding their model's success, but due to the uncertainty of the future, there is no real objective way to do this. Instead, an economist would be forced to assign a random, or subjective, probability to their conclusion. Huerta de Soto (2006: 386) criticizes the Bayesian insight as being inapplicable to the field of entrepreneurship, or human action. The mistake in misidentifying case probability for class probability, as demonstrated by Mises (1949: 110) is a key point in deconstructing Bayesian theory.

Empiricists must all agree that humans can learn from experience, otherwise, why undertake any empirical work? However, as any theory will be one based on past events, will it not describe a situation, or a human mind, which will be markedly different than one which could exist at a later date? Empiricism can only historically analyze something, it can never state anything meaningful about the future. As people can and do learn, it follows that empirical knowledge is contradictory when applied to the realm of human action.

Mathematics

Within this empirical realm, a movement started, and has taken hold, that views mathematical analysis as the proper approach economists should take. This approach has seen an increase in practitioners, and there remains no doubt that the approach favored today is mathematical-based inquiry. Backhouse (1998) shows us that between 1930 and 1980, two leading academic journals saw an increase from 10% of their articles being math-based, to over 75%. This percentage has only increased since then, with articles using pure-theoretical verbal logic becoming increasingly rare in the mainstream journals. Simultaneously with this increase in mathematical articles, there is a shift towards an increased emphasis in the University system to develop students apt at this type of research.[78] It has been noted that this shift was been a rational response to a demand from private-sector practitioners for this type of graduate.[79]

The mathematical movement was put into motion with Leon Walras and Augustin Cournot as they both furthered the systemic application of mathematics to our field.[80] In fact, Walras (2003: 43) would go one step further and assert, "it is only with the aid of mathematics that we can understand what is meant by the condition of maximum utility." Jevons would also help to bring mathematics to the forefront in the English speaking world, solidifying its dominant role in economic analysis.[81]

The 'infection rate' of mathematical analysis was, as could be expected, initially quite slow. Marshall (1890) in his *Principles*, for example, relegated mathematical inquiry to footnotes, preferring verbal logic for the bulwark of his work. Despite his shortcomings, Marshall held a quite limited view

---

78 It should be noted that the importance is not even placed on developing mathematical models that work! Black and Scholes, for example, won the 1997 Nobel prize partly for their ill-fated finance model.

79 It is outside the realm of the present paper to debate whether this is a truth, or if it is that more likely an increased demand exists for graduates who excel in real-world situations. Mises had commented that the increased mathematization of economics was due to statist influence.

80 Although early examples can be found scattered, for example, the first such example may be Quesnay's 1758 *Tableau Économique.*

81 See Jevons (1871) for this introduction. In particular, on page 27, "if only commercial statistics were far more complete and accurate than they are at present, so that the formulae could be endowed with the exact meaning by the aid of numerical data."

on the potential for mathematical economic analysis, and much like Austrians of his time and today, viewed economics as a study of a process, not some factitious equilibrium state. Likewise, his student, John Maynard Keynes, had a quite limited opinion as to the application mathematics would play in economic science.

Mathematical economics today has bifurcated into two separate fields: pure and applied theory. Pure theory focuses on furthering existing proofs, while the applied branch takes on a data-fueled approach. The two have increasingly departed paths, with pure theory being separated from the applied part even within individual articles.

Some economists have questioned the role mathematics should have in the field. The Keynesian Joan Robinson, for example, is famous for advocating the limited role it should have. The common question that mainstream economists (such as Robinson) raise against mathematical economists is to which extent human behavior is able to conform to deterministic principles. Even viewed stochastically, one general view has been that the approach has limited appeal with regards to individual matters (i.e., micro), but is generally valid when viewed in a wider, economy wide context (i.e., macro).

The shift to mathematical modeling has been the result of the success that the physical sciences have seen using it, particularly physics.[82] However, in making the methodological shift from the natural sciences to the social one which economics is part of, the question was not asked if the approach was transferable. Physics, for example, is a science based on the measurement of values. There is no question that this can be achieved in cardinal terms relative to other physical objects. Hence, we can say a piece of wire is *x* units long, with one unit being defined by some other physical object. Mises (1949, 208) would note that countable and measurable quantities (i.e., those which mathematics can manage) can only deal with objective value, never subjective.

However, measurement in the realm of economics, that which concerns with human actions, cannot be measured so simply. In fact, no measurement can take place, only ordinal ranking. When we say how much a person values an object, we can see that at a given moment, it will not have a purely measurable value. Instead, only an ordinal value can be achieved. When I buy one coffee, I can only state that I value $0.70 less than one coffee. But this value scale is not constant, instead it undergoes constant change. In the next moment, a coffee may not be worth $0.70 to me, and so I will forgo having another one.

[82] In fact, many economic terms today are 'borrowed' from physics: static, dynamic, equilibrium, etc.

In the natural sciences, we can say that equality exists. One foot equals 12 inches. But in the realm of human action, with subjective value scales, only inequalities exist. One coffee is worth more than $0.70, and this inequality is time-variant, the next moment one coffee is potentially worth less than $0.70. Furthermore, human value scales are influenced by other factors. One person buying a coffee for $0.70 may influence and alter my value scale to do likewise.[83] In the physical sciences, one foot will always be equal to 12 inches, independent of what influence it is under. This is, in part, due to a standard, absolute measuring unit available for the physical sciences. If a ruler is used to measure a piece of wire, and the result is questionable, a different ruler can be used to verify the result. Such an absolute standard does not exist in the physical sciences. If we measure a person's value scale, and something seems amiss, due to the subjectivity of that scale, we cannot turn to another person to delineate it for us.

Finally, it should be noted that using a mathematical model assumes that the language of human action can be translated into math. This is a highly questionable proposition. There are many concepts in life that defy translation, especially into mathematical terms. Even if human actions were able to be adequately translated into the language of math, it is still questionable whether this would be the most optimal approach. Results must be translated back to our language so that we may act on them. Just like when any translation takes place, there is a risk that something will be lost along the way. Leoni and Frola (1977: 109) sum this position as they state "translation of ordinary language is therefore not necessarily the most suitable way of dealing wit the empirical problems of human beings in the real world." Huerta de Soto (2005: 52) would agree by viewing the language that humans act upon as being tacit, and subjective in nature, thus, unable to be translated or articulated by anyone else. In fact, this knowledge may not even be translated into words, instead only being spread by the actions of the individual.[84]

Friedman and Assumptions

Arguably, the most influential article about methodology for mainstream economists in the 20th century

[83] Crusoe economics allows us to mitigate some of this 'influencing' effect, however, it can never be fully eliminated. It is not just human actors that influence human actions. The fact that it is raining one day would cause me to value an umbrella more highly than I did the previous sunny day.

[84] Rothbard (1997: 137) points out that "[i]n the market economy, qualitative knowledge can be transmuted, by the free price system, into rational economic calculation of *quantitative* prices and costs, thus enabling entrepreneurial action on the market." It becomes apparent that the entrepreneurial process can interpret, at least some, *qualitative* knowledge into a transferable *quantitative* form.

was Friedman's (1953b) *The Methodology of Positive Economics*.[85] The gist of the article laid out the foundation for what an economic method should be concerned with; precision, scope, and the track-record of a model's predictions. Economics should be concerned with objective data, and this positive approach should be closely involved with normative implications.

The goal then of a positive science was, to Friedman (*ibid.*: 7), “the development of a 'theory' or 'hypothesis' that yields valid and meaningful (i.e., not trusitic) predictions about phenomena not yet observed.” This theory would serve as a sort of “filing cabinet”, providing only tautologies that make the necessary abstract deviation from a complex reality. The comparison of a theory to the predictive value it had in the future was seen as the only true test of validity. He, much like Popper, viewed the facts as only being capable of disproving a theory, never proving it outright. Hence, a theory remains perpetually unproven, always waiting for the next test to prove it is not invalid. The positive approach must be more than just purely mathematical based (i.e., it must contain analytical theory as well), and would depend on empirical data at two key stages: constructing a hypothesis and testing its validity.

Most troubling was Friedman's view on the need for assumptions. For a theory to be useful, it would yield a lot of result, from a relatively small theory. Assumptions necessarily simplify theory. However, if assumptions were based on realism, they would complicate matters, therefore, assumptions should be “descriptively false.” He (*ibid.*: 15) would sum up by stating:

> the relevant question to ask about the “assumptions” of a theory is not whether they are descriptively “realistic,” for they never are, but whether they are sufficiently good approximations for the purpose at hand. And this question can be answered only by seeing whether the theory works, which means whether it yields sufficiently accurate predictions.

Furthermore, models should be very exact in objectively-based language, an object would be *x* lbs, for example, not merely heavy.

For Friedman then, economics is a purely positive science. The reality of the simplifying assumptions was not important, in fact, as an economic theory could never be tested in reality, the assumptions themselves could never be tested. Hence, the only way to truly test a hypothesis is continual empirical trials.

The view of economics as being a positive science has already been dealt with. Friedman's view

[85] It is unfortunate that while the general importance of this article to modern economic theory may be reviled only perhaps by Keynes' *General Theory*, as Hayek (1994: 145) tells us, it is also “as dangerous” as the *General Theory.*

of assumptions and hypothesis testing are troubling however.

Theory must describe reality. That is the goal of theory. Reality is complex, and simplifying assumptions are necessary sometimes to make theory workable. However, to state that the reality of the assumptions do not matter misses the point. Rothbard (1997: 102) tells us that "false assumptions are useful in economic theory, but only when they are used as *auxiliary* constructs, not as premises from which empirical theories can be deduced." The trouble lies with Friedman's view of the predictive value of a theory as being its true test. Empirical results can always agree or disagree with a theory, sometimes simultaneously. As Coase is known to have said, "if you torture the data long enough, it will confess". But if theory is to be tested on the empirical reality, how can we trust that reality to be true? This is where the trouble with assumptions manifests itself. If we do not even potentially test a theory against an objective reality, why should the assumptions themselves describe a reality.[86]

Theory and History

The above mentioned approaches are all concerned with using historical data to dictate theories about future action. The precedence for this type of approach in economics is long established. However, many writers have seen issue with using history to create theory, and instead advocate history having a different role to play.[87]

History can only concern itself with past events that individuals instigate. Analysis and recording of action is a task for history, not for praxeology. Robbins (1952: 38) would make a distinction stating that, in regards to the two fields, "theory describes the forms, economic history the substance." If we view economics as a realm concerned with human action, we must start by looking for the universal truths that govern our action. But these cannot come from our external world, but from within us. Hence, Mises would contend that introspection into our *a priori* knowledge (i.e., that preceding experience) is the proper way to deductively formulate economic theory. These self-evident truths are equal for all humans, hence, an absolute theory can be devised.[88] Empirical results can never

[86] Rothbard (1997: 102) would counter against the argument that Austrians, too, are capable of using none-realistic constructs – Mises' ERE for example. However, as he tells us, "[o]nly by analyzing a fictive changeless state can we arrive at a proper analysis of the changing real economic world." For a thorough critique of Friedman's theory on assumptions, see Long (2006).

[87] "New experience can force us to discard or modify inferences we have drawn from previous experience, but no kind of experience can ever force us to discard or modify a priori theorems. They are not derived from experience; they are logically prior to it and cannot be either proved by corroborative experience or disproved by experience to the contrary" Mises (1976: 27).

[88] Synthetic a priori truths are self-evident and need no observation to prove. This derives from the fact that the denial of

refute them, if a result does not agree with our a priori theory, our theory is not defective, but a condition of our theory may be. The only method for faulting a theory based on a priori truths would be to look at the deductive logic step by step until a flaw can be found. Empirical results on their own can never achieve what logic can. Thus, the mainstream confuses the concepts and roles of theory of history. As Hoppe (1995: 9) sums up:

> A distinction exists between a historical and a theoretical explanation only insofar as a historical explanation refers to events that have already happened, something that lies in the past, whereas a theoretical explanation would be an explanation, or rather a prediction, of an effect that has not yet occurred.

Or, we can turn to Friedman's own view on the praxeologic method:

> That methodological approach, I think, has very negative influences... [It] tends to make people intolerant. If you an I are both praxeologists, and we disagree about whether some proposition or statement is correct, how do we resolve that disagreement? We can yell, we can argue, we can try to find a logical flaw in one another's thing, but in the end we have no way to resolve it except by fighting, by saying you're wrong and I'm right. (as quoted in Long 2006, 19)

Friedman confuses a priori reasoning with subjective thinking (Long 2006: 20). It also appears as though he confuses praxeology with rhetoric.

Starting with the axiom that "humans act", deductive theories can be created with auxiliary knowledge that can be gleaned from this. That humans act to remove felt uneasiness, that we use means to attain ends, and that we prefer leisure to work are all a priori knowledge that flow naturally from the axiom of human action. Mises (1949) would term this method praxeology, the science of human action, of which economics was only a subset of. From the action-axiom, Mises would deduce the categories of action that are also a priori; values, ends, means, choice, preference, cost, profit, loss, time and causality. However, to understand what these mean, a person must act. Someone who does not act could not understand what these concepts imply. Furthermore, more complex laws have been

their truth leads to self-contradiction. Hence, that humans act is self-evident, as the action of denying this is in itself, purposeful behavior. Also, no observation is necessary to see this. See Hoppe (1995: 5).

formulated from the action-axiom; the laws of exchange, of diminishing marginal utility, the Ricardian law of association, the law of price controls, and the quantity theory of money.[89]

Rothbard (1957: 314) would summarize Mises' praxeological method as being four steps:

[1] There are fundamental axioms and premises of economics that are absolutely true.
[2] The theorems and conclusions deduced by the laws of logic are therefore absolutely true.
[3] There is therefore no need for empirical testing of assumptions or conclusions.
[4] Deduced theorems can not, and need not, be tested empirically even if it were desirable.

Hoppe (1989) would describe a three step approach that economists should undertake, based on a Misesian praxeologic approach.

[1] Understand the categories of action, and what the meaning of a change is.
[2] Describe a given situation within these categories of action, specifically, finding the specific meaning that individuals have when undertaking specific actions.
[3] Logically deduce the consequences that must follow from this situation.

If no flaw can be found in your deductive process, your result must be true a priori. If you must introduce assumptions into your theory, it will still be a priori true, but only for a given field. The only way to refute such a claim is to go through the deductive steps, and locate a mistake. If none can be found, it must hold true, *no empirical evidence to the contrary could prove otherwise*.

Additionally, methodological individualism is to be employed. All action, even large aggregates, are the result of individual's conscious actions. Hülsmann however argues that individualism is a tool for history, not for pure theory. For instance, in viewing historical aggregates, we can reduce them to smaller individual units until we find the exact causes. Praxeology, or theory, is concerned with describing invariant consequences of actions. To this end, methodological individualism tells us much of individual action, but little in regards to aggregative action as a whole. For example, individualism will tell us little of a general money supply increase, or the general division of labor. However, in historically understanding these events, it is quite relevant. Further, Hülsmann (2003: 70) argues that human action is comprised of two parts: the realized, and the unrealized.

[89] Hoppe (1995) would 'discover' an additional axiom, the *argumentation axiom*. Like the action counterpart, one cannot argue, that one cannot argue.

Realized through the actions of our bodies, and the activity of our minds. Unrealized in the ends sought after, and the foregone alternatives. The realized portion makes up the visible part of choice. This comprises the bulk of economic law. The existence of unrealized choice, or counterfactuals, allows the economist to hone theory, and create a more exact science.

If the mainstream and Austrians find agreement on some stances, it is merely a superficial agreement. The mainstream may believe their theory to be true based on well-tested facts, when, in fact, it would be true due to a priori propositions. The history of economics is littered with the ruins of methodologies which have not stood the test of time. To think that some of the more prominent ones of today will not suffer the same fate is unthinkable. Praxeology, based on time invariant causes and a priori deductions will be as valid 1000 years from now, just as it was valid 1000 years in the past. As Hülsmann (1997: 25) points out, what is true of logically deducted axioms is true for reality as well. The method of praxeology is sound. New methodologies are sure to continue being developed, but it will be the role of praxeology to remain the one true method for the science of economics forever.

Conclusion

Theory cannot be based on empirical results. Empirical results are what are interpreted through theory. The finance world has followed the inductive method, highly influenced by Friedman (1953b). We see however that the conclusions that have been drawn from this method are controversial, and quite often, erroneous.

Furthermore, the highly mathematical method used, in particular for CAPM but also dominant in EMH development, is of questionable validity. The assumption that humans exist within a static framework precludes the concepts of self-determination, or learning. Humans learn through experience, and adjust behavior accordingly. Thus, they cannot be seen as time-invariant automatons in all regards. They are conscious, and as such, unable to be viewed in a vacuum like the specimen of the physical world. The use of mathematics masks the true issues that confront economists. For instance, as Baumol (1968: 67) argues, the entrepreneurial function is hid through mathematical models. Also, the fundamentally uncertain future element is unknown at the present, and hence, cannot be reduced to a mathematical statistic. These issues preclude a purely mathematical method from being conducive to discovering, and explaining, the issues that confront economists. As Mises (1977: 99) most poignantly put the issue, "[i]n the field of human action, however, there are no such constants. The equations of mathematical economics are therefore useless for all practical purposes.” Or, as Rothbard (1996: 352) explained the trade-off:

> Economic theory, in short, must choose between formally elegant but *false and distorting* mathematical models, and the 'literary' analysis of real human life itself. [emphasis added]

Empirical results can never refute deductive theory. Instead, deductive theory can explain empirical results so that humans can gain a deeper understanding of the world. An economist's observation of 'randomly' moving prices is not cause to conclude that prices are random. Instead, it is cause to develop a theory to explain why prices were moving randomly in that situation.

There is a distinction to be made between theory and history that is crucial. History is where we have been, it is the results of our actions. Theory is a wholly separate realm. The theorist can function in the absence of all history, but the converse is never true. An historian without theory would be akin to a driver with no map. The search for the final destination would be hopeless, directionless, *unscientific*. It is only through deductive logic based on a solid praxeologic foundation that we can achieve the time-invariant theories required to analyze human action.

Lachmann (1973: 204) would view the role of economics as two-fold:

> Economics has two tasks. The first is to make the world around us intelligible in terms of human action and the pursuit of plans. The second is to trace the unintended consequences of such action. Ricardian economics emphasized the second task, the "subjective revolution" of the 1870s stressed the urgency of the first, and the Austrian school has always cherished this tradition.

The role of the academic to to make the world understandable for the practitioner; in Lachmann's words, to make the world intelligible in terms of human action. It is difficult to say whether either EMH or CAPM have been successful in that regard. It does not take a long glance in the *Wall Street Journal* to discover there is a giant disconnect between what it is saying *does* occur, and what the *Journal of Finance* says *should* occur. Take for instance this quote from the April 1995 edition of *Fortune* magazine:

> [Noted Harvard financial economist Michael Jensen writes] “there is no other proposition in economics which has more solid empirical evidence supporting it than the Efficient Market Hypothesis,” while investment maven Peter Lynch claims “Efficient markets? That’s a bunch of junk, crazy stuff.” (As quoted in Clarke, Jandik and Madelker 2001: 2)

Making the world intelligible necessarily means using relevant theoretical devices. Irwin Friend's (1973: 272) conclusion to his 1972 Presidential address to the American Finance Association seems to have fallen on mostly deaf ears: “Much of [this Address] simply represents support for the position that methodological elegance should not be considered a substitute for substance, and by substance I mean solution of real world and not artificial world problems.” We see the fallacy in Friedman's (1953b) assertion that assumptions can, and should, be made unrealistic for proper study. It is only be making our analysis mimic the world as closely as possible that we can gain insight into how it works. The completion of Lachmann's second task can only be achieved through theory funded on realistic assumptions, a necessity of his first role.

To the extent that many of today's finance theories are based on erroneous assumptions, or faulty *empirical proofs*, their conclusion must be found equally invalid. A reassessment is needed for any theory that today is built upon this questionable past, and hence, start anew with a solid foundation from which to rebuild the science.

II. THE EFFICIENT MARKET HYPOTHESIS – AN AUSTRIAN PERSPECTIVE

In chapter I we looked at four common fallacies in both EMH and CAPM. We can now shift our attention to three specific problems that arise in EMH.

The first issue to be examined is the concept of information. Information is not created and spread throughout the market instantly. Instead, there is a distinct temporal element to this process. Furthermore, the role of information in creating action is looked at. With no specific temporal element in the EMH, the problems of knowledge and the market are erased away. Instead, as Hayek made so clear, the treatment of knowledge is one of the gravest problems facing economists.

Second, the concept of efficiency is reviewed. EMH assumes an idea of static efficiency to hold true. We see this misplaced in light of Huerta de Soto's (2004) concept of dynamic efficiency. The idea that any static efficiency could be applicable for humans is criticized with the dynamic alternative offered. This particular point is also seen as a direct result of the earlier erroneous treatment of time.

Lastly, the specific determinants of price are given a closer look. EMH assumes that the only factor that can alter future prices is the arrival of new information. However, we see there are sources endogenous to the actor that serve to shift our value scales, and hence, create price alterations. We see that information is not the sole creator of price changes, but instead is in many instances, the result of price changes.

EMH faces increasing empirical questioning. However, to solidly refute this concept what is needed is not a general critique of the overall concept of efficient markets. Instead, by looking at the individual components that contribute to the hypothesis, we see that they are incompatible, and unrealistically suited for our dynamic world. Many models hold true given the assumptions that are incorporated into them. In the case of EMH, we find the assumptions to be wholly irrelevant for human action. The result is that this concept is inapplicable to a world with conscious, purposeful, human actors.

1. Information and Knowledge

The nature of the problem concerning EMH is knowledge. Before we can even begin trying to determine what effects knowledge may have on action, we must define what knowledge actually is. There is a great deal of confusion surrounding this task. This is troublesome as the twin concepts of information and knowledge are at the forefront of economic research.

Four Nobel laureate economists have stressed the importance information plays in the economy.[90] Hayek posited that the greatest challenges for economists of future generations would be in the field of knowledge and information. Akerlof, Spence, and Stiglitz stressed the importance of asymmetric information, as well as the role that information has increasingly played in the global economy. However, 60 years after the birth of the field of information economics, a great amount of confusion surrounds the topic. A failure to distinguish between two similar concepts, information and knowledge, has contributed to this problem.[91]

Economics has typically focused on an approach centered on the concept of perfect information and assuming that although this is not descriptive of reality, reality would perform reasonably close to this paradigm. However, as Stiglitz (2002: 461) points out, "even a small amount of information imperfection could have a profound effect on the nature of equilibrium." This paradigm must be challenged, especially as it creates a focal point in EMH. By first providing a foundation of information, we can later determine through what process this building block creates, or contributes to, action.

Information and Knowledge

We can state that information is the available body of facts that can become known. It is hence finite at any given time. Knowledge, in contrast, is the information that we personally possess. Polanyi (1958b: 35) would differentiate between the two stating an, "invention ... does not produce something that was not there before; but actually, it is only the knowledge of the invention that is new, its possibility [or information] was there before." To Polanyi the distinction is slight, but critical. The knowledge given

[90] Hayek was awarded his Nobel prize in 1974; Akerlof, Spence, and Stiglitz would share theirs in 1991.
[91] The American Heritage Dictionary of the English Language (2006 ed.), for example, defines information as "[k]nowledge derived from study..." and knowledge as "specific information about something." This circular reasoning adds in no way to aiding the discussion.

by a new machine was only the collaboration of information that was already in existence.[92]

If information is given, we can see there are two distinct types. One is given by reason and logic; it always was and will be forever in a state of existence. For instance, Einstein's general theory of relativity falls into this camp. Although not being discovered until recently, this type of information was always available to us. The Pythagorean theorem would be another example. This information is unleashed through the exercise of our minds. The second class of information is physical. This is information that is inherent in something; the length of a car, or volume of water in a lake for example.

Logic-information is available to all to utilize. Concerning this information, we find agreement with Hess and Ostrom (2003: 144) who conclude that information is a force that cannot be stopped from distribution, not through market forces, nor governmental interference. Physical information is different to the degree that an individual can control its dispersal. It is not necessarily available to all.

We see from these two classes that one is distinctly private, and the other public, *in an abstract sense*. However, logic-information despite is availability to all also has the property of opening itself to being controlled, or "privatized." The process of creating knowledge is the method we use to privatize information.[93]

Information-Knowledge and Action-Knowledge

Kirzner (2005: 77) distinguished between two types of knowledge an individual can hold: information and action. In a Kirznerian world, action-knowledge is that which propels us to act. In contrast, we can hold information in our minds that is knowledge, but it may not necessarily cause us to take any further course of action. Hence, action-knowledge can only be formed as the joining of two pieces of previously existing pieces of information-knowledge. These pieces need not originate from the same person, but the knowledge of them will necessarily always be of a single person.

The way in which information becomes knowledge is crucial to the discussion. When we assimilate a piece of knowledge, we do so through our own personal senses, and this input is influenced by our experiences. Hence, we can sympathize with Polanyi and Prosch (1975: 61) when they comment:

---

92 See Kirzner (2005: 76) where he states, "[i]nformation is an input that may be used in a process ... that results in the possession of knowledge." See also Mises (1985: 109) where he views inventions to be the product of not something material, but of the mental process reasoning and collaborating information together.

93 See Hoppe (2007: 14) whereby each human enjoys a privileged access to their personal, and necessarily private, knowledge.

> We should now be able to see that all our knowledge is inescapably indeterminate. First of all, as we have seen, the bearing that empirical knowledge has upon reality is unspecifiable. There is nothing in any concept that points *objectively* or automatically to an sort of reality.

All knowledge is specific to us, although the source of it is, in some cases, universal: the body of logic-information. Hayek's (1976b: 106) stress on the sensory order, that is, how information is transmitted from the external world into our minds brings new significance to us. The indeterminate emphasis we will place on any piece of external information is influenced by our previous experiences, and the effects these have had on our memories and minds.[94]

Information-knowledge through logic-information is fundamentally non-excludable at its core. Anyone can access it through the access of their own logic. It is, in this sense, merely the personal embodiment of knowable information. However, in this procedure of personally acquiring the information, we alter it so that it is of a different nature. This procedure personalizes the knowledge based on our own subjective interpretation of it. This process creates an intimately personal element in the knowledge, one that precludes it from being identical to any other's.

Where does all this leave action-knowledge? As we have seen, action-knowledge is the collation of two or more pieces of information-knowledge. The end goal of all action-knowledge is to enact action. Although all actions are undertaken to better our current state of affairs, some are undertaken to do this relative to others. In this case, we will see that another person having access to this piece of action-knowledge can preclude an individual from using it for that end if they utilize it first.

All action-knowledge stems from information-knowledge. This aspect creates a chain of events that excludes some from accessing it. Although the flow of some information-knowledge (particularly that based upon logic-information) cannot be stopped, due to the nature of its source, action-knowledge is very different. Each piece of information-knowledge is "owned" by someone, despite coming from a common source. Hence, we can see that each resultant piece of action-knowledge is necessarily dependent on the "owner" of a piece of information-knowledge "sharing" it.

---

94 Earlier in Hayek's (1936: 44) life he had written that there was no way to be certain that subjective data held by a person, or between people, could be identical unless it was sourced from the same objective facts. However, in light of his later work, we can see that, although it is possible that the same interpretation may result, it is still not predetermined merely by utilizing the same objective source. See, for example, Hayek (1976a: 11) where he comments that the only thing we share with our knowledge holding counterparts is a general and very abstract concept of the knowledge they have.

Polanyi (1967: 4) would bring to light that "*we can know more than we can tell*." This implies that there are some pieces of information-knowledge which are unexplainable, they can only be conveyed through our very actions. Hence, these pieces of knowledge are mixed with another piece of knowledge, and the very action of doing this creates the entity we call action-knowledge. Action-knowledge is hence, not only a mental idea, but a physical process as well. As Boehm (1994: 160) sums up, "[k]nowledge yielded by the market processes knowledge generated through the operation of the market order – that is, it cannot be generated in any other way."

Some Brief Notes on the Interesting Characteristics of Knowledge

As Shackle (1972: 156) so succinctly states, "[s]o far as men are concerned, *being* consists in continual and endless fresh *knowing*." Humans strive for knowledge as an essence of being, the lack of this simple fact would result in a lack of life. The result is a constant strive to make more efficient the avenues of knowledge delivery and use. Hayek (1973: 15) viewed one of the key points in increased efficiency of knowledge not in making knowledge more widely disseminated throughout society, but by joining together the scattered pieces of knowledge that could not be disseminated. As we previously saw, much action-knowledge embodies this exact process.

Knowledge, due to its personal nature, allows the holder to control 100% of the supply. We see that Hayek's (1960: 43) statement, "[k]nowledge, once achieved, becomes gratuitously available for the benefit of all," is not generally applicable, but only if the knowledge holder chooses for this to be the case. However, much knowledge is of a tacit, inarticulable form. This implies that its transmission can only occur from personal contact between individuals (Polanyi 1958a: 52). Alternatively, as Kirzner (1973: 162) points out, some information cannot be garnered until it is obtained. For instance, to know how well a washing machine works you must first purchase it.

Machlup (1972: 363) comments on the divide between knowledge used for consumption, and that used for production. When we join this idea with Kirzner's dichotomy of knowledge, we can see that action-knowledge is consumption oriented, while information-knowledge only serves a role as a producer of action-knowledge. With this in mind we see that, concerning the field of knowledge, information-knowledge always represents a means to an end. The end is always the attainment of action-knowledge that results in our individual actions.

There is an important link to be made between the processing of information into knowledge, and the entrepreneurial process. Huerta de Soto (2005: 52) outlines six features of knowledge that are relevant for entrepreneurship: (1) it is subjective, rather than objective or scientific, (2) it is exclusive, (3) is is dispersed throughout actors' minds, (4) it is mainly tacit, and not expressible through words, (5) it is created *ex nihilo*, by the entrepreneur, and (6) it must be transmitted. These six factors make the entrepreneur supremely suited to disseminate information throughout the economy. Rizzo (1979: 9) makes an interesting point on the topic stating:

> Those who sell or buy information-producing resources are not fully aware of the value of the knowledge that will be produced with these resources. If they were to be fully aware, it would be necessary for the knowledge to have been produced beforehand and then, of course, there would be no demand for their value... The more innovative (and hence, important) the knowledge that will be produced with the resources, the more doubtful it will be that the sellers will know all of the possible outcomes... When, as a consequence, sellers underestimate the value of these resources, opportunities for pure profit will emerge.

If a holder of a piece of information-knowledge knew of its potential as action-knowledge, *and the holder also valued this course of action appropriately*, it would be used as action-knowledge – an action would result. An entrepreneur can make pure profits in this sense through exercising their ability to join disconnected pieces of information together.[95] The linking of disconnected pieces of information-knowledge is the entrepreneur's primary role in the economy.[96]

Lastly, the supply of knowledge can only increase, never decrease.[97] This counters typical goods whereby the trade of one unit is akin to a zero-sum game regarding supply. If I exchange one car with another person, my supply decreases by one, and the other's increases by the same increment. Not so with knowledge. When I trade or transfer my knowledge to another, my supply remains the same, and the other's supply increases accordingly. Likewise, the exchange of knowledge is not only a one-way street; both parties benefit from increased knowledge *even if none is explicitly exchanged*. If I trade a

---

95 See Huerta de Soto (2005) for an in depth look at the entrepreneurs function being the collaboration of knowledge.

96 As Pasour (1989: 100) elaborates, "It makes no sense to talk about consumers' demand for information on an investment alternative about which they know nothing. Much information about financial investments, for example, is concerned with making the consumer either aware of unknown investment opportunitites [*sic.*]or of unperceived aspects of already known investment alternatives." Hence, we see the importance of the entrepreneur in the information dissemination process.

97 This ignores the effects of amnesia, or death that necessarily do reduce the supply of knowledge. For living beings however, as Huerta de Soto (2004: 27) points out, the entrepreneurial knowledge creating process will continue forever.

piece of my knowledge in exchange for $10, although not expressly receiving knowledge, I now have new knowledge concerning demand for my knowledge. Hence, as Hayek (1936: 38) pointed out, there is some degree of circular reasoning with knowledge creation, whereby the input of one buyer's knowledge may lead to the creation of the seller's knowledge that is offered for exchange.

We thus see Hayek's (1945: 519) important emphasis he personally placed on knowledge, "[t]he economic problem of society is thus not merely a problem of how to allocate 'given' resources... it is a problem of the utilization of knowledge which is not given to anyone in its totality."

The Creation of Action

Although action-knowledge is the impetus for all action, our three elements all play an instrumental role in the culmination of this point. We see that a temporal order exists, whereby each must precede another for the production of action. The root must always be information, whether based on some physical entity or on logic. This will be personalized, or internalized, as a piece of information-knowledge. When the two pieces of this knowledge join, and the conditions are ripe, a piece of action-knowledge is created. The creation of action-knowledge *always* must lead to action; *there is no other outcome of it*.[98]

The corollary to this process is that the creation of action-knowledge, and its impending action, in turn creates two things: information-knowledge for those aware of the action, and information for those unaware of it. An action has occurred that has enlarged the sphere of physical information realm, creating further possibilities for the discovery of information. We can thus see the distinct roles each entity serves:

Information

As the source of all physical knowledge, information serves as the starting point for all action. It is given at any given moment, however, when we look at the temporal world we live in, the total amount of information is in a constant state of flux. This is because the very act of action, the end result of this chain of events we are now looking at, results in the creation of more information.

[98] The condition to create action will always present itself when two pieces of information knowledge form to create a piece of action knowledge. In fact, this will always result in action occurring. However, we see that if the individual does not value the course of action high enough than they will elect not to act, instead preferring to exercise their other options available. The action of not acting is, itself, an action.

Information-knowledge

Before any action can take place, information must be personally gained, and hence, knowledge created. We must possess knowledge prior to using it. Information-knowledge is the result of the action that possesses information for ourselves. *It is the step we take to "privatize" information, so that we may use it for our own ends.*

Action-knowledge

Action-knowledge is the final thought we have prior to action. It is the combination of of two pieces of information-knowledge. The resultant action creates a form of information that adds to the existing stock, and hence, expands upon the future possibilities for knowledge creation. action-knowledge exists only for a fleeting moment between when we think it, and when we act. Action will only occur at the specific point when action-knowledge is formed.

The Flow of Knowledge

Richard Cantillon wrote about the effect which bears his name in the early 1700s. In particular, he pointed out that as money was introduced into an economy, its effects were never neutral. An influx must always start at a given point, and its effects radiate outward therefrom. This process may proceed so slowly as to be almost unnoticeable at times, but must still occur.

Hayek (1945: 536) would use his famous "tin example" to demonstrate this effect. A demand for tin has arisen on the market, but only one producer may know it. Their increased demand for tin increases the relative price thereof and transmits this information to the rest of the market actors. However, we should not ignore the possibility that this signal could be incorrect and hence, transmit false information throughout the market. After all, as Hayek (1936: 33) noted, before we ask why anyone should be correct, we should ask why they would be incorrect. There is a distinct possibility that the information accounted for in prices will be falsely given based upon the actions of one individual misinterpreting the true importance of the facts.[99]

Williamson (1975: 5) would emphasize that "[t]he 'marvel' of the economic system is that prices

[99] See also High (1982: 165) where he notes it is not relevant for progress whether an individual acts upon data, but only if one acts upon the *relevant* data with the *correct judgment* for a beneficial result. As Lachmann (1978: 71) reminds us, there is a natural process built in the market economy to ensure the information is utilized in the most dynamically beneficial manner as entrepreneurs who cannot read the signs of the times correctly are substituted for those who can.

serve as sufficient statistics, thereby economizing on bounded rationality." Hence, we see Williamson takes the Hayekian view that prices are sufficient tools for conveying the information they contain. However, the issue with this viewpoint is that it tells us nothing of how these prices were arrived at (Thomsen 1992: 48). As a result, prices may tell us how prior actors valued the information; their demonstrated preferences result in prices. However, the prices themselves can tell us nothing of the relative value between the previous actor and the present regarding the same information. Hence, although prices tell us much of what others believe the value of past information to be, it is confined (or bounded) by he ancillary information they may have. This may or may not agree with how the present actor values the information based upon their own knowledge.

Likewise, the effects of information do not just happen in an instant, but gradually spread throughout the realms of knowledge and information. As an action occurs, a new piece of information is simultaneously created. This piece of information is available for others to acquire, and convert into part of their own arsenal of knowledge. This dispersed character of knowledge is of fundamental importance to understanding the market process – it is the essence of knowledge's existence.[100] As Salerno (1994: 114) tells us, "dispersed knowledge is not a bane but a boon to the human race; without it, there would be no scope for the intellectual division of labor, and social cooperation under division of labor would, consequently, prove impossible." It is interesting to revisit Hayek (1946: 95) where he emphasized "the paralyzing effect really perfect knowledge and foresight would have on all action." By his own reckoning, we see that imperfect knowledge is essential to action itself – action's existence could not be possible without this lack.

We see that, for the mainstream, imperfect, asymmetric, or dispersed information, whichever term is used, acts as a constraint on the equilibrium issue. The information is assumed to exist, and must be found to reach the equilibrium state (Kirzner 1997a). For Austrians, this same element of imperfect information, or what Thomsen (1992: 61) coins "previously unthought-of-knowledge", represents an element of sheer ignorance, and an opportunity for an entrepreneur to exploit for profit.

When we look at action, we see that actions are never absolutely large, but are comprised of many relatively infinitesimally small sub-actions. Hence, the resultant information created from these actions is likewise tiny in both size, and scope.

As the influx of new information is always a flow of small entities, the results from any new

[100] Lavoie (1985) furthered the view of "social intelligence" that no one individual has, but society as a whole has as a result of the interactions of people. This concept of a social knowledge is non-realistic, all knowledge, *by definition*, must be of a single person. We may however speak of *social information*. The distinction is of great importance for the acting human; one will allow us to act, and one requires a transmission to the individual before allowing this possibility.

information are small and manageable within the scope of an individual's sphere of existing knowledge. In this manner, changes and alterations that occur under the natural flow of information resulting from the process of action are necessarily small and contained.[101]

A brief note on two coordinations: knowledge and action

EMH assumes a coordination of knowledge brings us to a coordination of action. As information relevant to a security's price is released, and disseminated throughout the market, our actions are modified accordingly to bring a cohesiveness to the system. However, the actual relationship may not be so clear; the two types of coordination may both serve to influence each others fruition.

Kirzner (1973: 216) gave the following example:

> [B]y A's not buying B's apples, and by B's not selling them to A, each party is, because of ignorance of the other's 'existence,' acting as if the other did not in fact exist. A knows his own tastes and assets; B knows his. But because the bits of knowledge are not coordinated, the actions taken by A and B are uncoordinated.

In this case we have the common view of informational disco-ordination causing a similar disco-ordination in our actions. This gap is bridged by the entrepreneur bringing these disparate pieces of information into a single mind.

However, coordinated action need not be due solely due to coordinated knowledge. As Thomsen (1992: 90) shows us, action coordination can exist in hierarchical settings without a significant coordination of individual pieces of knowledge:

> [T]he commander [of an army] is ordering certain strategic manoeuvres, and there is among his soldiers one individual who, were he consulted, could prove to be a highly capable strategist. But the commander is unaware of his existence and therefore does not make use of his talents. There will not then be co-ordination of knowledge. However, in the sense described initially, there could still be (a visible) co-ordination of actions.

[101] With this flow of information in mind, we are reminded of Hayek's (1946: 106) comments on competition, "[c]ompetition is essentially a process of the formation of opinion: *by spreading information*, it creates that unity and coherence of the economic system which we presuppose when we think of it as one market" [emphasis added].

Informational disco-ordination is one of sheer ignorance; there is knowledge that may be known to exist that we are unaware of. The disco-ordination of action is of a Simonian bounded rationality type. The solution is there, but the complexity of the circumstances preclude the individual from reaching the coordinating solution. We see then that an informationally coordinated event may necessarily trend towards a coordination of action. However, action *can* be coordinated in the absence of information (although within the bounded confines of what is known).

The Relevance of Knowledge

More information exists than is necessary for the action creation process. Furthermore, it is impossible for any one individual to know all the information that exists at a given time.[102] It follows that actors must necessarily economize on knowledge, limiting their search to that which they think bears the most relevance towards the creation of action. As Hayek (1973: 83) would note, "[w]e never act, and could never act, in full consideration of all the facts of a particular situation, but always by singling out as relevant only some aspects of it."

All knowledge that is gathered is done so to create action. There is no other reason for the attainment of knowledge. Stigler (1976) and Rizzo (1995) have shown that information is only sought until its marginal benefit is limited by its marginal cost. However, as Machlup (1972) demonstrated, it is very difficult to measure the marginal benefit of information, hence the difficulty in realizing the point at which marginal costs and benefits equate. This problem arises as we are collecting information into our personal knowledge, for use as action. Action is always at a future point, specifically unknown to the actor at the point in time they initially gather the information. This time differential creates the uncertainty that makes the benefits difficult, if not purely impossible (although we cannot theoretically state this with certainty), to calculate.

However, costs of information are always known, in a general sense, as this is the part of the process that exists in the present. The problem here is that some information may prove too costly to

[102] Although we can see that in a static setting, information is given and may be known in its entirety. However, we see that the process of learning information, or making it part of our personal knowledge, is inseparable from the element of time. We have also seen that the passage of time creates additional information through our actions. In a dynamic sense, the only way that we can view the world, it becomes clear that it is impossible to know all the information that exists, *even though it is statically possible*. See, for example, Hirshleifer (2001: 1539), "[t]he other possible reason for the persisted mispricing it that some relevant pieces of public information are ignored or misused by everyone. This can occur either because the signals are obscurely located or because our shared model of the world is just not sophisticated enough to make their relevance clear."

acquire. Indeed, as Stigler (1967: 291) states, "[i]nformation costs are the costs of transportation from ignorance to omniscience and seldom can a trader afford to take the entire trip." The benefits of knowledge are seen as embodied in their utility in servicing their respective end (action), which is in itself a means to another end (the fruit of that action). As action continues as long as the human existence continues, we can see that the search for knowledge through information will continue unabated forever.

However, due to the uncertainty between the point at which information is gathered, and action results, knowledge may lose its importance for us. This loss of relevance may preclude its direct serviceableness in creating the action we desire. This loss of serviceableness does not imply the negation of knowledge, only in its role in providing the impetus for action. We may call this an *ex post* error, the attainment of knowledge to create an action at a future time *which never occurs*. *Ex ante* we can never know with certainty what the future relevance of our knowledge will be. However, just because a piece of knowledge has lost relevance to an individual for their expected action, does not mean that the inherent information in it is not valuable for another individual. The multitude of ends actors pursue implies that information will not have an equal value to all, it will be subjectively derived from its ability to provide serviceableness towards an expected end.[103]

Although it is always the actor's individual decision to accept a piece of information into their knowledge realm, much more knowledge is generally learnt than is necessary for the desired end. A simple abstract illustration may help. A fictional book may be colorfully written with many adjectives to create a descriptive atmosphere. If the end sought is the enjoyment from reading a vivid story, the added information through these words may bear relevant value to the actor. However, if we look at an earnings report from a company, the same use of colorful adjectives may provide much more information than is necessary. The difference can be seen in the following two sentences:

[1] The company lost $10 last in the year 2000.

[2] The Seattle-based company of 13 employees unfortunately lost 10 fresh, green US dollars, continually facing a storm of devaluation in the tumultuous currency markets, in the year 2000, the Chinese year of the snake.

[103] See Kirzner (1996: 150) for the importance in the market economy of the ability to "*know the importance to others* of the goods and services one commits to that action, and the importance to others of the goods one will obtain from that action." Hence, *knowing what we need to know* is an important part of the market economy, but this end can never be completely achieved in a dynamic setting of continually shifting preferences, knowledge, wants, means or ends.

This is an extreme example of the excess knowledge provided through non-relevant information, but the standard remains. Relevance is only applicable towards a given end of an individual, and hence, there is no way we can predetermine if a piece of information is relevant or not in general, only if it is so to us. We see the further problem arising where to get the relevant information required from passage 2 above, we need to read the whole sentence, which entails gaining the knowledge of much *possibly* irrelevant information.

Although a lack of immediate relevance may preclude immediate action from occurring, the storage of this knowledge in our memories means that in the future, as situations change, the relevance may also alter. Hence, knowledge that seems to have little value for a given action at t_0, may become essential and highly valued at some future t_n. We see relevance is time-variant, but in a meaningful way – always regarding expected future action.

Hayek (1945) viewed prices as being important transmitters of information throughout the market.[104] That they perform this role at low cost is an added benefit, and the speed and ease at which an individual can grasp this knowledge is apparent. Weak-form efficiency under EMH implies that current prices reflect all information contained in past prices. Strong-form represents the more extreme definition where not only the information contained in the weak-form efficient price, but also all hidden information that may not be publicly available yet. It should be stressed that these prices may not incorporate all the information necessary in a flawless manner.[105] As Mises (1936: 115) pointed out, market prices can never be a perfect tool in this regard, however, they can be of extraordinary value to this end. The entrepreneurial process is centered on finding the inter-temporal spread existing *in the future*, based upon the prices of today. It follows that, although these prices may not be based on wholly correctly interpreted information, it must still be utilized by the entrepreneur. In fact, the entrepreneur's existence is owed to the fact that their predecessors will have *misjudged* some important piece of information, and thus created the inter-temporal folly needing correction.

If Hayek viewed prices as transmitters of knowledge, others, such as Soros (1995), hold that prices and information exist in a "feedback loop." Under this view, market value does not simply discount the data available, it actually creates the data (Soros 1994a). This "reflexivity theory" hinges

[104] Baetjer (2000: 148) notes that "[o]ur knowledge is to be found in practice not in our heads, but in the capital goods we employ. *Capital is embodied knowledge*" [emphasis added]. We can make an important distinction that these will only represent past *action-knowledge* revealed through the resultant action. They cannot represent the latent *information-knowledge* embedded in our minds, waiting to be acted upon. He has since recognized this temporal aspect of the knowledge embodiment through capital (see Baetjer and Lewin 2007: 8).

[105] Yeager (1994) stresses the difference between possessing this information, and fully assimilating it correctly. The successful use will depend on both of these conditions being fulfilled.

on market actors creating the market data that they are acting upon.[106] We see that the attainment of information is only possible through action. Action is a prerequisite for new knowledge to be attained, however, this very act shapes the nature of the future information available. Action and knowledge are an inseparable duality, one we have already looked at. That knowledge must always occur temporally after action is unassailable, although we can now see how action influences *the future information* available to be attained.

As it is also person dependent, we see relevance is independent of its holder.[107] An individual that holds a piece of knowledge may have no use for it, but this same knowledge may be of great use for another. As we will see in the next section, this differing relevance will have significant consequences regarding concepts of efficiency, and will highlight the role the entrepreneur serves in the economy.

The Meaning of Knowledge

Knowledge, then, is the personal attainment of information. It is the method that we personalize and subjectively interpret the information that previously bombarded the mind as stimuli. Kirzner's (2005) dichotomy between information and action knowledge allows us to see that there is much knowledge that remains dormant inside us, awaiting to be acted upon. Action knowledge, on the other hand, is that which directly causes the individual to act – it is the result and goal of all knowledge.

Knowledge may also exist in a very inarticulable form as *tacit knowledge*. In fact, the degree to which it is unable to be communicated may be so extreme that it is *impossible* for a holder of it to relay this knowledge to another individual. The pricing system may serve as a proximal transmitter of this knowledge, however, as has been shown, what is actually been transmitted through a price is not the knowledge itself but its residual. Hence, we may see a price fluctuation caused by a individual purchasing shares of a company, but this does not necessarily inform us as to the true motives, or knowledge, behind the share purchase.

Knowledge will forever be limited or partial for the individual. This is not to be viewed as a bane but a boon, as this limited knowledge is the reason we continue to act. However, there are implications that result from this limitation. In fact, previously, we looked at the absolute concept of

[106] See Hoppe (2007: 16) for a look at this feedback loop in knowledge creation and attainment.
[107] This independence comes to light when we view Chamberlin's (1957: 146) comment on the information youths seek when buying cigarettes, "they are perhaps more interested in knowing that a famous movie star smokes a certain brand of cigarette, than they are at knowing what that cigarette was made of."

logic. As logic can only exist in a complete, unalterable form, we could possibly be enticed to view the resultant actions of this logic to be similar, or at least coherent, with this objective concept. Instead, a cursory observation of the world around us reveals that nothing could be further from the truth. The reason is that logic presupposes that the necessary perfect knowledge is available to be at its disposal. The preceding discussion should make the reader aware that this is not a possibility. Instead, it is within the *bounds* of our knowledge that we must use our logic; a confine which considerably alters the results of this immutable process.

Conclusion

Information, knowledge and action form an intricate trinity, with each being necessary for another's creation. The process of creating action necessarily entails the attainment of knowledge from the existing information base. In turn, the creation of action is the process that expands this information base. We see then that as action is inseparable from human life, as long as we exist, information will continually expand.

A curious type of knowledge exists, tacit knowledge, that is inarticulable to others. This implies that it can only be spread through more intimate means than other forms of knowledge. As action is dependent on all types of knowledge, we see that the inability to gain this tacit knowledge may preclude action from occurring. Also, the personal nature of knowledge relevance may mean that although knowledge exists, it may not be acted on. If this knowledge is of a tacit type, than a significant period of time may pass before it is passed from a non-user to a potential user.

The implications are two-fold. The first is that knowledge exists that may not have been acted upon. As action is an ultimate end, and knowledge as a means is valued by its utility to service that end, each individual will place a different value on the knowledge they have. Although *ex ante* they will always value it towards a given end, *ex post* this may not be the case. As ends are individually created, relevance is also an individual measure. This gives rise to the second conclusion; knowledge that exists at time t_0 may not be acted on until an unspecifiable t_n. The mere existence of knowledge is not enough to entice action, only that knowledge which is held in an actor's mind who has the end goal of action available through that knowledge. Hence, an individual may have much information-knowledge, but this may do little to create the necessary piece of action-knowledge necessary for the creation of action. As knowledge necessarily requires time to be passed from one individual to another, we can see that a significant time-lag may occur between the initial creation of knowledge, and the final acting upon it.

Finally, an important corollary exists regarding both of these points. Some EMH opponents have offered the argument that if it were to hold strongly, there would be no incentive for any one person to create any knowledge. Malkiel (1973: 348) offered the example of a dollar bill lying on the street. If one believed EMH to hold true always, the dollar bill could not be real, as someone would have been alert to this knowledge, and picked it up accordingly.[108] Hence, the value placed on new knowledge must be zero, if it is assumed that it is already known. The fact that we would pick up a dollar bill if seen on the street is taken as evidence that EMH cannot hold in the real world.[109]

Action is the end that knowledge is the mean for. As action never ends, we see that there is a continual value placed on new information. This will preclude the eventuality that the value placed on knowledge will be eliminated, and hence, result in the elimination of action. Instead, we will pick the dollar bill off the ground because, (1) we are the first person to stumble onto this knowledge, or (2) we are the first person to stumble onto this knowledge *and* have the relevant end in mind that values its attainment more than the cost involved. Knowledge is thus continually sought, and acted upon, due to the individual natures of our pre-existing knowledge, as well as our subjective valuations of our individual ends. As Kirzner (1992: 117) points out, in a disequilibrium state the informational role of prices is quite distinct than under equilibrium circumstances. Hence, in disequilibrium, prices spur on entrepreneurial discovery, and therefore, create new information.[110] As the world we act in is necessarily one of disequilibrium, we see the importance of this viewpoint.

---

[108] That the dollar bill is picked up in reality is evident. One of Kirzner's favorite metaphors involved this $10 bill lying on the ground. “Many people do not see the bill; but the entrepreneur is more alert than his fellows, and so he is the first to see, and to snatch the bill. Superior alertness, alertness to truth out there, accounts for entrepreneurial profits.” As quoted in Rothbard (1997: 128).

[109] David Friedman (1996: 9) provides another analogy in his comparison with a supermarket checkout line. If we were in equilibrium, then every line would be equally fast, the lines would all be the same length, and customers would not have to make a choice in assessing the best line. However, if that were the case, no one would be choosing the a line purposefully to use the quickest/shortest line. In that case, the driving force behind equalizing the lines would be removed, and hence, the static situation would be unbalanced (or dynamic).

[110] For both Mises (1998; 1980), and Hayek (2002), this disequilibrium state would represent a circumstance of *false* prices permeating throughout the economy. It is Hayek's competitive discovery process that seeks the necessary knowledge to replace these false prices with somewhat less false prices. As we can never reach a state of full equilibrium, we see that prices will forever have to be considered as technically *false*.

2. Efficiency

The efficient market hypothesis rests on the tenet of informational efficiency. In fact it divides this concept into three parts: weak-form, semi-strong form, and strong form. Fama (1970) would define these three as follows:

[1] Weak-form efficiency – future prices are already fully reflect all information that can be derived from examining past market data, such as history of past prices, volume, or other metrics. These past data are publicly available and assumed costless to attain. The implication is that prices must adjust instantly to all new information, hence, the opportunity to exploit new information is limited.
[2] Semi-strong form efficiency – current security prices contain all publicly available information and reflect this accordingly. This includes information regarding the business the firm is involved in: sales, earnings forecasts, management, etc. As this information is also assumed public and costless, benefits derived from this form are correspondingly limited.
[3] Strong-form efficiency – current security prices reflect all relevant information to a firm, including that known only to company insiders. This is the most extreme type of efficiency for EMH. Interestingly, in 1934 the SEC enacted rule 14(e) of the Securities Exchange Act to combat against insiders from exploiting this information.[111]

Thus, the efficiency that is sought for EMH practitioners is to fully reflect all information. This idea precludes the possibilities of earning above average returns, as there is no way to utilize knowledge that has not been previously disseminated by other professionals.

Static Efficiency[112]

[111] See section 14(e) of the Securities Exchange Act of 1934:

> It shall be unlawful for any person to make any untrue statement of a material fact or omit to state any material fact necessary in order to make the statements made, in the light of the circumstances under which they are made, not misleading, or to engage in any fraudulent, deceptive, or manipulative acts or practices, in connection with any tender offer or request or invitation for tenders, or any solicitation of security holders in opposition to or in favor of any such offer, request, or invitation. The Commission shall, for the purposes of this subsection, by rules and regulations define, and prescribe means reasonably designed to prevent, such acts and practices as are fraudulent, deceptive, or manipulative.

[112] The EMH is similar to other situations incorporating static assumptions, and their implications on efficiency. Therefore, only a few brief comments will be made on this aspect. For a thorough critique from this viewpoint, see Pasour (1989: 95).

The presumption that all EMH efficiencies must be based, at least implicitly, upon is the concept of static efficiency. This paradigm was originally employed as a method to eliminate waste and use resources to their full extent. It can come as no major surprise then that this idea has a core similar to that used in physics. The roots of this influence can, from an economics perspective, be traced to Walras, who explicitly claimed in his 1909 paper, *Economique et Mécanique*, that he was using formulae identical to those then used in physics (Huerta de Soto 2004).

The use of these concepts, borrowed from the realm of the physical sciences, removed any inter-temporal or dynamic aspect from the concept of efficiency. This would open the door for Pareto to create the efficiency that now bears his name. A Pareto efficiency situation is one where no person can be made better off, without making someone worse off. The implication is that inter-personal utility comparisons can be drawn, from which to determine if a person is really made better or worse off by an action.

One of the grave consequences of this viewpoint is that utility is seen to be measurable. Instead of the ordinal ranking of wants, a cardinal ranking is imposed. When we view utility in marginal, ordinal terms, it is clear that these quantities are not measurable. In fact, they are not even quantities. Instead, a utility is an immeasurable, time-variant entity.

Kaldor and Hicks would relax Pareto's criteria slightly and achieve a new form of efficiency. For Kaldor (1939b), if a person who benefits from an action can compensate a person who loses from an action, the resultant situation will be more efficient. Hicks (1939) would add that a situation is efficient if those who are made worse by an action cannot prevent the action by “bribing” or otherwise compensating those who gain from it. Both gauge the respective total utility in monetary terms. The use of a monetary meter-stick had the advantage of eliminating the problem previous economists faced of being unable to directly measure happiness or satisfaction.

However, we can see the fallacy in this logic of static informational efficiency by turning to Mises (1949: 327) where he notes that all prices are past prices. Thus prices can only measure past conditions, operative at one given moment in time. As Pasour (1989: 99) notes:

> A problem arises in identifying inefficient markets under real-world conditions because of uncertainty, imperfect knowledge, and costly information. A meaningful efficiency test has not been devised even under static neoclassical conditions where a defensible criterion of

efficiency must be based on an 'appropriate amount' of information.[113]

Stringham (2001: 48) concludes that other measures are equally unsuitable, for instance, "Kaldor-Hicks efficiency is an unusable standard" as it requires inter-personal judgments which cannot be made.

Rubinstein (1975: 812) notes that there are three different concepts of efficiency that economists can use. First is exchange efficiency, a state he defines as existing when no participants wish to exchange with one another. Second is production efficiency, the more traditional viewpoint of value-maximizing individuals operating in Pareto optimal conditions. Last is his idea of informational efficiency. This is the efficiency defined by EMH, whereby present security prices are costlessly known to all, and all past information is reflected in this prices. All three of these definitions provides a static look at the market. There is no room for new inputs, they operate within a given Robbinsian means-ends framework.

Hence, as Huerta de Soto (2004: 23) points out, there are three broad fronts that these concepts of static efficiency are all susceptible to be disputed on. First is the assumption of a comparable standard of utility. We have seen that this standard fails to exist in actuality. Second, there is the assumption that utility rankings are known and unchanging. However, we know that in the dynamic flux of time, preferences and the inherent utility of them are also in a constant state of change; they are time-variant. Third, the prime assumption is that efficiency in economics (static efficiency) is the same as technological efficiency. This viewpoint thinks that the key is to reduce inputs, and can generally be done by treating the issue as an optimization problem. This viewpoint persists, despite the attacks forwarded to counter these arguments.[114]

The Maximizing Individual

A further deficiency arises when viewing static efficiency in the conception of humans. The typical viewpoint of humans has been one where we are identified as homo economicus. As John Stuart Mill would sum up:

> [Political economy] does not treat the whole of man's nature as modified by the social state,

[113] See Demsetz (1969) for this viewpoint of efficiency measured as an "appropriate amount" of information.

[114] For example, see Robbins (1952: 36) or Alchian and Allen (1964: 435) for refutations of this viewpoint. For a more recent example, see Sonsoles Huerta de Soto (2005).

> nor of the whole conduct of man in society. It is concerned with him solely as a being who desires to possess wealth, and who is capable of judging the comparative efficacy of means for obtaining that end. (as quoted in Persky 1995)

However, this view of humans, operating only within a given means-ends framework and trying to maximize their profit is a flawed way of looking at action. The Robbinsian maximizer, or homo economicus, is more aptly described as Mises' (1949: 253) *homo agens*. This actor is able to see a new means-ends framework, one that failed to exist prior. Given this, actors are only Robbinsian maximizers to the extent that this framework remains static and unchanging (Kirzner 1973: 33).

When we view actors in this light, we see them not as passive observers to the world, but as active creators of it. In fact, the existence of these actors creates a feedback cycle of new entrepreneurship (Holcombe 1998: 54). The result is a constant creation of new information through their actions.

Dynamic Efficiency

A new conception of efficiency has been forwarded by Huerta de Soto (2005). This dynamic perspective focuses on the concept of entrepreneurship and the changing world. By viewing the world as a process inextricably linked to the passage of time, we can see a very different perspective emerges. Humans are now viewed as *homo agens*, creating new possibilities in this changing environment. Kirzner (1997: 67) would view dynamic efficiency as the "ability to encourage entrepreneurial alertness to valuable knowledge the very existence of which has not previously bee suspected." The implication is that information is constantly created by entrepreneurial functioning humans.

In Huerta de Soto's (2004: 27) eyes, this occurs in six ways.

First is the generation of new ideas. This occurs as every entrepreneurial act entails the discovery of new knowledge which the actor did not previously possess. This knowledge is of a personal nature to the entrepreneur, subjectively interpreted and internalized from the external world. The implication of this is that much knowledge is tacit, as it is held within actors, and diffuse, as it is spread throughout the world in the minds of many individuals.

Second is the fundamentally creative nature of entrepreneurship. As actors search out new profit opportunities, they must create new knowledge in their minds that failed to exist prior. The creation of this knowledge is the ultimate source of all pure entrepreneurial profits.

The third and fourth points involved the transmission of information, and the coordination that results. True entrepreneurs transmit information between individuals who previously had no knowledge of its existence. This in turn acts to coordinate these individuals as to where they are valued most. The signal is thus created of an opportunity that may be taken advantage of.

Fifth is the competitive aspect of entrepreneurship. The rivalrous aspect of the entrepreneurial function continually has individuals trying to better others and acquire, and utilize, available information before they can. This competitive process breeds the condition for the final point, that the entrepreneurial process never stops. *No final equilibrium state is reached*. This is due to the fact that the mere act of acting creates fresh information which the entrepreneur must adjust to. In this way, knowledge, and hence resources, are never given. They continually expand and require constant management through the entrepreneurial function.

If the goal of static efficiency is to acquire and use all resources with the least amount of waste in the present, the goal of dynamic efficiency is to continually create new resources through the entrepreneurial process to create new possibilities for action. These two concepts are not mutually exclusive. As Hayek (1946: 100) insightfully put it:

> The real economic problem in all this is not whether we will get given commodities or services at given marginal costs but mainly by what commodities and services the needs of the people can be most cheaply satisfied. The solution is always a voyage of exploration into the unknown, an attempt to discover new ways of doing things better than they have been done before.

Dynamic efficiency also continually pursues the goal of static efficiency. The coordinating factor inherent in the process will erase the existing maladjustments, thus eliminating, or reducing, static waste. However, given the endless flow of new information, we see that pure Pareto efficiency can never be achieved. The entrepreneurial process cannot, however, ever be perfect in the sense that it eliminates all inefficiencies from the economy. This can be equated to the concept of X-Efficiency (Leibenstein 1966). A degree of inefficiency exists supposedly due to the incompleteness of information available. This raises attention to an important source of inefficiency that fails to exist with an equilibrium viewpoint. Stigler (1976) would try to show that there will always be an amount of inefficiency in the market, however it will always be an optimal amount. This is due to the fact that the search for new information will cease when its marginal cost nears its marginal benefit. Rizzo (1995: 12) notes that "[i]nformation difficulties... are not perfectly offset by the greater potential for profit." As Grossman and Stiglitz (1980: 393) would also point out, "there is an equilibrium degree of disequilibrium."[115]

We can see, however, that this will only exist in the static sense. Hülsmann (1997: 48) shows that we are unable to calculate the marginal value of knowledge. The reason is that in a dynamic view, inefficiencies are not yet known to have existed yet. As Sautet (2000: 64) shows us, preferences, resources, technologies, and the like, are all constantly changing. These new variables imply a state of rest can never be achieved, and that these possible inefficiencies can only exist in an unknowable future. As entrepreneurs seek to erase inefficiencies once they are discovered, future inefficiencies are possible, in fact they are probable, but we will not know in the present what they will be, or when they will occur. They represent an unknown unknown. The point of calling an unknown inefficiency an inefficiency is lost once viewed from this perspective.

The seeming paradox of Grossman and Stiglitz (1980) also assumes that present prices are always a valid representation of the underlying information, a condition that we have have seen to be not necessarily true. The actual result, in terms Grossman and Stiglitz might understand, is that a "noisy equilibrium" results, where prices cannot convey all the information assumed. The source of this confusion, according to Boettke (1997: 31), is that they assume information is available but dormant, like a book situated on a bookshelf, waiting to be pulled off and used. Obviously, this ignores many of the aspects that make information so unique, as have previously been looked at.

Individualism and Efficiency

Efficiency applies itself towards a given goal. Goals are always individual, they can never apply for all individuals equally, only to their creators. It follows then that any idea of efficiency that we seek must be individual in nature. Rothbard (1997: 267) may have put it best, "efficiency only makes sense in regard to people's ends, and individuals' ends differ, clash, and conflict."

[115] Sautet (2000: 64) shows that there are three main sources of entrepreneurial error, which will all act to deny the arrival at any static equilibrium. First is the error through missed opportunity. Second is the spurious discovery error, where entrepreneurs act as if something was discovered which hasn't been (an addition to this is when the entrepreneur only errs in magnitude, for instance, is more optimistic than reality warrants). Last is when an individual steps in and exploits the opportunity before the entrepreneur is able to. Of course, the third source of error applies only to the individual acting entrepreneur, and not to the further disequilibrium of the economy.

It follows then that, much like the concept of utility that efficiency is based upon, efficiencies are not cardinal nor additive. We cannot apply a concept of general efficiency that applies to all equally. The trichotomy of concepts we find in EMH efficiencies – weak, semi-strong, and strong – imply that all individuals share and practice these same goals. We see however that ends are ultimately given, and as such, of a deeply personal and individual nature. One may wish to pursue a different end, and hence, a different type of efficiency than another. For example, while indeed one person's end goal may be towards the dissemination of information, and will seek efficiency in this action, a different person may well prefer to try keeping the maximum amount of information they have private. The implications are stark. One person will be trying to achieve the exact opposite of what another believes is the efficient action.

A prerequisite for general equilibrium is full certainty and perfect knowledge of the future; our expectations must equate perfectly with the unfolding of time. As Rizzo (1979) points out, disequilibrium situations are surrounded by divergent and inconsistent expectations. This world of disequilibrium is the world that humans exist in. This inherent inconsistency in expectations gives rise to a plethora of different end goals that actors aim at achieving. As was previously seen, the existence of different goals implies differing efficiencies sought. Contra the prior belief that entrepreneurial actions are solely equilibrating, many are just as likely to be dis-equilibrating.[116] Hence, the market is an endless flow of opposed forces, denying the creation of any type of equilibrium. As long as the world exists in a state of disequilibrium, efficiency will be a unique end.

Rothbard (1997: 270) would note that “[n]ot only is 'efficiency' a myth, then, but so too is any concept of social or additive cost, or even an objectively determinable cost for each individual.” In fact, earlier in the same writing, he would explicitly refer to efficiency as a “chimera.” This is true from a dynamic perspective. In a static world, ends would be given, and efficiency could be sought as these final ends remained constant. In a dynamic world, one marked with varying ends in a constant state of flux, we see the whole concept of efficiency called into question.

Implications for Information

With this new paradigm of dynamic efficiency, what becomes of the efficient market hypothesis? Traditionally, efficiency has been attached to a static system, one where the given means-ends

[116] However, we see this only applicable towards other's individual ends. An entrepreneurs action may dis-equilibrate the plans of another entrepreneur, but society as a whole will necessarily be moved to a state closer to equilibrium.

framework is known in its entirety at any one point in time. This assumed world is one of fancy however, wholly detached from the dynamics that occur on a regular basis.

The information creation process is constant and without end. It is impossible to fit an informationally efficient framework within this world. The static viewpoint, as is taken in the extreme form of EMH as strong-form efficiency, views information as being entirely known, and acted upon to the full extent possible the instant it is created. However, even were humans capable of utilizing each piece of information correctly and wholly at any given moment, we could still never consider ourselves to be operating in an efficient world. As Huerta de Soto (2005) has articulated, the continual information producing process by entrepreneurs implies that all the information can never be fully acted on, *in the dynamic sense*.[117]

Furthermore, efficiency only exists for a given end. Individuals each utilize their own ends, which may not be conducive to the ends sought by others. For instance, we have seen how the two cases of one person striving towards information dissemination, and the other striving towards information privatization, are mutually exclusive The success of one precludes the success of the other. To speak of a blanket static efficiency, applicable to both actors simultaneously, is nonsense.

Lastly, as was previously seen in the previous section, the essence of information inhibits it from being viewed in an efficiency-maximizing way. Information may be tacit, and hence inarticulable. The implication is that it may wait latent, although existing, for a period of time before agents can physically interact to share their knowledge.[118] Butos (2003: 303) sheds light on the point that some market participants may not wish to share their knowledge, taking on a "discretionary policy" instead. Action such as this will necessarily limit the amount of available information in the economy. Kirzner (1973) and Huerta de Soto (2005) view the entrepreneur's primary role as the discoverer, and collaborator of this scattered information. However, when we speak of efficiency, in particular efficiency regarding information, we assume that the end is known. It is clear that the information that is needed may not exist yet, and hence efficiency as the maximization of its reach becomes, in Rothbard's words, "a chimera."[119] We cannot view future events in a maximizing way, we can only

[117] As Lachmann (1977) points out, the market process moves through periods of both equilibrating and disequilibrium forces. These preclude any notion of equilibrium from occurring. The "Lachmannian problem" is that we cannot know if an individual's actions are creating equilibrium or disequilibrium (Sautet 2000: 68).

[118] As Baetjer and Lewin (2007: 25) correctly point out, exact knowledge can never be shared between individuals due to its subjective nature. We can say that the knowledge person A has is person B's information, and that this information is able to be shared. It will then become the new knowledge B has is distinct of the knowledge of A, due to the subjective process of acquiring and interpreting it.

[119] Indeed, as Buchanan and Di Pierro (1980: 700) point out, "[w]hen individual investment decisions are considered in the context of Knightian uncertainty, where the set of outcomes is not and cannot be determinate, the very notion of

move toward the future, allowing the process to determine its rate of growth independently.

Conclusion

EMH rests on three patterns of efficiency. All three necessarily assume a given means-ends framework, where all respective knowledge is known to all and utilized fully. In reality, we have seen that all knowledge is not even knowable to all, much less utilized by them. As Pasour (1982) notes, there is a large difference between market efficiency and the existence of profit opportunities. Due to the fact that market inefficiencies at any given time are not seen does not imply the non-existence of entrepreneurial activity in the pricing process. In a dynamic sense, as humans are not omniscient, plans can never be fully coordinated, or successful, leading to a continual stream of inefficiencies, prime for the attentive entrepreneur to exploit. This dimension marks a grave deficiency in the traditional framework, one which emphasizes its static underlying assumptions.

expected-value maximization is ambiguous and misleading."

3. The Determinants of Price

One of the most troubling conclusions of EMH, both practically, empirically, and theoretically, is that as all knowledge is disseminated fully at any given time, future movements in prices must be random. This stems from the belief that the generation of future knowledge, as it lies in the indeterminate future, is itself a random variable. Practically, we see the problem arising that the implication is that no one should make an effort to predict price movements, as they are deemed unknowable *ex ante*. Empirically, we can see that there are many individuals who continually forecast correctly over the short-term, and some who see success over longer periods of time also. Theoretically however this issue has received scarce attention. Most arguments, and ensuing debates, surrounding this point of EMH are empirically-based, thus lacking the necessary rigor to forever disprove this contention.

There are several issues which we will look at regarding this point. The first will surround the allocation of goods. The entrepreneurial role is commonly seen as planning for the future. In this regard, their ability to mitigate the uncertainty inherent in the passage of time is seen as being purposeful, and not random. The true nature of prices must be understood before discussing implications for future prices. As prices are indirect signals as to future expected conditions, we can see that these can be neither fully random, nor fully correctly forecast at any given time.

Second, we see that changes in prices are not solely caused by the arrival of new information, and hence, the inherent assumed randomness of this factor. Instead, two particular factors inherent in humans determine prices. The first to be looked at is our time-preference scale. As this determines the demand we have for present over future consumption, we can see the relative spread on prices that will result from this. Second is our value-scale. As prices are necessarily a monetary measure influenced by the relative rank on this scale, it can be drawn that any independent change in an agent's value scale will correspond to a relative change in the price of a good.

Finally, it will be instructive to revisit the marginal insight. The case when determining a price is never of a homogeneous group of people buying all the stock of a good, but of each individual given their personal information creating a value for a given amount of a good. As prices are set on the margin, we see that the EMH implications for value theory are unfounded. The assumption that all actors hold identical information and hence determine value in a cohesive manner is wanting. Instead, we see each individual unilaterally values good, and it is the marginal pair of buyer and seller who equate these valuations with total demand.

The EMH conclusion that the future is a random variable is found lacking in substance. Future prices may never be fully predictable, this would be a condition that precludes action, but they can be approximated given the expectations we have that are themselves based on purposeful behavior.

Uncertainty and the Future

The element of time is an unalterable part of human action; it constrains every act we do. The prime complicating factor of time is the uncertainty inherent in it. This uncertainty is not full however. The future is always the result of human actions in the present. It may be said by some that the long run is only a series of short runs joined together. For acting humans, this is erroneous; action can only exist in a present.

Action takes into account future expectations, in fact, as Mises (1949) noted, all action is future oriented. However, the fact that we take future considerations into account when formulating plans of action does not negate the fact that action must necessarily exist only in the present; there is no other way for it to exist.

So we can see that although the future represents an unknown factor (an unknown or known unknown), the action that mitigates this uncertainty exists only in the present (a known known). We find then that in viewing the alleged randomness of the future, we must look to the source of our expectations. As we will see in the following sections, these must necessarily be based upon non-random variables.

Price Determinants

Price is concerned with only two physical factors: supply and demand. *Prima facie*, we can see that these cannot be random variables. If future supply and demand variables were randomly given, there would exist and immense problem for entrepreneurs in calculating the needs of consumers; *chaos would ensue*. That the market economy works, for the most part, without significant chaotic trouble implies that these future factors must be predictable to some extent.

Shostak (1997: 30) notes that one of the gravest fallacies of the EMH is the total disconnect it has between the financial and real world. It is assumed that the financial realm exists solely independent of the real world. In reality, nothing could be further from the truth. The financial realm is a reflection of the real world. When we look at an individual security, we see that its value, in part, derives from the real world in which it operates. A company's real world valuation is reflected in its financial world valuation.

As Mises (1949: 520) would put it:

> [S]tock exchange transactions produce neither profits nor losses, but are only the consummation of profits and losses arising in trading and manufacturing. The profits and losses, the outgrowth of the buying public's approval or disapproval of the investments effected in the past, are made possible by the stock market... It is ultimately the consumers' attitude that makes some stocks rise, others drop.

Hence, the connect that security prices have with the real world is through profitability. This is, in turn, determined by the business' ability to meet the future demand and supply considerations imposed on it by the market. Although we can have a general indication of what these considerations may be in the future, it will not be until this unknown reaches the present that we are certain of their values.

However, security valuations do not move randomly to keep equated with the underlying business concerns. Instead, it is the entrepreneur working to find discrepancies that brings coherence to the market. We may return to Shostak (1997: 36) where he states:

> Profit occurs whenever an entrepreneur discovers that prices of certain factors are undervalued relative to the price of the final product. Once an entrepreneur acts upon this, he eliminates the potential for further profit. For an entrepreneur to make another profit, he would have to be engaged in a different activity. Also, *no entrepreneur can know what ideas he will have in the future*. [emphasis added]

The future is never determined, but the way that we progress into it is. When we view securities, we see that the prices do not move into their future state randomly, but through the purposeful speculation of entrepreneurs. We have seen that all action is concerned with the present. Although the future is of an uncertain dimension, this is of no critical concern to the entrepreneur. They can only act in the present. They do know what demand and supply considerations of the present are. From this they can determine if the price is a good approximation of real needs, or if an opportunity is presented to move the price towards what its underlying business concern suggests it should be.

Ex ante, we can see then that every action is both purposeful, and expected to yield a profit. *Ex post*, the case may differ. The action will always remain a purposeful, non-random event, but the profit may not materialize, or not to the extent that the entrepreneur initially expected. The entrepreneur is one who believes that the conditions exist to yield a profit. When the opportunity manifests itself, the entrepreneur's role is to act upon it, and hence, bring prices of the financial realm towards the limit that the real world requires.

Entrepreneurial profit is not random, but determined by three things. The first is the needs of consumers. These are not random, but will always be conditioned by the available alternatives. Nobody who eats purchases food for a random, unknown reason. They purchase food as the conscious decision to sustain their life. Second is the non-random variable that marks the supply of a good. Goods, and services as well, are physically given. There is never a random number of haircuts available, or automobiles for sale, but a definite number conditioned by the market's needs, and the availability of the physical world. Lastly, the entrepreneur's own ability to exercise alertness to these opportunities. It is not that enough that these opportunities solely exist, but that the entrepreneur has the alertness to know they exist (Hülsmann 1999: 64). Prices do not align themselves with the needs of consumers automatically; they require the entrepreneur to accomplish this feat.

Time Preference and Valuation

If prices in the financial world are conditioned by the underlying real demands that exist, they are also influenced by the time preference scale of the individual actors. This occurs in two ways: directly and indirectly. Indirectly, time preferences influence the temporal element of our demand for goods. Directly, we see that demand for financial assets compared to real assets are effected by our time preference as well. The interplay of both factors work to determine prices in a non-random manner.

When an individual has a high time preference, we may say they enjoy present consumption more than future consumption. The converse is a low time preference, where there is no such pronounced preference for the future or the present. Time preference is not an abstract concept, but manifests itself through the actual demand and supply of goods in the present. Hence, our time preference dictates, in a general sense, the quantity of goods available in the present. This is where part of the source of the entrepreneur's profit opportunity comes from as was previously discussed. The coherence that exists between the present demand for goods, as dictated in a sense through time preference, and the supply that businesses have created to satisfy this want.

Second, an individual's time preference will direct their portion of their resources available for saving. This savings is generally invested into the financial realm: securities, bonds, etc. As a general increase in the quantity of savings directed towards the procurement of investment vehicles increases, this is equivalent to an increased demand for these same instruments. As supply and demand considerations for the actual financial instruments alters, prices will need to be adjusted through entrepreneurial action to account for any lags or discrepancies.

Some may argue that time preference is a fundamentally random variable, that we have no control over it directly, and hence, it shifts uncontrollably as a random variable. This is patently false. We can see, for instance, the effects of age on time preference. Retired people generally have a higher time preference, spending more and saving less than the average suggests. Likewise, middle-aged people may have a very low time preference. They prefer to save a higher portion of their resources for use later in life. Likewise, the personal situation of an individual may affect their scale. Suppose, for example, that an individual wins the lottery, and suddenly finds themselves with one million dollars. Two possibilities may occur with these new resources. The first is that the money is saved, thus signaling that there is a generally low time preference for the actor, and a desire to use the money later. Conversely, if they use the money in the present it signals that the actor has a high time preference, preferring to consume in the present over the future.

Which ever case may occur in particular, the impetus is the same. A purposeful decision was made on the part of the individual. Whether they decided to spend or save the money is not randomly given by the throw of a dice, but rather decided given the opportunities and avenues to utilize the new money that exist at the present compared to the future. Time preference is always, and will always be, a purposeful trade-off.

Value Scales and Valuation

Value scales dictate how important the utility of a good is to an individual. This always refers to utility towards servicing a given end. As such, as ends change, or the importance an individual places on an end changes, the relative rank on their value scale will be altered accordingly. Much like time preference, there are two effects that concern us here. The first is the relative rank of financial assets compared to real ones. As individuals acquire ends that include the use of financial assets, they will place these instruments relatively higher on their individual value scales. Additionally, as independent ends change, the underlying goods needed to service those ends will also alter. This provides a shift in the value scale position of these goods, and as a result, the ultimate valuation of a financial asset representing a physical good will also change. Neither of these changes can be viewed as random-variables.

In a broad sense, like we saw regarding time preference, individuals can value assets in the financial realm, or assets in the physical realm. The relative location of these respective ends will determine the comparative level of valuation that results. For example, an individual decides to make an investment for the future. There are two broad methods they could use to pursue this end. The first is investing in some physical entity. Maybe they will buy a house in the expectation that the house will increase in value and deliver a respectable return in the time span they intend to hold it for. Alternatively, they might expect the same fundamental action to occur on the market place, that is, the appreciation of housing prices. However, the method they use to capitalize on this end is to buy a financial asset, maybe a security in a mortgage company that they think will enjoy successful and profitable business from the housing boom. Which ever choice the actor pursues, one or the other method will be ranked higher on their personal value scale.

In each case above, the end goal is identical – the capitalization of the expected profits resulting from the housing market's boom. The method employed to pursue this end may differ as well. It may well be that each choice may yield the same profitability *ex post*. We might be able to say that in this event, the actor would have been indifferent between the given delivery method. But action is not concerned with *ex post* results, it is concerned with *ex ante* expectations. We know that indifference does not exist, otherwise we would live in an actionless world. The result is that one avenue will always be chosen, and this "chosen one" will be the method ranked higher on the individual's value scale. This in turn affects the relative demand for given means, and hence, the valuations that will be placed on them.

The second value scale shift we see concerns demand for individual goods. As ultimate valuation is dependent on a good's ability to service an end, as the ends sought by an actor change, the relative valuations change as well. Changes in ends desired are not without reason. The shift from whale blubber to petroleum as a fuel may serve as a useful example. The shift put significant downward pressure on the valuation of whale blubber. The reason is simple, it occurred as a result of the increased emphasis that was placed on petroleum. This was a conscious decision of many market actors working together. The advantages of petroleum created opportunities for entrepreneurs to exploit. As they worked to place it higher on their own value scales, while simultaneously moving the existing whale blubber down, we can see that in doing so they provided the pressures needed to adjust prices to their relatively more desired positions.

The decision that marked the shift from blubber to petroleum was not random. It was calculated and conscious, the result of the serviceableness of a specific good, relative to others, in attaining an end. If there existed two companies at the time, and the whaling company's security's price declined, and the oil refining company's security price advanced, we can see how futile it would be to attribute this to a random event. It was not an unknown event that altered the respective prices, but rather the process of substituting one good for another on the value scale.

Value scales are determined by the respective utilities of given goods relative to one another. As the serviceableness of a good changes, always relative to an alternative, it position on the scale will alter. Additionally, we can see that it is not just the position on the scale for goods that effects the ultimate valuation of a financial asset, but the position between the desire for real and financial assets that exists. As a shift moves towards valuing financial assets higher, this increased demand will manifest itself as a higher position on the scale, and create an opportunity for an entrepreneur to adjust the market price accordingly. Just as the original decision to alter the relative location of a good on the value scale is not random, the entrepreneur's decision to exploit the existing price differential is not randomly given either. Instead, it is purposefully determined, through the entrepreneur's desire to maximize profit, through adjusting prices to the wants of consumers' utilities functions.[120]

The Marginal Insight and Valuation

One of the main tenets given through EMH is that all actors are imbued with identical knowledge. As information is instantly manifested into the pricing of assets, all information is available to all actors in exactly the same way. The result is that the pricing of assets is not achieved on the margin, but is rather the result of the aggregate average of the actor's expectations. The loss of the marginal insight results in a major failing of the way that EMH must view valuation.

[120] Mueller (2001: 14) points out that price 'bubbles' exist on a threefold basis: (1) a monetary policy that allows excessive money growth, (2) authorities inducing a moral hazard element, and (3) investors are exposed to a learning process where their expectations are altered and risk perceptions diminished. We see that the academic debate as to rational versus irrational behavior is lost, there is only purposeful behavior causing events to pass.

The foundation of EMH is that all information that is relevant to the pricing of an asset is instantaneously and perfectly reflected in its price. The implication is that all actors must have access to the same information. Hence, if all have access to the same information, and all prices currently reflect all the relevant information, then prices must all be set according to the same information. A problem emerges.

The marginal insight generally identified that a buyer is never faced with buying the whole stock of a good at once, only a given amount. Hence, they never value the whole outstanding utility associated with it, only the utility they can garner from the limited amount they desire to purchase. The corollary is that sellers are faced with the same dilemma. They are never faced with selling all the available quantity of a given good, only a portion is at their disposal.[121] The part that receives little attention, in some circles, is that available information is also marginal. As Hayek (1973: 30) would point out, we can never act with all the information of a given action, only a portion of it. Action on the margin applies not just to control of the stock of good, but to the information available concerning it.

The assumption, or conclusion depending on your point of view, of EMH is that all actors act with the same information, and also that the information they have is the entire relevant amount. The aspect of *informational marginality* is removed.

In a marginal world, the interplay between demand, supply and price is mostly determined by the marginal pair of actors. Demand and supply are dictated in relatively absolute terms, and the buying and selling of a good will continue until the number of units available to be bought equates the number of units available for sale. This in turn is determined by the price at which the marginal pair of buyer and seller operate at. Hence, for marginal valuation, the key aspects are the two individuals who equate the supply and demand. Individuals selling at a price higher than the market price, and buyers offering lower than the market price, are thus seen as essentially not a direct part of the market. Likewise, actors of the market who are not the marginal pair influence the market price, but only in an indirect way through their exerted supply and demand pressures.

When we view all individuals as having and utilizing the same information, we can see that the actors at the margin are removed; *everyone is now the same marginal actor*. As all actors will have identical expectations, given by the same available information, all will be interested only in buying and selling securities at an identical price.[122]

[121] We have previously looked at the special case where the supply of information is wholly determined by any one individual holder of it.

[122] Although, as Grossman (1981: 556) points out, one of the roles of prices is to aggregate information. This is the exact opposite of what EMH proposes, whereby information is aggregated and prices are homogeneous as a result.

The real world of asset pricing is markedly different, and the difference stems from the treatment of information availability. We have already seen that information is not given to anyone identically or in entirety. In fact, as knowledge (action-knowledge to be precise) is the direct determinant of action, we see that information has no direct influence on the pricing process. It is only that information which actors have personally obtained, and amalgamated into their knowledge base that is utilized towards action. Also, as Polanyi and Prosch (1975: 61) have pointed out, all knowledge is necessarily distinct of all other knowledge. There can be no objective homogeneity inherent in it, unlike its information predecessor. As each actor is individually in possession of knowledge, which they have personally obtained, and which is conditioned by their own subjective valuation of it, their expectations will also differ.

In the real world marked by such asymmetric knowledge, we see that marginal pairs of actors must exist. The assumption that is given through perfect, homogeneous knowledge is that all valuations will be identical, and hence, marginal actors are eliminated. In our world of imperfect knowledge we find that marginal actors must necessarily exist, through the differing expectations each will possess. Prices are no longer the result of one shared belief attributable to all actors, but to each acting as a unique individual given their own available means. In early 2008, when crude oil broke the $100/bbl mark, it was not because every actor saw the same expectation for the future. It was due to the sole reason that one individual expected to make an expected profit that was desirable from buying at that price, and that one individual expected to make a desirable profit by selling at that price. That they both utilized unique sets of knowledge must at this point become evident. As Lachmann (1978: 40) explained it, "[t]he *market* is no substitute for the decision-making unit"[emphasis added].

Conclusion

The idea of random prices that result from EMH adherence can be seen to be based on fictitious, unrealistic assumptions. Most importantly, the belief that all information is available, and acted on the same way by all actors is found wanting. Prices can never be a random future variable.[123] They are, in

[123] The devastating effects this conclusion bears can be seen from the emphasis placed on random variables in financial modeling in this passage from Dothan (1990: viii):

> The central mathematical concept of the theory of financial markets is the stochastic integral. The stochastic integral is basic to the theory because it describes the gain from trading in securities. A large part of the mathematical development ... leads to the notion of a stochastic integral, its properties, and its role in the theory of financial markets.

nearly every respect, the result of purposefully human action. As Shostak (1997: 30) has demonstrated, much of the confusion has stemmed from a disjointing of the real and financial arenas. Hence, prices are no longer tied to their underlying goods. However, as Mises (1949: 810) pointed out, "[t]he success or failure of the investment in [securities] depends ultimately on the same factors that determine the success or failure of the venture capital invested. There is no such thing as independence of the vicissitudes of the market." However, we see that at the root of valuation is the subjective emphasis placed on the serviceableness of a mean for a future end. As such, we see no reason to disconnect the financial realm from its real counterpart. As Lachmann (1978: 86) points out, capital markets are not markets for physical goods, but titles to them. That these are based then on the values of the representative physical goods, and serve to promote capital change, goes without saying.

Malkiel's (1973) position that a monkey throwing darts at a board could choose stocks as effectively as professional analysts missing one major point. Monkeys throwing darts randomly are not an equilibrating force (Pasour 1989: 102). Prior to reaching this stage, we must assume the monkeys already know much knowledge to narrow their actions down to this specific one. For instance, the monkey must know what it is aiming for, what a dart is, where the board is, how to throw, and much more. A monkey throwing, hitting, and selecting stocks, is not a random action, but presumes much integral knowledge already exists. Investors likewise must consider much knowledge, and, purposefully utilizing this knowledge, act in a way that drives the market towards its unreachable equilibrium.

Additionally, we have seen that valuations are not solely determined directly by the utility of a mean, but also by two additional factors: time preference and the individual's value scale.

Time preference dictates two main points. The first is when an individual's consumption will occur. The differential that exists between present and future wants will partly determine the valuation placed on these means. The corollary to this is that what is not consumed in the present is saved. Savings manifests itself primarily through the financial realm. Hence, a decrease in time preference can place increased emphasis on saving, possibly manifested in financial assets.

Individual value scales also act similarly to time preference. Previously, an emphasis has been placed on the place of an individual good on an individual's value scale. We have seen that the shifting of groups of assets on the value scale can also occur. As financial assets become more valued in

---

As stochastic calculus concerns itself with the ability to model randomly behaving systems, we can see that the importance placed on its usage may be misplaced in the realm of finance.

relation to their physical counterparts, an upward relative pressure will be seen on their prices.

Both time preference and the value scale are not given by random variables. They are the product of human action. Hence, we can see that all input factors to the value of an asset to be purposeful decisions, implying that future prices, although embodying the degree of uncertainty given by their existence in the future, will not be purely randomly determined.[124]

Lastly, we saw how the marginal insight of valuation has been eradicated within the EMH framework. This closes the door for any gain from additional information, and thus precludes the entrepreneur from functioning. We see that this is clearly not reflective of the real world. Future prices will forever bear the uncertainty inherent of the passage of time. This is a much different statement than stating that future prices are a random variable. As Hoppe (1997: 65) states, "I can predict the general, logical structure of each and every one of my actions, whether past, present, or future." With the successful attainment of the relevant information, the future price can be approximated, due to the intimate relation that exists between valuation and purposeful human action.

[124] This ignores the influence of random "acts of God" that may wreak havoc on valuations. To the extent that these are relatively infrequent occurrences, we shall exclude them from further analysis with only a minor comment. Even random variables attributed to weather, for example, are able to be mitigated somewhat by insurance companies. Although unable to perfectly determine when these events will occur, we can see that insurance companies cope with their inherent possibility through probability analysis.

Appendix A: Logic and Reason

A great deal of confusion surrounds two similar concepts: logic and reason. Economists are often fond of saying that individuals are endowed with rational expectations, or that they behave rationally. Mises (1949: 19) would state, "[h]uman action is necessarily always rational. The term 'rational action' is therefore pleonastic and must be rejected as such."[125] Often times, logic is used as a substitute for reason, as if they were two interchangeable concepts. They are not, and a short clarification should satisfy in rectifying the situation.

EMH rests on the assumption that actors act rationally. Rationality is always a relative concept. It depends on the specific end that an individual seeks given the available means they have at their disposal. As such, there is no gauge with which to define true rationality. A fall-back position for EMH advocates could possibly be to alter and say that investors may not always act rationally, but they do act logically. This is also erroneous. Humans have a distinct ability of being able to suppress our logic if we wish.

When we discover the fallacy of the rational actor that EMH rests on, we can see a significant problem in its message. The assumption that all actors act rationally is true, but the definition of rationality is false. Rationality can only exist for a given end. The conclusion yielded through EMH that all actors act in a homogeneously rational way is incorrect, and leads the the erroneous conclusion that all will act towards the same goal.

Reason

The American Heritage Dictionary (4th ed.) defines rationality as "consistent with or based on reason" or "exercising the ability to reason." Rationality is a personal criteria. It can only be known to the individual. This is necessarily so as whether an action is rational or not will depend on the viewpoint of the actor. It will thus depend on their means available, their knowledge levels, and expectations of the future.

Due to this personal nature of rationality, we can never say if an action was irrational or not. We can only say that, at the moment in time when the action existed, it was rational in the mind of the actor. If it was not, the actor would not have undertaken it. We can therefore never predetermine what

[125] For Mises, rationality and purpose would be one and the same. As all action is to be considered purposefully directed, all action must also necessarily be rational.

rationality is. It is always a case by case basis, taking into account the momentary circumstance that the individual is faced with. Huerta de Soto (2005: 49) argues that this is not so, that outsider can always deem another's actions as being irrational, based on their own subjective interpretations. However, the very concept of rationality is one of relative knowledge, which precludes the possibility of another, other than the actual actor, passing judgment on an action; *only one with identical information and an interpretation thereof could be in such a position*.

Mises would refer to ends as always being the viewed as rational. They are "ultimate givens", and thus can only be viewed from the mind of the actor. Rationality is a concept that can be applied to both means and ends, but only from the perspective of the actor. At any given point when an action is undertaken, we know that it was a rational decision from the eyes of the actor. This may change in the *ex post* sense, but only an the individual actor could decide this.

Logic

The America Heritage Dictionary (4th ed.) defines logic as "the principles of reasoning, distinguished from their content."[126] The "distinguished from their content" part is what separates logic from rationality. Logic is an absolute; it is separate from the relativism that marks reason. We can observe actions of another individual and determine if these are logical or not in a way that will never exist with rationality. One of the distinguishing characteristics of the human condition is our ability to set logic aside and perform functions that are illogical.

We can, in this sense, say that logic is an absolute standard. It involves all the information that exists, not solely that at the disposal of the actor. Furthermore, logic is only applicable to means, never ends. We have seen that the use of means must always be deemed rational from the viewpoint of the actor, however, with logic we can absolutely state if they were logical or not. Take an example of trying to drive from point A to point B. It could be that an individual without a map will try to complete the trek, but get lost halfway through. They will stop and ask for directions, and eventually make it to point B, hours later than was possible. An observer may have knowledge of a road that connects the two points directly, and, if used, would have allowed for the trip to be completed much sooner. The observer does not know the reason why the actor took the route they did. Only two conclusions can be drawn from the observer's perspective then:

[126] Likewise, Rothbard (1998: 10) notes that "[m]an's reason is *objective*, ie., it can be employed by all men to yield truths about the world."

[1] The route that was taken must have been fully rational in the actor's eyes. There exists a reason why they traveled the way they did, and this is unknown to all but the actor. It is shaped and conditioned by their personal knowledge and circumstance.

[2] The route taken was wholly illogical from the observer's perspective. There was a method that was much simpler, and faster that was not utilized. The end was the arrival at point B. There was an easier method to fulfill this end than the one taken.

Both cases involve the observer knowing what the end involved is. It is only in this way that an individual can gauge the means necessary to obtain it. In one case only the means at the actor's disposal can be analyzed, the other concerns itself with total available knowledge. Also note that a separate observer may also call the action illogical, but for a different reason. Perhaps there is a flight that would have achieved the trip at less cost, and shorter time, then the route undertaken via land. Also note that the end must be identical to all actors. If the original end sought was to move from point A to point B using only a car, then the second observer's viewpoint would not matter. It would be outside the realm of analysis *given this particular end*.

Mises (1949: 103) would clearly delineate the difference between logic and reason:

> One must not confuse the logical concept of consistency (viz., absence of contradiction) and the praxeological concept of consistency (viz., constancy or clinging to the same principles). Logical consistency has its place only in thinking, constancy has its place only in acting. Constancy and rationality are entirely different notions. If one's valuations have changed, unremitting faithfulness to the once espoused principles of action merely for the sake of constancy would not be rational but simply stubborn. Only in one respect can acting be constant: in preferring the more valuable to the less valuable. If the valuations change, acting must change also... A logical system must be consistent and free of contradictions because it implies the coexistence of all its parts and theorems.

Logic presupposes an amount of absolute knowledge of an entity. We know however that in a physical sense all the knowledge of something cannot be had. Action therefore may not be viewed as logical at all times. However, reason, or our rationale for action, depends solely on the finite amount of knowledge we have regarding it. From this viewpoint of the individual actor, all action will be necessarily rational – there could be no other known manner it could be. Knowledge of something not known by us cannot be an influential factor in establishing rationality. To an outside observer with knowledge of this relatively unknown information, an action can most definitely not be considered logical *in any absolute sense*.[127]

Bounded Rationality

Simon (1957) proposed that actors are fully rational, but within a constraining sphere. His theory of "bounded rationality" posits that individuals do not reach fully rational decisions, however, they reach rational decisions based upon the limited knowledge they may have concerning an event. As Kirzner (1973: 159) points out, choices are made not only on what is known, as some options exist that will not be known. Hence, full-rationality, at least in a neo-classical sense, is impossible to achieve due to the incomplete information available for any decision. This comes mainly as the result that:

> [T]he capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behavior in the real world – or even for a reasonable approximation to such an objective rationality. (Simon 1957: 198)

For Simon, the main limitations this incomplete knowledge may serve manifest in two primary ways: (1) limits in the utility functions an individual can formulate, and (2) inability to properly assess the costs of gathering and using information.[128]

Under bounded rationality, a decision is achieved in light of the fact that an individual may not be able to assess all the possibilities in an environment. Therefore they cannot make an optimal

[127] In a sense, every entrepreneurial action may be considered as *illogical* from an outsider's perspective. As Mises (1949: 501) states, "[t]he [profit] goes only to the dissenters, who do not let themselves be misled by the errors accepted by the multitude. What makes profits emerge is the provision for future needs for which others have neglected to make adequate provision." As any entrepreneurial action will by necessity preclude an outsider from knowing its full cause or reason, we see that outsiders must always consider entrepreneurial action as unnecessary and hence, illogical. See also Hazlitt (1974: 759) whereby it is shown the entrepreneur does not act because there *is* an opportunity, but because they *perceive* there to be an opportunity. If their perception or alertness if not sufficient, they could act in an inappropriate way, and that seems illogical to the outside observer.

[128] For Knight we saw that true uncertainty and novelty constrains the entrepreneur, for Simon, we see that complexity in the decision making process constrains their ability (Sautet 2000: 51 n64).

decision in light of all the facts, hence in Simon's view they maximize the part of knowledge they do know, or “satisfice” upon it. As Sautet (2000: 29) summarizes, “[t]he implication of bounded rationality is that individuals will economize on their (bounded) rationality in their decision processes on the one hand, and setting up corporate governance structures on the other hand.” Hence, Sautet views the limitations that bind us as being responsible for the creation of social structure in society (particularly firms), instead of allowing all processes to be performed through market processes. He (*ibid*.: 29n38) views rationality as a scarce resource. This is true, in the sense that logic is *quasi*-scarce. We all have access to all of our supply of it at any time if we wish, but it can be developed over time, and hence, used more efficiently.[129]

We see that what Simon refers to as bounded rationality is really just an exercise in logic. Rationality always assumes that we know only what we know. Dequech (2000: 171) views rationality and knowledge as inextricably linked; rational is that action based upon knowledge pertaining to that action. All action is rational as a result, from the perspective of the actor explicitly, and from the perspective of the observer implicitly. It is, in fact, only this actor who can determine if an action is rational or not, and by the act of performing the action, we know it must be rational in their eyes. Logic, on the other hand, concerns all the options available. In an absolute manner then, we see that logic is necessarily always limited by the amount of knowledge an actor has at their disposal. As it is absolute, and outsider has the ability to state if an action is logical or not, as they may have access to subsidiary information that is essential for the decision making process. Hence, bounded rationality is really a term to describe logic, as outsiders to an action view the same action.

The Negation of Logic

An important corollary to our example above is that logic is able to be negated by the acting person, but rationality will always be fully operative. Now let us change our example and assume the same ultimate end is involved, getting from point A to point B. In our previous example our actor had limited knowledge of the road network. Now we will change it so that the actor knows the shortest route in existence, that is, the “optimal” or logical route. However, they still choose to take the previous path, which involves a greater distance, and more time and cost than the known direct route.

[129] As Mises (1975: 102) viewed economic calculation as the method we economize, or satisfice, on this complexity. “[N]o single man can ever master all the possibilities of production, innumerable as they are, as to be in a position to make straight-away evident judgments of value without the aid of some system of calculation [i.e., economic calculation].” See also Mises (1949: 698).

We can see that regardless of the explicit goal being the attainment of point B, the path taken besides knowingly being a sub-optimal path, will be fully rational in the eyes of the actor. They will have taken this route for reasons known only to them. However, this same actor may realize that this choice is fundamentally illogical. A technically more optimal means of getting there is available *and known to be* available. They have decided to act against this logical setting and take a separate route. The logic can be bypassed, always knowingly, but rationality cannot.

The suppression of logic may have no *explicit* reason for the observer, and hence may not appear to be rational. But to the acting human, suppression of logic will always be a fully rational action.

EMH and Rational Expectations

EMH assumes that all actors utilize, and formulate, expectations rationally. The rational expectations hypothesis (REH) is inextricably linked with EMH in this regard. All market actors thus achieve expectations based on present information that faithfully represents this information and shows no systematic biases. However, as Hoppe (1997: 57) demonstrates, our knowledge and expectations can not be homogeneous, as this would have devastating effects on action:

> If everyone's knowledge were identical to everyone else's, no one would have to communicate at all. That men do communicate demonstrates that they must assume that their knowledge is not identical.[130]

This is true, but only from the standpoint of the acting person. Rationality is an individual concept, it can never be gauged by an outside observer. What the purveyors of EMH are confusing are two concepts, albeit very similar, but different in scope. Logic is a metric that can be assessed absolutely. EMH advocates really imply that an actor will always act logically toward a given end. But as we have seen, a defining characteristic of humans is the ability to suppress logic, and to act fully rationally at the same time. The assumption therefore that all actors choose logical actions is erroneous.

Take an investment example. An individual may buy a stock for the sole purpose of generating

[130] See also Hoppe (2007: 15) for the view that human's communicate to acquire others' personal knowledge. Stiglitz (1991: 137) defends the use of full rationality in economics: “The deficiencies in the 'rational actor' model have long been recognised, but economists have defended their pursuit of the rational actor model on the grounds that it was the best game in town... while the alternative was a Pandora's box – there was an infinity of possible irrational behaviors.”

the highest return possible for their risk-profile. Let's assume there are two options to choose from:

[1] Stock A pollutes the air, tests chemicals on animals, and uses child labor. Its return is 10%/annum.
[2] Stock B is “ethical” and does none of the above. It exists to help needy children. Its return is 5%/annum.

If both stocks offer similar risk measures, we can see that with the end in mind of earning the greatest return possible, the individual would prefer company A. This is the fully logical choice for the attainment of that particular end. Suppose that an investor decides instead to invest in stock B. This action is still fully rational, and logical to the acting person. However, for the onlooker, we can now say that with the end in mind of attaining the highest return possible, this is a purely illogical choice. The investor has sacrificed 5%/annum of return. In no way can the observer ever say that this choice was not fully rational however. Rationality can only exist in the mind of the acting human.

We can further assume that the individual does not know why they have chosen B over A. The satisfaction of ends must occur in sequence, one at a time. The most pressing desire will be satiated before moving on to the next. Suppose the individual's most pressing desire was to help needy children. They also could have donated money to an organization functioning to that end. We can see that the individual may not realize why it was that they invested in the company, instead of donating the money directly – they may not personally realize their most pressing desire.

Or, in another situation, an individual may explicitly know what their particular most desired end is. Furthermore, they may know the most logically way to attain this desire. However, for a reason unknown even to themselves, they may choose a different path. They may thusly be acting in a manner they personally deem illogical. This too will be rational action for the actor though. Rationality comprises all the unknowns of our decision making process, the inarticulable benefits and costs that enter the decision making process. *Implicitly we know that a decision is rational, when all of its benefits and costs operating on the actor are factored for.* These may not be known to the actor personally, but we know that by the very act of acting, the actor has deemed the benefits *ex ante* to outweigh the costs. Logic concerns itself with all the known criteria that go into a decision for a given, known end. Logically, we can say there is a most correct way to use a mean to attain an end.

Arrow (1978) argued that the rational expectations assumption is diametrically opposed to Hayek's (1945) assertion that prices serve a role as summary statistics of information. Hence, Hayek argued, absent a market price system, there would be no way to collaborate all the necessary information to plan an economy in any cohesive manner. Arrow points out that a rational expectations assumption seems to imply that price-signals are not needed in the market economy. Instead, actors can, supposedly, correctly guess or infer what prices should be, or would have been. We see, however, that although prices may not be wholly accurate statistics, they are necessary to convey large bodies of information throughout the market in an efficient manner.

Hence, we see the EMH assertion that actor's behavior as being rational is correct, but rationality is not the correct expression EMH advocates mean to use. Logic is the more apt descriptor of investors' behavior, whereby each is using the means of information in the absolutely optimal manner. We see this is false however, as humans act in ways which are not always fully logical. They may purposely ignore information, regardless of the importance it serves in their investment decisions. They may over-emphasize information, even though they know its importance or relevance is not so great as they give it. This is action for humans, rational as it is.

Conclusion

The human decision-making process is one of relatives and absolutes. Rationality is a concept only applicable to the individual. It can only be personally gauged for oneself, and only for a given end. Logic also applies only regarding a given end, but can be assessed in absolute terms. All action is necessarily rational, but not necessarily logical. As such, economics is not a theory of choice, both of which are necessarily always to be considered rational, but a theory of processes which provide coordination despite the fact that these decisions may not be fully aware of all possibilities (Kirzner 1992: 201). Action may be illogical, that is, it may not consider all possibilities external to the actor that are known, and as such, may provide a disequilibrating effect in need of explanation.

EMH rests on an assumption that all humans assess information in a rational manner. This is an erroneous way of viewing human action. All humans act rationally, but this can only be seen from the view-point of the individual. Instead, the comparable standard that EMH really implies is that all humans act logically. We see this too leads to erroneous conclusions.

For a specific end, we see a distinctly human quality is the ability to suppress our logic, knowingly or not. We can use means in a way that are not the most logical way for the attainment of ends. This is an opinion that an outsider can make. Logic can be deemed in absolute terms regardless of the knowledge level of an individual and in light of the knowledge that an observer may have.

Some view an illogical action as a failing, something to be criticized or eliminated. In the monetary economy, many illogical actions *may* result in corresponding monetary losses. However, as Mises (1980: 114) points out:

> Nobody has the right to take offense at the errors made by the entrepreneurs in the conduct of affairs and to stress the point that people would have been better supplied if the entrepreneurs had been more skillful and prescient. If the grumbler knew better, why did he not himself fill the gap and seize the opportunity to earn profits? It is easy indeed to display foresight after the event.

As perfect information is never endowed in humans, actions deemed as illogical will forever occur, when viewed from an outside observer. These do not construe a problem for humans, instead, they only signal a constraint we shall forever be bound by.

Humans need not reflect information in a fully logical manner, but this action will necessarily always be rational from their own point of view. Greater emphasis may be placed on a given piece of information than it warrants. Its relevance may be incorrectly gauged. Humans can err in this manner. The explicit bedrock of rational action that EMH rests upon cannot be supported, and the implicit assumption of fully exercised logic is inappropriate as well. Pricing theory must account for the fact that humans need not assess all information logically in any absolute manner, but are instead able to err and create mistakes, whether wittingly or not. This is the essence of human action.

Appendix B: Comparative Constructs – EMH versus ERE

Mises (1949: 247) created an artificial construct, the evenly rotating economy (ERE), from which to ascertain changes that occur in market data. In particular, the ERE is characterized by two distinct elements. The first is the elimination of the temporal element, the second is the removal of changing market data. We can see that the second point necessarily arises from the first.

Is it possible that EMH, despite its flaws, may be used as a similar construct? If we envision a similar state of affairs as under the ERE, is it possible to grasp more fully the effect that information has on prices? We argue that it cannot, for two main reasons. The first is that the necessary assumption of a changeless environment necessarily precludes the existence of information. Second, we look at the final conclusion that we wish to gain insights from in this model. Its use would be to gain a better understanding of information on prices. However, like in the ERE, under a similar EMH world comparable money prices would be eliminated. Instead, a system of direct exchange would arise.

The use of an EMH construct, as a counterpart to Mises' ERE, is counterproductive. The one role that it would serve to elaborate – the effect of information on prices – would necessarily be excluded by its very construction. We find that EMH is not even useful as an abstract model from which to draw ancillary conclusions of the world.

Mises' Evenly Rotating Economy

Mises (*ibid.*) created an artificial construct, the evenly rotating economy (ERE) in order to explain the entrepreneurial function better by removing two elements, time and change. In the course of removing time from this conception, Mises does not mean that time ceases to exist. The removal of time completely would imply the non-existence of everything it is that we wish to see through the use of the ERE. Instead, the removal of time implies the removal of *Bergsonian* time. Newtonian time still passes, but it progresses as a static wave. It exists as the antithesis of time. It exists, market actors move through it, but they do not feel it. The essential point that makes time come alive to us as humans is removed – it incites no new knowledge in us, and alters nothing of us. We see that as a result of this removal of the essence of time, the economy of the ERE still functions, and progresses, but it does so without change. In Mises' (*ibid.*) words:

> The system is in perpetual flux, but it remains always at the same spot. It revolves evenly around a fixed center, it rotates evenly. Therefore prices – commonly called static or equilibrium prices – remain constant too.

Hence, as change is removed from the ERE, entrepreneurially acting humans are likewise removed. There are no longer disco-ordinations that require attending to, nor is there an equilibrium in the distant future to be reached; it already exists in the present. In Mises' (*ibid*., 245) view, this was not a wholly artificial construct, but instead represented an "adequate description of what happens again and again on every market." If time ended at any given moment, this would be true. This state of the ERE is achieved in the sense that the state of affairs that has occurred, and has taken all of history to reach this point in time, ends at every moment. Hence, the ERE can describe accurately the world that exists at any given point in time – Newtonian time. However, we see that the continual passing of this temporal element, the ERE can never be reached. *It is a realizable construct, but only within its own known boundaries*.[131]

Mises utilized this construct, well aware of its limitations for one purpose; as a purely abstract construct to demonstrate other principles, in particular, entrepreneurial profit and loss (Gunning 1989: 124). It was only in constructing a system that removed the entrepreneur wholly, that changes concerning this role could be introduced and analyzed.

The use of this construct has seen increased use over time, but has also come under fire. Cowen and Fink (1985) argue that the ERE fails in its role as it: fails to be totally unrealistic, and is internally inconsistent. However, they have demonstrated a poor understanding of the ERE's construction, and its proper use. It was never created to be wholly nondescriptive of reality, this would serve against everything Mises forwarded as being methodologically necessary for economics. Instead, it was to be sufficiently unrealistic so as to demonstrate the problem intended – entrepreneurial profit and loss in the market. Secondly, it is only viewed as internally inconsistent if it is misused, as they portray it to be. Mises made note of the fact that money ceases to exist in the ERE, as uncertainty is a natural precondition for the existence of this factor. This only precludes the possibility of explaining monetary changes through the ERE, something that it was never originally proposed to do.

Furthermore, Cowen and Fink (*ibid*.: 867) fault the ERE as not sufficiently explaining how it

131 As Garrison (1984) reminds us, there is no significant issue *theoretically* by viewing single actions in this static manner. All action does, after all, transpire within its own single moment. It is only the masking of the processes that connect singular actions together, or the market process, that creates a significant issue when viewing the world to exist in this unrealistically static construct.

would be reached. In their eyes, the removal of change and time does not sufficiently explain why an equilibrium, such as is described in an ERE, would come to exist. However, as Garrison (1991: 95) has countered, “[i]t is not necessary for the initial conditions to preclude all kinds of disequilibria but only to preclude systematic intertemporal disequilibrium – the kind of disequilibrium for which the theory itself accounts.” Again, we see that when used *within its own limitations* the ERE provides a consistent analytical tool. Lastly, as Gunning (1989: 126) points out, the definition that Cowen and Fink use regarding equilibrium is mistaken:

> The problem with [Cowen and Fink's] criticism is that it is based on a *mathematical* definition of equilibrium and not a *logical* definition consistent with Mises' pure logic of action. In the logical definition, the concept of disequilibrium is meaningless. To say that there is a tendency toward disequilibrium is like saying that individuals do not make choices.

Hence, we see that Mises' ERE is an invaluable tool, provided one knows its limitations, and uses it appropriately. Its construction as a partly unrealistic representation does not fault its results, but instead gives added meaning and clarity to them.

Fama's EMH as a Static Construct

With this in mind, is perhaps the conception of EMH as a static construct a meaningful way to look at the world? We may envision a world where there is no *Bergsonian* temporal element, but the creation of information is still occurring. In this world, we see that information is disseminated instantly and costlessly into the market. Hence, as preferences are not changing, what is the result that is achieved?

We see that EMH is insufficiently suited to gauge changes that occur in this world it creates. An EMH world would, by definition, limit itself to effects that occur on information within its realm. In particular, we are interested in seeing what occurs regarding prices when influenced by a new informational element. However, we see that, like in the ERE, in an EMH world prices cease to exist. The uncertain element is removed from this world due to the change in temporal conception. As a result, information in the present must necessarily conform to this known future. As prices converge upon their 'long-term equilibria' we find that there equilibria will be, like in the ERE, monetarily valueless. As Thomsen (1992: 37) informs us, it is only in disequilibrium that prices provide their informational role. A movement to equilibrium would negate prices, thus eliminating information from having a value in determining them.

Thus, the removal of uncertainty from the sphere removes two things which are central to the analysis we wish to undertake. The first is information. The essence of information is that it is fresh, and has an effect on our current, and future state of affairs. The corollary to this is that information exists, and is sought, in the present partly as a result of an uncertain future. In an EMH world, as this future uncertainty is removed, information loses this important influence that it once held. In fact, it becomes valueless – we already have all the information we need.

Secondly, as in the ERE, money prices are no longer necessary in the EMH world. A system of direct exchange commences, precluding the need, and use, of an indirect factor. But the effect on prices is what we are supremely interested in by using this construct. The removal of the existence of this factor seems to leave us further away from the goal we search for – the reason behind the formation of prices.

Conclusion

Is the EMH a proper static construct to start from when viewing changes occurring in the market, similar to Mises' ERE? Beechey, Gruen and Vickery (2000: 23) believe so and forward that "[t]he efficient market hypothesis is almost certainly the right place to start when thinking about asset price formation." However, despite being theoretically similar, we see that EMH cannot assume a role as an abstraction similar to the ERE. Part of the reason why may be gleaned from Mises' (1949: 248) own description of the tool:

> In order to grasp the function of entrepreneurship and the meaning of profit and loss, we construct a system from which they are absent. This image is merely a tool for our thinking. It is not the description of a possible and realizable state of affairs. It is even out of the question to carry the imaginary construction of an evenly rotating system to its ultimate logical consequences. For it is impossible to eliminate the entrepreneur from the picture of a market economy. The various complementary factors of production cannot come together spontaneously.

EMH also cannot be taken to its full logical conclusion, but if it could, the result would take us further from the answer we seek than when we begin. We seek to determine the effect that information serves on the prices of securities. However, we see two things develop as we enter the EMH world that complicate this possibility. The first is that information ceases to exist, and hence, loses all value to the acting human as the future is assumed certain. Second, prices are eliminated as a result of this same removal of uncertainty. Hence, the two subjects we wish to study become an impossibility: prices and information. In contrast to the ERE, we find that EMH is too unrealistic to provide insights into these elements, *even when used as a purely artificial construct*.

III. THE CAPITAL ASSET PRICING MODEL – AN AUSTRIAN PERSPECTIVE

Like EMH, CAPM suffers from several specific fallacies stemming from its assumptions. Also, like we have previously seen, many of these false ideas are continually viewed in a general way. The result may be that they are coherent, based upon their assumptions. However, when we stop to look at the individual building blocks of many ideas, we can see significant cracks beginning to form.

Since Markowitz (1952) the central idea of financial pricing has been that return is the price of risk. It is also assumed that the only choice facing investors is between this trade-off. However, we find that there exist a multitude of options that actors are concerned with when purchasing assets. Some of these will be articulable, and others not. Likewise, even if consumers were only concerned with this trade-off, producers are continually searching for methods to differentiate their products and reap further value from them. This implies that values do not automatically trend in a market towards an equilibrium, as CAPM must assume, but instead are purposefully shifted through the action of market participants.

Additionally, we see the importance of the risk-free asset may be misplaced. In reality a truly risk-free asset cannot exist. First, a look at risk is offered. The idea of variance as a true measure of risk is found wanting. However, even if we allow this point to pass unscathed, we still find that risk must continually exist. This is so as the passage of time and uncertainty of the future virtually guarantees it. Variance will exist as it is indeterminate if the *ex ante* duration an asset is expected to be held for may not match the reality revealed *ex post*. Additionally, even if we allow the fact that no variance in returns exists in a nominal sense, we find that in real terms, values are constantly shifting. For a domestic actor, this consideration means that any inter-temporal action will have a dynamically changing value inherent in it. Foreign actors face an additional uncertainty due to foreign-exchange risk.

Even if the risk-return trade-off is viewed as valid, and that risk is an acceptable measure, we still find fault in the method used to calculate risk. By viewing variance as the cause of risk, we can see that positive deviations from the mean are as negatively viewed as positive. However, the largest issue we find with the use of the 'variance' measure for risk is that it masks the task at hand behind a generalized metric. The thinking economist is not concerned primarily with the results of action, but the reasons thereof. As such, looking at variance as the source of risk underscores the source of the variance. The search for this should be the true task we face, not the retelling of a metric readily available to all.

CAPM rests on assumptions both faulty and erroneous. When we look at these two – the risk-return trade-off and the existence of the risk-free asset – we find that the issues have far-reaching effects for the model as a whole. Roll's (1977) critique of the non-existence of the market portfolio has been the most comprehensive critique of CAPM to date. However, we find that these additional two points make the model wholly unworkable in the world humans live in.

1. Homogeneity and Choice

The prime conclusion from Markowitz (1952) was that return is the price of risk, and that not all risks are created equally. A higher return would be have to be given for a higher non-diversifiable risk level. All development from that point on assumed that investors sought increasing returns by increasing their risk exposure. Hence, the trade-off for the investing actor is binary – risk or return.

However, we can expect that many factors comprise the decision to purchase an asset besides the expected risk inherent in it. Choice is concerned with a multitude of variables, some articulable and some not. These variables may not be time-invariant either. As circumstances change, an individual's criteria for choice will change. It follows naturally that these criteria are not only different for different time periods an actor experiences, but between different actors as well.

The existence of arbitrage is assumed to equate expected returns for assets of similar risk profiles. Entrepreneurs will keep asset prices coherent with what the expected risk dictates the return should be. The existence of arbitrage within this framework implies however that expected risk will be priced identically between different securities. This leaves no room for subjective valuations on the part of actors concerning securities that are not homogeneous, but rather heterogeneous entities. It also implies that all actors must share a homogeneous expectation of the future expected return and risk inherent in a security.

Roll's (1977) famous critique of CAPM rested on the assumption that the market portfolio was knowable. As he pointed out, the market portfolio must include all assets an individual may own, for example their house, car, watch, etc. We see also that the construction of the market portfolio is not only impossible, but the construction of the optimal portfolio must be as well. When looking at returns that must be incorporated, the individual must also look at expected returns for every asset. For instance, a raise in their job may be seen as an expected return much as an expected dividend is for a security. The fact that the returns of such eventualities are unknown creates additional complications for the computation of the risk-return profile.

CAPM rests upon a binary choice. It assumes individuals can only choose between the possibilities of risk or return. The real world contains an endless stream of factors that an investor would chose from. Additionally, when we view securities in a competitive environment, we see that these choices are not static. Entrepreneurs continually enter the market creating new differentiating factors for goods, or in our case, securities. Marketers continually create new exposure levels for securities, enticing individuals to purchase them for reasons unrelated to risk or return. In a world of dynamic, shifting preferences, the risk-return trade-off falls short of the plethora of options the acting individual has at their disposal.

The Nature of Choice

Markowitz (1952) tried to demonstrate that securities are fundamentally homogeneous. As such, when selecting one to choose to hold as an investment, actors face a binary choice: risk and return. Thus, an individual has no other factor with which to differentiate securities by. As Shostak (1997) demonstrated, however, there is a glaring disconnect between the financial realm and the real-world these assets are based upon. The result is an ignorance of the factors that contribute to the physical entity that a financial security is based on.

Firms exist to fulfill a multitude of consumers' needs. As a result, there exist numerous, or more aptly, countless, options that firms use to differentiate themselves from one another. For instance, we can look at the breakfast cereal industry. Every grocery store has a shelf solely dedicated to this product. We can see that the consumer is faced with a multitude of options concerning their choice of which one to purchase. If all breakfast cereals were valued according to the same metric, we could see that all consumers would gravitate towards this cereal. Eventually, one of two things would occur. First, one cereal would exist to the detriment of all others. As this one cereal fulfilled the needs of consumers the best, no one would consciously purchase a different one. Hence, a sole survivor would be the sole market option, *provided that it continued to be the best option for consumers*. We see that this is decidedly not the case. In fact, the breakfast cereal aisle is in a constant state of flux. Brands are constantly changing, with new ones being offered and old ones being replaced.

We see there are two general reasons why no one cereal completely dominates the aisle over time. The first is the continual shift in consumers' wants. The second is the constant competitive force being offered between producers to differentiate their products.

Consumer wants and desires are not time-invariant. Instead, they constantly shift throughout the passage of time. Looking at breakfast cereal, a child may have purchased for them (as they themselves cannot make the purchase) a colorful cereal to make it more attractive to eat. A young adult may prefer a high-energy cereal to help them start their day. A busy middle-aged individual may forgo eating cereal altogether. And an elderly person may prefer to purchase a hot porridge over a cold cereal in the morning. As consumer preferences change, we can see that their specific ends change as a result. The means they use to attain those ends also alter. A person purchases breakfast cereal as an intermediary mean to satisfy an end (perhaps the nutrition they require in the morning, although there could be other ends involved). Value is subjectively derived from the individual ends sought. It follows that the value that an individual places on a means is constantly shifting, relative to the value placed on the end.

It also follows that just because a comparable product is on the market, it is not part of the market for the individual. We can not homogenize in any meaningful way a group of goods into a class of goods to satisfy a want for an individual. This is due to the personal nature of want satisfaction. Let us assume that every single individual has a want for a breakfast cereal. Can you group every breakfast cereal produced into one homogeneous group, to be compared against each other? The unequivocal answer is: *no*. This is so as means never exist in a general sense, they only exist relative to a goal, and their ability to service a goal considering the needs of an individual. Hence, a diabetic breakfast cereal buyer's breakfast cereal market will not include sugary cereals. We often say that someone is 'in the market' to purchase a good. The only market that we can meaningfully speak of is the market that satisfies the individual's specific end, given their wants and criteria for means.

Therefore, we see the multitude of markets that are able to exist concurrently. Our diabetic may only be in the market for sugar-free cereals. Their choice will only consider those options that satisfy that specific want – cereals high in sugar will not be considered. An individual may work for a cereal company, and as such they wish to support that company by purchasing their cereal. Hence, their cereal market would consist solely of cereals produced by that specific company.

However, we may, for illustrative purposes, define a world where consumer preferences are identical and static. In such a world, there will continually be homogeneous desires on the side of consumers. Let us assume that all consumers are in the market to buy cereal, and that they all want the most colorful cereal – this is their only criteria. We know that consumers always prefer to pay less than more, hence, the trade-off that exists for the consumer will statically be assumed to be price for color. In this situation, we should see one brand of cereal, the least expensive/most colorful choice becomes the dominant, if not eventually the sole offering on the market. In reality however, we see this is not the case, why is this?

This situation never comes to fruition due to this continual competitive nature of businesses driven by the profit motive.

To return our analogy back to the financial world, why is it that one security offering the best risk-return profile is not the only one offered on the market? Firms, in an attempt to increase their own profits, continually seek to earn a more profitable share of the market. Hence, they continually refine and modify their product to better meet the needs of consumers. In our example, even though consumer needs are static, producers will continually differentiate themselves *within the demands of consumers* to earn a higher market share. Given our consumers are only concerned with color and price, the producers will continually change their product to make it the most superior option on the market, *in light of the preferences of consumers*.

What does breakfast cereal have to do with financial securities? Securities are valued according to their underlying assets, they do not exist wholly separate of the physical world. As Mises (1949: 270) tells us, consumer sovereignty rules supreme in the market. Consumers do not passively accept the goods offered to them by producers. In distinction, producers continually strive to offer consumers the goods that they demand, and hence value, the most. Nobody would ever assume that consumers assert a single, time-invariant choice in selecting goods. Our example of all consumers demanding the most colorful cereal all of the time would be deemed ridiculous. However, for over fifty years, the financial realm has assumed that consumers of securities only demand a single choice in purchasing securities: risk, and its corollary, return.

Purchasers of securities make use these issues as means to an end. The end is always specific. They may like the management of a company, support the product they produce, be forced to invest in them through an investment plan, or other numerous, countless options. Risk is certainly an option that is considered, but it is not the sole option. Likewise, producers (issuers of securities) are continually searching for methods to get consumers to choose them over their rivals. More information is provided in the annual reports, different voting rights are offered with security purchases, shares may be convertible, or any number of differentiating methods are used to continually set one company's shares apart from another's. The profit motive over time drives companies to continually strive towards making their offering the most attractive.

The choice of a risk-return trade-off is purely arbitrary. Why not instead make the trade-off return-name, or return-number of employees? The conclusion that actors are solely monetary profit maximizers is found wanting. Instead, actors have a multitude of personal, subjective ends conditioned by their individual desires. They are not passive takers in the market place, but instead demand what will be offered by the producers. As these demands change and shift over time, the values placed on them will also shift. Furthermore, producers continually strive to differentiate their products from one another. This implies that the supply side of the equation is continually being altered to satisfy the wants of consumers better, and hence, increase their profits. The trade-off facing consumers in the financial realm certainly could be risk and return, however, there is no way to state that this is the sole trade-off that exists. Furthermore, even if it were found that this trade-off existed at a given point in time, in a dynamic sense this choice will be subject to change.

The Non-Sequitur of Arbitrage

When we view choice as existing within a multitude of distinct criteria, we see that arbitrage as a pricing concept loses credibility. The application of arbitrage lies on several assumptions that prove erroneous in our dynamic world. The first is that all consumer preferences, and hence, the trade-off inherent in their choice are identical and time-invariant. We have previously disproved this notion. Second, it assumes away the existence of subjective valuations. As only a common, objective criteria could be compared between individuals, there can be no room for subjective valuations occurring simultaneously with arbitrage. Lastly, future expectations must become homogeneous between actors for this concept to be fully employed. The existence of time denies this final possibility from ever occurring.

The concept of arbitrage exists on a foundation of shared decision criteria. Each actor must share the belief that one trade-off, or group of trade-offs, will be the sole selection criteria of a good. We have just dealt with this concept accordingly. However, it will prove instructive to employ an example assuming this fallacy to reign supreme again. Arbitrage assumes that values for identical goods will tend to coincide. If they did not, profit opportunities would exist for entrepreneurs to exploit. An overriding assumption is that there is an objective value that can be compared for individuals. This creates drastic problems for the thinking economist.[132]

Value is always subjectively derived as based on the individual actor. Furthermore, means are always valued according to the end goods (first-order goods) that they will serve to directly satisfy. Individuals will never have the same final desires. Their individual requirements could never lend themselves to being homogeneously given. For example, we have seen that not all individuals may have a need for breakfast cereal, or that not all have the same definition of breakfast cereal. The value

[132] Patel and Zeckhauser (1990) argue that arbitraging tendencies exist for goods between U.S. cosmopolitan areas, but no similar tendency is apparent between OECD countries, *even closely related ones such as the Netherlands and Germany.* They conclude there are more criteria for price determination than are currently thought, giving rise to a poor understanding of equilibrating tendencies in market prices.

that we place on these final goods is never objectively determined by some formula. Instead, it is the result of our subjective minds, ever keeping in mind the relative use these goods will provide in serving our ends compared to the other options that may exist at any given time. This subjective valuing element precludes the possibility of comparing values, *despite the unlikely case where the decision criteria are identical*.

How would arbitrage be achieved in this case? We may know the value placed on a good at one location, and we may know what its value is at another location. However, these two goods are not the same good. A good becomes a final good (good of first-order) at the point where it is available to directly satisfy a want. A breakfast cereal in a store in Cleveland is not the same good as the *exact same* breakfast cereal in a store in Cincinnati.

Some may argue that there is the general tendency for arbitrage to exist however. If cereal in Cleveland were relatively cheaper than in Cincinnati, and even after including all associated costs of making the good the same (transporting it to the same location), this action would be achieved for a profit. We see that this assertion rests on the assumption that the values of the goods will be the same in both places. But what if the subjective valuation of the good (the same good) in the final location is not high enough to warrant moving it – the expected profit will not compensate for the costs involved. In this situation, no arbitrage will result.

Arbitrage relies on the existence of two identical goods. However, as we saw previously, financial assets are not identical. They are conditioned by underlying variables of the businesses they represent, as well as the factors employed to entice consumers to purchase them. As Stigler (1961: 214) points out, "there is never absolute homogeneity" in goods as producers will continually seek to differentiate their offerings. In his own example of a car dealership, identical cars would cease to be identical goods as dealer added benefits, like after-sales support for example, could differ between suppliers. As Rothbard (2004: 282) reminds us, "two goods cannot be "perfect substitutes" for each other, since if consumers regarded two goods as completely identical, they would, by definition, be one good." We know explicitly that financial assets are not the same good, therefore, we can see the problem that arises between trying to arbitrage their differing values by treating them as the same underlying asset.

Lastly, we see that future expectations serve a large role in pricing assets. The future returns, or risk levels, for example, both serve to condition the value that an individual will place on these assets. The existence of inter-temporal arbitrage also assumes that individuals share homogeneous expectations as to the future conditions, and the effects on the value of the asset they are purchasing. Future expectations as to returns, risk, or any other metric may never be identical between actors. Expectations must always be opposite for trade to occur. This is so due to the uncertainty of the future. Each will assign their own subjective probability to each individual criteria as to how likely it will be to occur. Then these criteria will be priced into the asset accordingly in consideration of the future state of affairs. The result will be that no two individuals will share an identical idea of what the valuation will be in the future of any given asset. Given this, we can see that any concept of inter-temporally arbitraging an asset is fraught with danger.

CAPM assumes that all assets with similar risk-return profiles will be *valued* identically on the market. This is erroneous, and rests on a bed of faulty assumptions. The first, already looked at, is that risk and return are the sole metrics used to value an asset. Second is that a comparable objective value can be achieved. This too is found lacking. The strongest criticism is, however, that arbitrage requires the existence of identical assets, separated by a temporal, or physical, element. We see however that securities, although similar, are not identical. Further, competition will continually create differentiating factors to separate and make heterogeneous entities, whose values will likely be heterogeneously determined. Financial assets are viewed as being a case of perfect competition. The neo-classical viewpoint of this is that wholly homogeneous products are offered to consumers who do not value differentiating factors. In reality, we see that the financial arena is a model of the Austrian definition of perfect competition. Wholly heterogeneous products, constantly being further differentiated in order to entice buyers to value them more highly. As products are continually being differentiated, the possibility of arbitrage is erased.

The Market Portfolio and Roll's (1977) Critique

By far the most devastating critique of CAPM thus far was given in Roll (1977) known now, affectionately, as Roll's critique.[133] The critique has been widely cited, and discussed in the finance literature. Hence, only a few remarks will be offered here.

Part of Roll's critique centered on the fact that the market portfolio in CAPM is fundamentally unknowable. The market portfolio must necessarily include every asset available in the economy. This

[133] Two critiques were actually provided by Roll, but only one seemed to grasp the attention of the profession. The 'forgotten' critique centers on the fact that the CAPM is a tautology. Hence, if a market return is assumed to be mean-variance efficient, a portfolio will be by definition also. See Roll (1977) for the formal proof.

could include real estate, stocks, bonds, precious metals, rare art, collections, or the life. Many of these assets will have no market price available, or readily available to be observed. Not every asset will be held on the market for sale, hence precluding the possibility that it will have return to be discovered, or derived for it.

The fact that all the investment opportunities cannot be observed in the economy implies that the CAPM is not testable in any real manner.[134] To mitigate this issue, proxies are given for the market portfolio return and risk based on large indexes, the Dow Jones or S & P 500 for example. The use of these may closely approximate a real market portfolio return, but can never emulate it fully.

Conclusion

We have examined the nature of choice as CAPM assumes it to occur. Every part of has been found wanting. Choice is not a binary decision, instead we have seen that it incorporates a multitude of variables. Consumers demand many criteria when they purchase a good. Furthermore, these demands are not time-invariant. We followed the cycle of a breakfast cereal purchaser to see how their decision-making process alters over time.

Additionally, we see that producers do not passively offer products to consumers. Instead, they continuously strive to create better offerings that differentiate their products from others on the market. The result is that the value that consumers view a product containing is forever in a state of flux. Choice cannot be based solely on static criteria that individuals will have, but will also be based on choices that producers offer that they didn't realize to exist prior.

At the same time, it is assumed that arbitrage will equate every security with a similar risk profile to have the same expected return. Three significant issues arise with this idea. The first, the issue of the binary choice given by risk-return, has already been criticized.[135] We now focus our attention on the idea that a comparable, objective value exists to be arbitraged. We find that this lies on

---

[134] Although, this caveat is ignored by some. Ross (1978: 885) paradoxically notes that "the attractiveness of the CAPM is due to its potential testability."

[135] As Alonso Neira (2004: 40) explains the use of different criteria for choice:

> Los agentes utilizan dos tipos de información: una es costosa y se basa en el uso de los modelos estructurales de determinación del tipo de cambio (análisis fundamental), y otra se basa en la extrapolación de tendencias pasadas (análisis chartista). En este ultimo caso, dado que la información sobre los precios pasados es de dominio publico, puede utilizarse libremente (sin coste).

We see then that not all individuals utilize the same information, even in a general or explicit sense when analyzing values, hence, the criteria they base their choices upon must differ as well.

the assumption that an identical good exists. However, as we have seen, producers are continually searching for ways to differentiate their offerings. The idea that all financial products of the same type are identical is found lacking. Additionally, arbitrage on risk-return profiles involves an inter-temporal element. However, the idea that future expectations of value could be equated is also not positive. There exists a great deal of subjectivity when viewing returns, and the option to compare inter-temporal values may not exist.

The existence of multiple betas, such as is utilized in arbitrage pricing models, does little to rectify the situation. The reason is that, as we have seen, there are not general criteria individuals use in their decision-making process. Instead, each individual has personal determinants they search for, and subjectively assess, when pursuing an asset. Furthermore, as these will be time-variant, we see that even if the individual determinants could be singled out in a static setting (and multiple betas utilized), the dynamic world of human action precludes this option from becoming a reality. As Rothbard (2004: 453) put it, "[a]s long as the factors are all purely specific, economic analysis can say little more about the determinants of their pricing."

Nor does the use of variable betas (as in Krueger and Rahbar 1995) rectify the issue. Models incorporating this element of variance in beta still must rely on a relative stability between beta and the exogenous variables explaining beta (Basse 2006: 16). As these exogenous variables cannot be constant, and cannot be modeled in the manner with which human actors alter the criteria they deem important for 'beta', we see that these approaches also fail.

Lastly, following Roll (1977), we looked at the hypothetical market return. As the discovery of a mean-efficient variance portfolio rests on the idea of knowing the market portfolio's risk-return profile, we can see that a lack of this possibility has severe consequences for both testing the CAPM empirically, as well as using it in reality.

CAPM assumes that a faulty trade-off exists for actors. Furthermore, it assumes away individual choice, and leaves the consumer as a price-taker in a market where their specific demands are not satisfied.[136] At the same time it considers producers to be passive price-takers, unable to differentiate their products and hence, extract added value from them compared to competitors products.

Hayek (1952: 31) reckoned that "every important advance in economic theory during the last hundred years was a further step in the consistent application of subjectivism." CAPM can thus be seen

[136] Arrow (1959: 43) would understand, and outline, the implications of this before the CAPM was developed, "[e]ach individual participant in the economy is supposed to take prices as given and determine his choices as to purchases and sales accordingly; there is no one left over time whose job it is to make a decision on price."

as a step-backward for the science. By replacing the subjective valuations of the market actors with concrete objective criteria, such as beta and variance, we see that an important element has been removed from the pricing process. The implications are that the CAPM rests in a world were actors are assumed to have none of the qualities that make us uniquely human. Instead, actors are replaced by unthinking automatons, unable to consciously seek, or create, value in any way other than defined by the models own assumptions.

2. The Risk-Free Asset

Tobin (1958), in his separation theorem, would note that portfolios should be divided into a risky part, and a risk-free part. As CAPM was later developed, it was seen that the trade-off in a portfolio was between holding 100% of the assets in a risk-free asset, and dividing proportionally between the this and the risky portion. As risk in the CAPM world is variance of return, a risk-free asset is deemed to be variance free. Typically, short-term government debt is used as the proxy for the risk-free rate. Additionally, it is common to match durations of risky and risk-free assets. This eliminates the risk involved in the inherent variability that a risky asset would incur being sold prior to maturity.

There are several problems with this conception of the risk-free asset. The first is the idea that a short-term government note can be considered risk-free. There is always an element of risk inherent in every passage of time, it is inescapable. There is no way to fully eliminate risk, even by assuming that a government bond will always be paid in full.

Variance of returns is the crux of the issue CAPM focuses on. The assumption that a risk-free asset experiences zero variability of return is wanting, even in a nominal sense. Nominal returns, even if the asset was absolutely assured of bankruptcy protection could not be variance free. The essential factor for a risk-free asset is the protection of risk of bankruptcy, or in other words, the guarantee of repayment. This repayment must necessarily not only apply to the principle amount, but the expected yield from the interest-rate. In a nominal sense we see that *ex post* yield can only be the same as the *ex ante* expectation if the actual holding periods are equivalent. In the real world of uncertainty, this is a dangerous assumption. Perfect foresight is not given to humans, and as a result, the holding period, although being broadly defined as an expectation, can by no guarantee be the same *ex post*.

In a real world, we see that returns are significantly affected by purchasing power considerations. Some practitioners have countered with international CAPMs, or utilize an inflation adjusted bond (a TIPS) as the risk-free asset, that counter for the effects of purchasing power considerations. However, purchasing power is not only a cross-currency occurrence. Goods are always priced in currency units, and the corollary is that currency is always priced in terms of goods. With changes in either side of the equation comes changes in the relative purchasing power over time. Hence, we can see that the real-return on a risk-free asset will vary over time in comparison to goods that can be acquired.

Once we see the inherent variability that even a risk-free asset has, we see that Tobin's distinction between risk-free and risky portions of a portfolio is fictitious; it is merely a theoretical construct with not basis in reality. All assets are risky, it is only a matter of degree. The assumption that we can ever guarantee a return can never exist in the dynamic sense. This essential part of the CAPM, the actual crux of risk reduction, is therefore seen as non-operative. Risk in the dynamic world cannot be eliminated.

Time and Risk

Tobin (1958) made the distinction between portfolios holding risky assets and risk-free assets. Prior to this, the want for liquidity had given individuals the desire to hold cash as a risk-free asset. Tobin asked the question of why anyone would hold a non-interest bearing asset, such as cash, in their portfolio instead of an equivalent with a yield.. Hence, a risk-free government bond was viewed as being of equivalent risk to cash, but entailing a return to be earned at the same time. We have previously looked at the failure of CAPM to correctly identify the risk inherent in investment returns. However, even if we assume that the definition of risk used – that risk consists of statistically knowable unknowns – we can see that additional problems surface.

The idea of risk as being best defined as variance of returns may be erroneous however. When we look at a concept of risk, we can not view it as an absolute term. In fact, the word risk, used alone, is meaningless. For the acting human, they can only be concerned about risk toward a given end. If we view the end of an investment as earning as high a return as possible, a second question must be answered – what is the time horizon that the investor is expecting to hold the asset for? As was seen previously, time is an element that influences every decision humans make. Action is never undertaken without conception of time. Instead, there is always an, at least implicit, recognition of the temporal duration the actor expects to be active during. Hence, we see that the investor does not purchase an asset for an unknown amount of time – implicitly they must realize that the asset will not be held for an infinite amount of time.

In fact, the exactitude of expected duration is much more fundamental in the decision making process than this. Each actor realizes that they have a finite time with which to operate towards their ends. Let us assume the end is the maximization of returns by retirement. The actor may not know with any exactness at a young age what this retirement age will be. They know it to be finite, and they have expectations as to when it will occur. We can state that perhaps most people expect to retire between 55 to 70 years of age. Hence, the maximization of returns will not be over an infinite time-span, but instead will be conditioned by this expected amount of time.

Risk does not exist outside of time. It can only be viewed as a product of time. If we lived in a world of no time, we may be inclined to consider this existence in an absence of risk. Commonly risk is referred to as the danger of loss, or the probability of such loss.[137] Two notes can be made about this definition.

The first is that the concept of probability is inseparable from the concept of time. Probability is always, implicitly at least, the probability that something will occur over a finite amount of time. Suppose a person expected to hold an asset for eternity, and also suppose that they had decided that there was a 0.5% chance they would loss all their money in the investment. The 0.5% chance of losing money is meaningless in this concept. If they plan on holding an asset forever, and there is *any* chance of losing money, we see that there is, in fact, a 100% chance of losing money. The element of time will assure it.

The second is that only the danger of loss is important. Nothing is said about the variance of a return. As we will see next, this definition of risk is found wanting.

Variance and Risk

The nature of risk, for asset pricing purposes, is the risk of losing money. It becomes clear that losing money involves the temporal element inherent when the asset is sold, respective of when it was bought. We find that three separate periods of time are important. The first is the moment in time when the asset is purchased. This is the moment when an individual has fated themselves to accept the uncertainty of the future. As time represents an irreversible flow, there is no way after this point to erase the action that has occurred. The second is the passing duration of time that the asset is held for. It is over this period that an individual can watch variations in value occur. However, we see that any variation during this period only exists as a mental construct. The actual return that the individual is trying to maximize is, during the whole of this period, identical: zero. That is to say, no return is earned on an asset until the third temporal element enters the picture. This is the moment when the individual decides to liquidate their position. Hence, it is at this finite point when return can finally become physically calculated with certainty.

[137] See for example, *Collins Essential English Dictionary, 2nd Edition 2006*, where the first listed definition of risk is "the possibility of bringing about misfortune or loss."

Risk, in this sense, can only be calculated with certainty *ex post*. But what is the true nature of risk that we speak of. Two definitions come to mind.

The first is the risk that the actual return will be zero. The possibility that the asset invested in goes bankrupt, is destroyed, or suffers some similar fate will result in a total loss of the investment. This total loss would be the least desirable outcome for the investor. The possibility of this type of risk can be subjectively assessed prior to the actual outcome. We see however that whether an actor continues holding their asset in light of their *ex ante* subjective risk probability will depend on the relative return they expect to earn in light of continuing to hold the asset. It is at the point where the investor believes, without doubt, that the asset is doomed to yield in the future less than they can earn now by liquidating that they will end their position. Hence, risk can be seen as a return less than what could be yielded in other investments – this is a concept of risk relative to other assets.[138]

The second definition we can look at is the risk that the future return will be less than what was expected at the point the asset was purchased. Hence, when an asset is purchased at the initial time, it is done so because the actor expects *ex ante* to receive a higher return on that particular asset than they will on any other. Return in this case cannot be gauged solely in monetary terms however. It will include all incomes expected to be received – monetary and psychic. Therefore, we see that a risk is constantly inherent in the unknowable flux of time that the *ex post* return will not be as high as the initial expectation held.

We see however that both concepts of risk apply only in an *ex post* sense; why is this? It is due to the fact that this definition of risk is related to return, which can only be determined after the moment in time where the sale of the asset has occurred. In the holding duration we can see, however, that the value of this asset will fluctuate. In fact, as values are constantly in a state of flux, we will see that the value that the investor deems they will receive at the time they liquidate the position will also be constantly shifting. However, this fluctuation of value will not entail a strict element of risk for the investor. Instead, it may condition when the investor expects to finally liquidate their position, but until such a time passes, no risk will occur in the sense of disappointment of loss.

Hence, although we have already discussed why risk is not a good measure in the realm of human action, we see that even if it was a good measure, in this instance it is wholly meaningless. The

[138] However, as Mises (1949: 810) points out, entrepreneurs do not invest in the projects that loss them the least, but the projects they expect to earn the most. Given this insight, it is difficult whether to say expected loss is a pricing factor at all for decision-makers directly. To the extent the prospect of a loss creates a lower demand for investment in this area, the value would reflect this. However, an individual will not invest in a decision, expect a probability of loss, and demand a higher rate of return as a result.

variance of return does not affect the return that one earns on an asset. Instead, the element of time that dictates when and hence under what conditions, is the factor that influences the true risk in holding an asset. The passage of time, for the human actor, contributes the element of uncertainty. How "risky" an asset is, is fundamentally given through the temporal element. It is only at the moment that an asset is liquidated, and action completed, that risk can be assessed. The variance in return in the meantime may not contribute to this risk in any way.

Let us take an example of a store that sells fur coats in Norway. Let us assume that the operating costs for the store are approximately constant throughout the year. However, we notice that all the sales the store incurs practically occur over the course of three months – December through February. It becomes apparent that the profit of the store will likewise be very erratic. Large profit margins will exist for these three months, with the remaining nine suffering large losses as costs are not compensated for. What is the risk in owning this asset?

We know that risk is relative. If the owner expects that the profit margin will behave in this way, there will be no significant issue with this fluctuation of returns. However, what if the owner is wholly unprepared for this variance in profit, and hence, experiences great surprise when they realize how returns will be distributed throughout the year. This will be conditioned by the amount of time that the asset owner expects to hold on to the asset for. If we assume that, for the sake of simplicity, the store (asset) can be sold for a high price when sales are brisk, but that there is a lack of buyers when the sales are slack, we will see that the true risk that the owner will suffer a relative loss will only occur when the timing of their sale is not conducive to earning the expected return on the asset. If the owner expects they will sell the store when sales are brisk and the price high, there will be no significant risk in holding the store through a period when sales are minimal and the price low; *the variance in return will not significantly affect the true nature of risk facing the owner.*

Hence, we find that variance is not a good measure of risk for an asset.[139] What matters to the asset owner's risk profile is only two facts: the price when the asset was purchased and the price when it is sold. Variance of return in the duration between when an asset is purchased, and when it is sold, may affect the time when an individual will sell an asset, and hence, complete the initially embarked upon action. Even if risk was an applicable measure for the realm of human action, it is still

[139] Previously we looked at an alternative risk definition, as established by Rothschild and Stiglitz (1970). The inherent problems that their three new risk definitions created in a truly uncertain world have also been looked at. Paradoxically however, we note that Gehr (1979) found that increasing risk, when viewed in the Rothschild-Stiglitz sense, may not necessarily lead to an increasing amount of required return in the aggregate economy, and may in fact lead it decrease in aggregate. These kinds of paradoxes have plagued the risk-return trade-off assumption empirically for since its inception.

inapplicable to be used describing investment returns.

The True Risk of Risk-Free Assets

The assumption that there exists an asset with zero-risk (risk-free) that can be incorporated into a decision is also erroneous. As we looked at in the earlier section, risk is the possibility that there will be either no return on the investment, or that the return will be less than the expectation when the action is initially undertaken. A closer look into the true nature of risk must be undertaken.

Risk in the realm of acting humans is actually uncertainty. There exists a Knightian "fog" in the future which includes eventualities that are unknown in the present – unknowns that are unknown. It follows that any decision that incorporates an inter-temporal element must also recognize that there is forever an element of uncertainty inherent in every action (even if the entrepreneurial function can mitigate this uncertainty). Eventualities will forever appear that could not be foreseen in advance.

The element of risk, that is to say, that a return will be either zero or less than expected will forever remain. Commonly, high-grade government debt is used as the risk-free asset in portfolios prices with CAPM. The assumption is that these are the least "risky" assets available, and hence, exhibit the least variance of returns. Ignoring the previous section and its refutation of the theory that variance of returns is a good measure of risk, let us pose the question of whether government debt can be seen as risk-free from this variance standpoint.

Commonly, the risk-free asset is chosen with a duration approximately equivalent to the expected holding period of a portfolio. This is so because we see that debt can only be variance free if held to maturity. That is to say, we may in the present purchase a bond yield of 5% over one year, however, we will only assuredly earn this return if we hold the bond for the full duration – one year. Assume we sell before this one year has passed. We see that the return *may* be significantly different. This is due to the conditions that surround the sale of the instrument at its early date, conditioned by the relative prices of other asset's that are available. Return of a "risk-free" asset is only variance free if held to maturity.

However, we see that due to the uncertainty of the future, it is unlikely that an asset will be held for the *ex ante* expected time. Eventualities will arise due to the future fog that require a shift in the individual's time preference schedule. An asset may have to be liquidated at a time different than what was previously expected.

Second, this assumes that there is a 100% chance that the asset will be fully available to be liquidated at the future time, for the expected amount. We can see that this can never be assuredly the case. Even the highest grade debt faces uncertainty as to its future conditions. As time changes the business conditions, there will always be a possibility that the asset may not be able to be redeemed at its expected, or full, amount. The probability of this occurring may be viewed as finite. There is, after all, a chance between 0 and 100% that an asset will not return its *ex ante* projected amount. However, we see that this is not possible to reduce the problem to one of mere statistics. This is so as the underlying conditions that will affect the change in return are uncertain in nature, and hence, defy statistical measures.

All assets have an element of risk, to make the distinction between a class with none, and one with, is erroneous. A risk-free asset, even if we consider it as being possible of having risk inherent in it, is impossible to be reduced to a statistical measure predetermining the possibility of risk occurring. This is due to the fact that the underlying nature of risk is really uncertainty – an element that precludes the use of statistical measurement. A risk-free asset, as determined by lack of variance in returns, will never occur absolutely in an *ex post* sense. This is because the concept is intimately tied to the idea of time, specifically, the time at which it is expected that an asset will be sold in the future. To the extent that it can never be known with absolute certainty that an event will occur in the future, we see that variance of returns of even the most risk-free assets will be possible – indeed, they are likely.

Purchasing Power and Risk

Two additional significant issues arise regarding the nature of risk in an asset. We live our day to day lives in terms of nominal prices. If we preclude the element of time completely, we can see that we earn a return at a point in time, and spend it at the same point in time. There will be no difference in the purchasing power of our money in this static sense. However, the temporal world we live in, with its inherent constant flux of changing preferences, implies that changes in purchasing power will continually occur. Additionally, they may occur in uncertain ways. The idea of purchasing power considerations raises significant difficulties and complications for both domestic as well as international users of the CAPM.[140]

[140] Two of Böhm-Bawerk's (1959, vol. II: 238) six determinants of price were related to purchasing power considerations: the magnitudes of value that both the buyer and seller place on the medium of exchange. Of course the other four criteria were supply, demand, and the subjective valuation of the commodity for both the buyer and the seller.

Suppose that a portfolio is being reckoned by the CAPM. The actor who will be purchasing the portfolio, as well as every component within it, can be assumed to be all of the same pool of currency users. That is to say, we see no exchange-rate risk, or other purchasing power complications. However, we see that purchasing power concerns not only an international actor operating with a foreign currency. Purchasing power also affects domestic actors through the continual change in the spread between asset prices.

Suppose that an asset is assumed risk-free, and yields a 5% return per annum. This 5% yield reveals nothing to us, it can only become meaningful when used for a given end. Money is not held solely as money, instead, money is held as a shield against the uncertainty of the future (Mises 1949: 418).[141] Hence, we can see that this uncertainty involves ends that may, or may not be known to the actor in the present time. We also know that values are constantly changing, therefore the spread between values must also be in a continual state of change as well.

This implies a significant amount of real purchasing power risk for any individual regarding the future. Even if an asset were held for exactly the expected duration, and there was zero-chance that it would return the expected yield, this yield could never be forecast with certainty in real terms, *especially* not in real terms concerning a particular end. As preferences change, the valuations between these preferences are also in a state of change. The return that is yielded on an asset will always have an expected use, but the value spread between the expected yield and the purpose the money is saved for has an element of continual shift and uncertainty embedded in it.[142]

Purchasing power is primarily affected through supply side issuance of additional money. As a greater supply of money chases a set amount of goods, spreads will develop between relative values. Even using Friedman's “helicopter experiment,” we see that spreads would develop through the issuance of money. This is due to the fact that individuals' preferences, and hence, the marginal utility of money, would be set into motion. As some ends would be monetarily sought at the relative expense of others, valuations would be effected in a non-linear way. More pointedly however, we see that money would be spread throughout its users and a Cantillon effect would develop. Users first endowed

[141] Samuelson (1947: 123) reckoned that money was held due to uncertainty and transaction costs. Hence, if there was not transaction costs involved, any asset could be used as a money substitute. Although he was correct about the existence of money being due to an uncertain world, we see that in a certain world nothing would be money, instead of his viewpoint where everything would be money. Leontief (1947: 238) would criticize this viewpoint originally by pointing out that demand for money must necessarily be zero in equilibrium, hence, its non-existence in an equilibrium marked by total certainty of the future.

[142] This is demonstrated by Garrison (2001: 62) as he points out his concept of “saving up for something.” Savings, or investments have no meaning independent of ends. It is only with these expected ends in mind that we can assess valuational changes.

with the freshly created money would spend it on the ends they deem more attractive, and hence these would increase in value relative to the ends that the new users of money decline to spend on.[143]

Purchasing power is, however, significantly affected through demand side issues as well. Even if no new quantity of money is introduced to the economy, we see that demand side issues also create valuational shifts. Time-preference changes alter the spread between money saved and spent throughout the economy (Rothbard 2004: 444). As preferences are altered, as well as savings rates, we see that valuations continually change in a dynamic setting. The continual shift between money saved and circulated for spending purposes will continually shift the price dynamics of goods.

This problem is minimized when placed in a solely domestic setting. More realistically, in a developed economy relying on global relationships, is that trade and hence, asset purchases and sales, will be made in foreign currencies. The problem then arises that not only is there purchasing power risk in a real sense, but there is also the risk of exchange-rate fluctuations. However, even as some have argued that fixed-rate currencies eradicate this issue, we see that it can never be removed completely.

As a result, there can be no constancy in prices. Future returns suffer from a significant real-risk to the degree that the ends that the return will be used for may change in value over the passage of time. The longer duration the asset is held for, the greater will be the uncertainty attributed to this temporal element. We find that even if all the previously dismissed unrealistic assumptions regarding the risk-free asset held true – that it offered 100% security in return, that expected duration equated actual duration, that variance was non-existent, or that risk was a good metric to be used in the model – a degree of purchasing power risk would still exist regarding all assets. As this element of risk is unable to be eliminated from any asset involving an inter-temporal element between its point of purchase and its point of sale, we see that no truly risk-free asset can exist *in real terms*.[144]

A Brief Note on Liquidity and Variance

The issue of variance involves the ability to sell an asset at a price that is cohesive with an actor's initial

[143] Bagus (2007) has brought light to the quality of money considerations that are commonly dismissed in the recent literature. It is not only the quantities supplied and demanded of money that affect its purchasing power, but the quality of the money itself that effects this process.

[144] Sechrest (2006: 36) forwards the idea that as credit expansion decreases the market-rate of interest, actor's perception of risk decreases. Hence, a divergence occurs between the actual, and the perceived capital market line (CML). Actors demand a higher return to compensate for their reduced risk expectation, that is to say, returns are higher at every level of risk. As credit expansion *may* increase returns *nominally* on assets, thus giving the perception of higher returns available at all risk levels – *especially on longer duration assets*. Sechrest's conclusion results in a bust as the market incorporates more risk than the real-return is able to bear.

expectations. One of the primary determinants is attributable to liquidity issues. Liquidity represents how little the spread is between the asset wished to be sold, and the price that will be received in exchange for it. Highly liquid assets will experience little variance in their prices, relative to those assets that experience limited liquidity.

Liquidity as a measure between the bid-ask spread on an asset implies that assets selling at prices near both expectations are highly liquid. In fact, we see the only issue that a lack of liquidity brings is that the bid-ask spread is increased. The corollary is that assets with high liquidity will enjoy prices near their asking and offering prices. We have noted that the risk that exists is only due to the difference between the expectation of what an asset's return will be *ex ante*, and what the real return is *ex post*. If the spread between this expectation is narrow, we can see that risk will correspondingly be small – *or mitigated*. Hence, assets in highly liquid markets may offer an opportunity for lower risk through the elimination of wide bid-ask spreads.

However, as we have already looked at, this spread cannot be observed before it occurs; it must be based upon expectations. As these expectations will be based upon an unknowable, foggy and uncertain future, they can never be determined in advance with any degree of certainty. Therefore, although highly liquid assets may be viewed as being less risky after an action is complete, at the onset of the action this is an impossible statement to make.

Conclusion

We have seen that the inclusion, and concept, of a truly risk-free asset is of fundamental importance for the elimination of risk in the CAPM. Following Tobin (1958) the assumption that a risk-free asset can be found and incorporated into a portfolio has been theoretically unquestioned. Further, his assertion that an interest bearing bond can be substituted for non-interest bearing cash with no additional risk has likewise been accepted without concern.

However, several factors make the existence of a truly risk-free asset questionable. Time is inseparable from the idea of risk. Risk viewed as variance in return only exists in a temporal setting, based upon the expectation that the actor has as to when they will sell their asset and complete their action. However, we see that variance as a measure of risk is not an applicable measure. Instead, it has been shown that the expectation of the time to be spent holding an asset is what conditions the true risk; the chance of an asset yielding a less than expected return. Variation after an asset is purchased is not of immediate consequence to the actor. The true determinant of risk occurs only at the future point in time when an asset is sold, and the true return can be solidified. *Time is the instrumental factor in determining risk*.

Purchasing power considerations have been examined. Even if an asset satisfied all the previously criticized viewpoints, we find that in a world of shifting purchasing powers, real-risk exists in every action. As preferences, and hence, valuations are continually being shifted we find that, in real terms, there will constantly be a state of variation between asset values. Complicating this factor is the use of foreign currencies which introduce an element of foreign exchange-rate risk to the model. Furthermore, we have looked at the role of liquidity in assigning risk. As risk is commonly viewed as a variance of return, we see that highly liquid assets offer lower variation of return. This is due to their reduced bid-ask spread in prices. This spread is, however, observable only in an *ex post* sense. The result is that we can never know in advance the true risk inherent in an asset through price variation.

CAPM relies on the existence of a true risk-free asset. However, short of a purely static world, with constant consumer preferences this cannot be achieved. As we have previously pointed out risk is a poor metric to be used in the model. The true measure of "risk" that is sought is one of uncertainty. However, even if we accept the viewpoint that all future uncertainty is knowable, and hence, reducible to a statistical measure, we find that risk can never truly be eliminated. Additionally, risk as is commonly defined in the model may not be the object of an actor's concern. As Buchanan and Di Pierro (1980: 700) allude to, "[e]ach entrepreneur confronts a unique situation, and he thinks he can see opportunities, can create opportunities, to make profits. He acts quite simply because he thinks he can win." That these "unique" opportunities preclude the use of risk as a measure has already been discussed. That actors even consider this when making their initial decision is debatable. As Mises (1949: 810) already pointed to, the entrepreneur acts towards the decision that is expected to yield the most, not loss the least:

> There is no such thing as a safe investment. If capitalists were to behave in the way the risk fable describes and were to strive after what they consider to be the safest investment, their conduct would render this lone of investment unsafe and they would certainly lose their input A capitalist never chooses that investment in which, according to his understanding of the future, the danger of losing his input is smallest. *He chooses that investment in which he expects to make the highest possible profit* [emphasis added].

Additionally, as risk defined as variance only takes on a meaningful definition regarding at least two separate assets, we find that real variance can never be fully eliminated. Even if risk were measurable in the realm of human action, there is an infinite amount of asset combinations to give an equally infinite amount of risk profiles. As it can never be known to what extent variance will exist *ex ante* we find that this metric as applied to the CAPM leads to erroneous conclusions.

3. A Brief Note on Some Complications with Risk and Return

A problem arises in CAPM even if one assumes the theory to be fundamentally correct – how to calculate the variables. Risk, as measured through beta, must somehow be established through one of two methods. The first would be a forward looking expectation on the part of the investor. This would allow them to exercise their subjective beliefs to achieve a probability expectation as to this measure. Alternatively, and more commonly, the historical basis provides the source of the beta measurement.

We have already looked at some of the methodological concerns with the treatment of using past data to extrapolate the future trend. However, there may be two additional aspects that are ancillary to these cautions.

Further, we see that returns may not be normally distributed, and that this assumption hides the results of entrepreneurial action. As we shall see, the assumption that returns follow a symmetric, or even static, distribution is misplaced, and leads to erroneous conclusions.

Second is the assumption that risk, defined as deviation from the norm, is appropriate. We see that CAPM has been developed knowingly assuming that any deviation is appropriately labeled risk. However, we see that this belief may be misplaced, only a special type of deviations are to be negatively treated – negative deviations.[145]

By having a misplaced view of a 'normal amount of profit', we will see that erroneous conclusions are drawn in the model. Particularly damning is the conclusion that beta is a mean-reverting variable. This treatment downplays the significance of entrepreneurial foresight, and hence, creates results ill-suited for the real world.

Historical Returns

Commonly, the method used to establish beta is historically-based by comparing past returns with the variance of return. The rationale is that these measures alter so slowly over time that their projection based upon past results is acceptable.

Early in Markowitz's career he faced the problem, and the realization, that no individual can know what the probability distribution of a security's return will be. However, as he (1991: 470)

[145] Although, we had previously looked at the possibility of using a measure such as semivariance instead of absolute mean standard deviation as a measure of risk. For reasons already explained, due to the assumption of returns being symmetrical around the mean (or normally distributed), the use of absolute mean deviation is utilized commonly as the description of risk.

recounts:

> [I was] convinced by Leonard J. Savage, one of my great teachers at the University of Chicago, that a rational agent acting under uncertainty would act according to "probability beliefs" where no objective probabilities are known; and that these probability beliefs or "subjective probabilities" *combine exactly* as do objective probabilities. [emphasis added]

This input from one of his professors led to the fateful assumption that there was no difference between a historical probability distribution, or one subjectively derived by the individual. They would lead to the same conclusion.

However, one of the underlying, implicit, assumptions in using historical returns to determine deviations (and hence risk), is that there exists some level of comparable, objective return. However, the idea of some stable, or normal, rate of profit is lost in the dynamic world. Only in the confines of the ERE could such a state be achieved. We find that entrepreneurs are constantly shifting resources from areas where profits are lower, to those where they are higher. This is forever a movement towards a general state of equilibrium, however, this equilibrium is forever out of reach. As consumers continually change their preferences, entrepreneurs are continually shifting resources to meet this need.

We see then that profits in year t are not comparable with profits in year t+1. The reason is that these profits were achieved under different conditions, with different consumer preferences, and a knowledge level which is now a thing of the past. Hence, to say that the return this year was 10% over last year's, and last year's was 5% over the year before that, is not a comparable statement. There is nothing we can say about these distinct profits except that – yes, indeed one is higher than the other – however, we find that they will be caused by different reasons. It is not the existence of the profit or loss that matters; it is the cause thereof that is of prime importance to the acting human, and the economist.

However, even if we assume a static state where the profit rate is stable, and preferences are constant, are profit rates directly comparable under such conditions? We can see this unrealistic world will never exist, but we will assume it does, and look at the risk derivation thereof in the next section.

Standard Deviation Derivation

Hence, from this collection of historical returns, we find that risk is derived. The measure is based upon standard deviation, or results different from the mean. Thus, the, at least implicit, assumption is made that any deviation from the expected value is undesired. However, we can see that if an investor received a return above their expectation, they would not be dissatisfied. More likely, they would be delighted!

This issue has been recognized, however, justification has been offered statistically. By assuming that risk, as measured by standard deviation is symmetrical, it is reckoned that this usage will approximate real risk. However, while the assumption that the standard deviation of returns will be symmetrical across a whole cross-section of securities *may exist*, this will likely be decidedly less so in view of individual securities.

The market process operates on a basis where successful entrepreneurs are rewarded, while unsuccessful ones are eliminated. The result is that there will likely be companies continually offering returns that are either skewed to the positive side over extended periods of time, or skewed negatively for short periods of time. The reason for this result is that entrepreneurs who are successful are rewarded with profits, and hence, not only experience higher returns, *but stay in business*. In contrast, unsuccessful entrepreneurs are penalized with losses, earn below average returns and find themselves out of the market eventually.

Over a cross-section of the whole economy, one may find that the standard deviation of returns is normalized and symmetrical. In fact, this will be likely. As successful ventures will be offset by less successful ones, we see that there will be a representation on both sides of the return scale – negative and positive. However, CAPM is intimately concerned with the concept of the risk level that one individual stock adds to an existing portfolio. We see that this general risk level throughout the economy is not a valid measure to this end.

Furthermore, a generalized stochastic process cannot be applied to expected returns or variances as a projection. As was shown previously, there exists no randomly generated behavior in the realm of human action. All is purposeful, and relies on events that have specific casual relationships to explain their occurrence.

Independence and Risk

CAPM relies on risk-reduction through *independent* sources of risk. True independence may be difficult in the realm of a tightly-knit economy. Cowen (1997: 21) notes that an economy of fifty independent shocks would be safer than an economy whose return is determined by a single shock. However, the independence of the *shocks* is the key here. It is unknown in an *ex ante* sense the true dependence of shocks on returns.

Hence, Cowen notes the erroneous conclusion that an economy will be less risky than a single sector – the source of shocks may be related. In fact, multiplying a single shock several times may not result in a decreased aggregate risk if the source cannot be isolated and treated as independent.

Conclusion

We may find it instructive to remember what Gordon and Hynes (1970: 377) had to say about data and the decision making process:

> [A] formal decision process for this learning is not possible in a world where the underlying stochastic process is not stable. It is true that the response sellers make to new *data* can, ex post, be described as a rational response to subjective prior distributions. However, since there is not sufficient information to accumulate relative frequencies, these subjective estimates will depend, in part at least, on 'judgment', will differ among rational persons confronted with the same measurable *data*, and will also alter from period to period in an unpredictable manner on the basis of information external to the individual's own sampling experience.

It is possible that even subjective expectations of the future will be based upon, at least some, information that is not applicable to the decision making process.[146] As the reason for action is continually in a state of flux, the results of one period may have no bearing on the results of another – many relationships that seem to be casually linked could possibly be only accidentally so. It is only by examining the underlying causes, and not the results that of them, that this can be established.

We have seen that the concept of a normal rate of profit is unsound. No such thing could ever exist short of the ERE. The idea of comparing returns of one period with another to establish a measure of risk is futile. Past changes in returns or profit rates were due to specific factors that occurred at a

[146] Perold (2004: 5) criticizes the inherent subjectiveness involved in forecasting future growth rates in order to use the Gordon -Shapiro growth model. However, he equally fails to realize the subjectiveness in applying past deviations, or returns to a future setting.

point in time, but may fail to continue into the future. More likely, given the ever constant change in individual actors on a process, the reason for a change cannot even be linked to a single, constant actor.[147]

The measure of standard deviation as a risk measure assumes that individuals are averse to returns that deviate from their expectation. However, we see that only negative deviations are to be viewed negatively. If a stock continually returned 1% more than the expectation of the previous period, the investor would no doubt view this as an enjoyable event. Only *some* types of variations are to be penalized, not all.

If indeed a trade-off between risk and return existed for all investors, there should be at least a definitive correct construct for what risk entails. As general variance *may* not be of concern to many investors, we can see that the implication for a general asset pricing model based upon this concept is in trouble. The economist is concerned with the cause of action, not necessarily the result. By viewing risk by a resultant measure – variance of returns – CAPM loses sight of this fact and under-emphasizes the role true, purposeful human action serves in the dynamic world.

[147] Tirole (1985) showed that as new individuals constantly enter the investment arena, time preference alterations are continual, and hence, factors at one point in time may cease to be important in another. In the Lucas (1972) model, new generations receive old money and are unsure if price fluctuations are nominal or real. Younger generations are unable to estimate the new money supply in any reliable way. Cowen (1997: 94) argues whether this extreme view is plausible.

IV. CONCLUSION

Mises completed his opus *Human Action* by surmising that if actors ignored the teachings available through economics, they would "stamp out society and the human race." We should be hesitant to assert the same true in finance. We have seen that finance has never had a solid base from which to advance from. As such, there is no solid base from which to fall from grace. The fact remains however that theoretical finance is forever a realm of economics; it relies on the relationships of the physical market, and valuations are ultimately determined by individual consumers in want of satisfaction.[148]

The two most fundamental ideas underlying most financial literature are that markets are efficient, as is dictated through EMH, and that risk and return are the sole trade-off facing investors, and that this risk is able to be eliminated through return covariance, as is espoused by CAPM. Both these theories are coming under increasingly hostile empirical testing. Academics and practitioners alike are beginning to question their validity, however, they lack the proper theoretical tools to offer a substantiated critique. It is only through sound theory that this can be forwarded, never through empirical testing with historical data.

We have found primarily that these two ideas have been based upon erroneous, unrealistic conclusions. Such fundamental ideas as the role time plays in action, the function of the entrepreneur in the economy, or the real metric of risk applicable to the physical realm are all left unexplained.

In the past, empirical testing of both concepts has found significant flaws in not only their predictive value, but in their explanatory value of the past. The largest failing of these tests lies in the fact that the concepts of EMH and CAPM are continually tested as wholes. Their independent building blocks are never given the attention necessary. The result is that serious questions are raised as to why these concepts prove empirically difficult to prove, but no insight is gleaned as to what it is that is incorrect. The economist can not be concerned with the result of action, but the causes thereof. Menger stressed that everything existed in a state of causality. The best method we can use to discern these causal relationships are by dissecting theories into small blocks, and testing the logic inherent in them. This is what has been, by and large, omitted from past attempts to critique EMH and CAPM, and this missed step has yielded inconclusive results.

We have attempted to shed light on this problem by assessing individual facets of each concept.

[148] We feel the need to point out that while theoretical finance is forever an academic discipline, this cannot be confused with finance, which is "first and foremost a profession" (Calandro 2004: 45). If finance is the ultimate end of finance theory, we see the need to establish a new footing from which to draw correct conclusions from in practice.

The result has been that both EMH and CAPM are unable to tell us anything about the world that humans act within. EMH ignores the importance of heterogeneous information and knowledge, it confuses concepts of efficiency, and misses some of the most important determinants of price. CAPM suffers just as badly. It over-simplifies the choice that faces consumers, and the methods producers use to satisfy these wants. It over-emphasizes the risk-free asset, without examining whether the risk-free asset has a viable existence, or is even an applicable influence on price.

There is no doubt that one of the primary determining factors of these models fallacies, and their continued success, has been a lack of methodological rigor in the realm of finance. Unlike its parent field of economics, finance has never been exposed to a proper methodological base; one based on time-invariant principles and purely deductive logic. Instead, it has continually focused on empirical results to derive theory from. In fact, one of the main points that spurred on the creation of a separate field of finance was the plethora of empirical results from which to derive theory from.

We see however, that this is the wrong approach – in fact, it is backwards. History does not make theory. Theory neither makes history, but it describes it. Only the use of deductively true propositions can be used to base a science of human action on. As finance is most definitely concerned with the actions that individual humans undertake in selling, buying, and valuing securities in the market, we see the dire need for this type of approach.

A disconnect exists between the real world of physical assets, one governed by the laws of economics, and the realm of finance aiming to price these assets. Securities do not exist *in absentia* physical assets – the dichotomy is false. Some think this distinction is warranted, even that it should be expanded and made more apparent. Ross (1987: 34) contends that the divide should be widened, with two separate fields of competing theories:

> I believe that it would be productive to maintain some distance between the two areas [finance and economics]... Clearly, financial theorists should master modern economic theory and look to apply it to problems of interest in finance... But, much of what finance has accomplished and contributed to economics has been a result of working a a somewhat isolated and eccentric tradition.

And while we should see no problem with increased competition in any aspect of our lives, the distinction becomes unwarranted – undesired even. Before it was ever asked if we *can* separate these two worlds, the question of *why* we should was never answered properly. The justification for the dichotomy has always been the plethora of detailed data that exists in the world of finance, the perfect substitutability of assets, frictionless transactions, and the like. But these are only applicable criteria given the point of view the approach one uses. Given the highly empirical approach that has always been a distinction of finance, we can see the impetus that drives the split. Financial assets rely on the physical realm for their very existence; there could be no other order of business. To make the distinction between the two is not only detrimental, it is dangerous.

In a world of human action, both these models are found wanting. The conclusions of the current "laws" of finance are totally distinct from those of economics. This is so as they are not "laws" in the strict sense we wish to ascribe to the word. They are forever contingently true or false hypotheses, awaiting continued empirical validation or refutation. We have just offered a theoretical refutation of both, this eliminates the need for further empirical testing. Mises worried of the state of the world if the laws of economics were to be forgotten, or not adhered to. If the "laws" of finance, based on EMH and CAPM, were ignored and forgotten, the world would be in a better position to create a more sound, and solid, theory for financial assets. As Hayek (1944: 246) noted, "[w]e shall not grow wiser before we learn that much that we have done was very foolish." It is hoped this work is a step in the right direction.

REFERENCES

Alchian, A., and Allen, W. R. 1964. *University Economics*. Belmont, CA: Wadsworth Publishing Co.

Allen, H., and M.P. Taylor. 1990. Charts, Noise and Fundamentals in the London Foreign Exchange Market. *Economic Journal* 100(400), (Supplement): 49-59.

Alonso Neira, Miguel Angel. 2004. *Teoría Economica de las Crisis Monetarias y Financieras y de los Controles de Capital*. Madrid: Instituto de Estudios Económicos.

American Finance Association. 2008. American Finance Association Award for Excellence in Finance – Eugene Fama. *The Journal of Finance* 63(1): v – viii.

Arrow, Kenneth. 1959. Toward a Theory of Price Adjustments, in (*eds*.) Moses Abramowitz et al., *The Allocation of Economic Resources*. Stanford, CA: Stanford University Press.

Arrow, Kenneth. 1974a. Limited Knowledge and Economic Analysis. *The American Economic Review* 64(1): xiii-xiv+1-10.

Arrow, Kenneth. 1974b. *The Limits of Organization*. New York: W. W. Norton.

Arrow, Kenneth. 1978. The Future and the Present in Economic Life. *Economic Inquiry* 16(2): 157-169.

Arrow, Kenneth. 1994. Methodological Individualism and Social Knowledge. *American Economic Review* 84(2): 1-9.

Bachelier, Louis. 1900.Théorie de la Spéculation. *Annales Scientifiques de l'École Normale Supérieure*: 21-86.

Backhouse, Roger E. 1994. *New Directions in Economic Methodology*, (*ed*.) Roger E. Backhouse. New York: Routledge.

Baetjer, Howard Jr. 2000. Capital as Embodied Knowledge: Some Implications for the Theory of Economic Growth. *The Review of Austrian Economics* 13(2): 147-174.

Baetjer, Howard Jr., and Peter Lewin. 2007. *Can Ideas be Capital: Can Capital be Anything Else?* Mercatus Center George Mason University, Working Paper 83.

Bagus, Philipp.2007. *Deflation – Is it Really Harmful? A Theoretical Analysis of Deflation with Two Historical Examples*. Ph. D. dissertation. Universidad Rey Juan Carlos.

Banz, R. 1981. The Relationship Between Return and Market Value of Common Stocks. *Journal of Financial Economics* 9(1): 3-18.

Basse, Tobias. 2006. An Austrian Version of the Lucas Critique. *The Quarterly Journal of Austrian Economics* 9(1): 15-26.

Basu, S. 1983. The Relationship Between Earnings Yield, Market Value and the Return for NYSE Common Stocks. *Journal of Financial Economics* 12(1): 126-156.

Baumol, William J. 1968. Entrepreneurship in Economic Theory. *American Economic Review* 58(2): 64-71.

Beechey, Merideth, Brian Gruen, and James Vickery. 2000. *The Efficient Market Hypothesis: A Survey*. Research Discussion Paper 2000-01, The Economic Research Department Reserve Bank of Australia, January.

Bernstein, Peter L. 1992. *Capital Ideas: The Improbable Origins of Modern Wall Street*. New York, Free Press.

Bernstein, Peter L. 1999. A New Look at the Efficient Market Hypothesis. *The Journal of Portfolio Management* 25(2): 1-2.

Bernoulli, Daniel. 1954. Exposition of a New Theory of the Measurement of Risk, (*trans*.) L Sommer. *Econometrica* 22(1): 23-36.

Bikhchandani, S., D. Hirshleifer and I. Welch. 1992. A Theory of Fads, Fashion, Customs and Cultural Change as Informational Cascades. *Journal of Political Economy* 100(5): 992-1026.

Black, Fischer. 1972. Capital Market Equilibrium with Restricted Borrowing. *The Journal of Business* 45(3): 444-455.

Black, Fischer, Michael C. Jensen, and Myron Scholes. 1972. The Capital Asset Pricing Model: Some Empirical Tests, in *Studies in the Theory of Capital Markets*, (*ed*.) Michael C. Jensen. New York: Praeger Publishing Co.

Blanchard, Olivier. 1979. Speculative Bubbles, Crashes and Rational Expectations. *Economic Letters* 3(4): 387-389.

Bodie, Zvi, Alex Kane and Alan J. Marcus. 2001. *Essentials of Investments, International 4th Edition*. Singapore: McGraw-Hill

Boehm, Stephan. 1994. *Hayek and Knowledge: Some Question Marks,* in *Capitalism, Socialism and Knowledge: The Economics of F. A. Hayek, Vol. II.*, Colonna, M., H. Hageman, and O. F. Hamouda (*eds*.). Aldershot: Edward Elgar.

Boettke, Peter J. 1997. Where Did Economics Go Wrong? Modern Economics and the Flight from Reality. *Critical Review* 11(1): 11-64.

Böhm-Bawerk, Eugen von. 1959. *Capital and Interest*. *Volume II: Positive Theory of Capital*. South Holland, IL: Libertarian Press.

Brunner, R., K. Eades, R. Harris, and R. Higgins. 1998. 'Best Practices' in Estimating the Cost of Capital: Survey and Synthesis. *Financial Practice and Education*, (Spring/Summer).

Buchanan, James M., and A. Di Pierro. 1980. Cognition, Choice, and Entrepreneurship. *Southern Economic Journal* 46(1): 693-701.

Budd, A. P., and Robert H. Litzenberger. 1972. The Market Price of Risk, Size of Market and Investors' Risk Aversion: A Comment. *The Review of Economics and Statistics* 54(2): 204-206.

Buffett, Warren E. 1984. The Superinvestors of Graham-and-Doddsville. *Hermes, the Columbia Business School Magazine*.

Buffett, Warren E. 1996. *Annual Report*. Berkshire Hathaway.

Butos, William N. 2003. Knowledge Questions: Hayek, Keynes and Beyond. *The Review of Austrian Economics* 16(4): 291-307.

Calandro, Joseph Jr. 2004. Reflexivity, Business Cycles, and the New Economy. *The Quarterly Journal of Austrian Economics* 7(3): 45-69.

Casson, Mark. 1982. *The Entrepreneur: An Economic Theory*. Oxford: Martin Robinson.

Chamberlin, E. H. 1957. Some Aspects of Nonprice Competition. Reprinted in *Towards a More General Theory of Value*. London: Oxford University Press.

Chan, L. K. C., Y. Hamao, and R. Lakonishok. 1991. Fundamentals and Stock Returns in Japan. *Journal of Finance* 46(5): 1739-1764.

Clarke, Jonathan, Tomas Jandik, and Gershon Mandelker. 2001. The Efficient Market Hypothesis – An Overview, in R. C. Arffa (*ed.*) *Leading Financial Experts: Investment Strategies from Industry Leaders*. New York: Wiley.

Cochrane, John H. 1999. New Facts in Finance. *Economic Perspectives* 23(3): 59-78.

Cootner, Paul H. 1962. Stock Prices: Random vs. Systematic Changes. *Industrial Management* 3(Spring): 34-105.

Cornell, Bradford, and Richard Roll. 1981. Strategies for Pairwise Competitions in Markets and Organizations. *Bell Journal of Economics* 12(1): 201-213.

Cowen, Tyler. 1997. *Risk and Business Cycles*. New York: Routledge.

Cowen, Tyler, and Richard Fink. 1985. Inconsistent Equilibrium constructs: The Evenly Rotating Economy of Mises and Rothbard. *The American Economic Review* 75(4): 866-869.

Cowles A., and E. H. Jones. 1937. A Posteriori Probabilities in Stock Market Action. *Econometrica*

5(3): 280-294.

de Long, J. B., A. Shleifer, L.H. Summers and R.J. Waldman. 1990. Noise Trader Risk in Financial Markets. *Journal of Political Economy* 98(4): 703-738.

DeBondt, Werner F. M., and Richard H. Thaler. 1987. Further Evidence on Investor Over-reaction and Stock Market Seasonality. *Journal of Finance* 42(3): 557-581.

Demsetz, Harold. 1969. Information and Efficiency: Another Viewpoint. *Journal of Law and Economics* 12(1): 1-22.

Dequech, David. 2000. Asset Choice, Liquidity Preference, and Rationality Under Uncertainty. *Journal of Economic Issues* 34(1): 159-176.

Dornbusch, R. 1976. Expectations and Exchange Rate Dynamics. *Journal of Political Economy* 84(6): 1161-1176.

Dothan, Michael U. 1990. *Prices in Financial Markets*. New York: Oxford University Press.

Douglas, George W. 1969. Risk in Equity Markets: An Empirical Appraisal of Market Efficiency. *Yale Economics Essays* 9(Spring): 3-45.

Dow, James, and Gary Gorton. 1997. Market Efficiency and Economic Efficiency: Is There a Connection? *The Journal of Finance* 52(3), Papers and Proceedings Fifty-seventh Annual Meeting, American Finance Association, New Orleans, Louisiana January 4-6, 1997, (Jul.): 1087-1129.

Elton, Edwin and Martin L. Gruber. 1987. *Modern Portfolio Theory and Investment Analysis*. New York: Wiley.

Fabozzi, Frank J., Sergio M. Focardi., and Petter N. Kolm. 2006. *Financial Modeling of the Equity Market: From CAPM to Cointegration*. Hoboken, NJ: Wiley.

Fama, Eugene. 1965a. *Random Walks in Stock-Market Prices*. Graduate School of Business, Selected Papers, No. 16. Chicago, The University of Chicago.

Fama, Eugene.1965b. The Behavior of Stock Market Prices. *Journal of Business* 38(1): 34-105.

Fama, Eugene. 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance* 25(2): 383-417.

Fama, Eugene. 1971. Risk, Return and Equilibrium. *Journal of Political Economy* 79(1): 444-454.

Fama, Eugene. 1991. Efficient Capital Markets II. *Journal of Finance* 46(5): 1575-1617.

Fama, Eugene F., and Kenneth R. French. 1992. The Cross Section of Expected Stock Returns. *Journal of Finance* 47(2): 427-465.

Fama, Eugene F., and Kenneth R. French. 1993. Common Risk factors in the Returns on Stocks and Bonds. *Journal of Financial Economics* 33(1): 3-56.

Fama, Eugene F., and Kenneth R. French. 2004. A Capital Asset Pricing Model: Theory and Evidence. *The Journal of Economic Perspectives* 18(3): 35-46.

Faust, J., J. Rogers, and J. H. Wright. 2001. *Exchange Rate Forecasting: The Errors We've Really Made*. Prepared for the Conference on Empirical Exchange Rate Models, Sept. 28-29. University of Wisconsin, Madison.

Fisher, Irving. 1906. *The Nature of Capital and Income*. London: Macmillan.

Flood, Robert P., and Peter M. Garber. 1980. Market Fundamentals Versus Price Level Bubbles: The First Tests. *Journal of Political Economy* 88(4): 745-70.

Frankel, J. A., and K. A. Froot. 1990. *Exchange Rate Forecasting Techniques, Survey Data, and Implications for the Foreign Exchange Market*. NBER Working Paper No. 3470, October.

French, K.R. 1980. Stock Returns and the Weekend Effect. *Journal of Financial Economics* 8(1): 55-69.

Friedman, David. 1997. *Hidden Order: The Economics of Everyday Life*. New York: Harper Collins.

Friedman, Milton. 1953. The Case for Flexible Exchange Rates, in *Essays in Positive Economics*. Chicago, University of Chicago Press.

Friedman, Milton. 1953b. The Methodology of Positive Economics, in *Essays in Positive Economics*. Chicago, University of Chicago Press.

Friend, Irwin. 1973. Mythodology in Finance. *The Journal of Finance* 28(2), Papers and Proceedings of the Thirty-First Annual Meeting of the American Finance Association, Toronto, Canada, (December 28-30): 257-272.

Friend, Irwin, and Marshall Blume. 1970. Measurement of Portfolio Performance Under Uncertainty. *American Economic Review* 60(4): 561-575.

Froot, K. A., and M. Obstfeld. 1991. Intrinsic Bubbles: The Case of Stock Prices. *American Economic Review* 81(5): 1189-1214.

Garrison, Roger W. 1984. Time and Money: The Universals of Macroeconomic Theorizing. *Journal of Macroeconomics* 6(2): 197-213.

Garrison, Roger W. 1991. New Classical and Old Austrian Economics: Equilibrium Business Cycle Theory in Perspective. *The Review of Austrian Economics* 5(1): 91-103.

Garrison, Roger W. 2001. *Time and Money: The Macroeconomics of Capital Structure*. London: Routledge.

Gehr, Adam K. Jr. 1979. Risk and Return. *The Journal of Finance* 34(4): 1027-1030.

Gennotte, G. and H. Leland. 1990. Market Liquidity, Hedging and Crashes. *American Economic Review* 80(5): 999-1021.

Girlich, Hans-Jachim. 2001. *Bachelier's Predecessors*. Paper presented at the 2nd World Congress of the Bachelier Finance Society, (June 12-15): Crete.

Godfrey, Michael D., Clive W. J. Granger, and Oskar Morgenstern. 1964. The Random-Walk Hypothesis of Stock Market Behavior. *Kyklos* 17(1): 1-30.

Gonedes, N. J. 1976. Capital Market Equilibrium for a Class of Heterogeneous Expectations in a Two-Parameter World. *Journal of Finance* 31(1): 1-15.

Gordon, D. F., and A. Hynes. 1970. On the Theory of Price Dynamics, in Edmund S. Phelps (*ed.*) *Microeconomics Foundations of Employment and Inflation Theory*. New York: W. W. Norton & Co.

Graham, Benjamin and David L. Dodd. 1934. *Security Analysis: Principles and Technique*. Columbus, OH: McGraw-Hill.

Grossman, Sanford J. 1976. On the Efficiency of Competitive Stock Markets When Traders Have Diverse Information. *Journal of Finance* 31(2): 573-585.

Grossman, Sanford J. 1981. An Introduction to the Theory of Rational Expectations Under Asymmetric Information. *Review of Economic Studies* 48(4): 541-559.

Grossman, Sanford J. and Joseph E. Stiglitz. 1976. Information and Competitive Price Systems. *The American Economic Review* 66(2), Papers and Proceedings of the Eighty-Eighth Annual Meeting of the American Economic Association, (May): 246-253.

Grossman, Sanford J. and Joseph E. Stiglitz. 1980. On the Impossibility of Informationally Efficient Markets. *American Economic Review* 70(3): 393-408.

Gunning, J. Patrick. 1989. Mises and the Evenly Rotating Economy. *The Review of Austrian Economics* 3(1): 123-135.

Harrison, Paul. 1997. A History of an Intellectual Arbitrage: The Evolution of Financial Economics. New Economics and its History, Annual Supplement to the *History of Political Economy* 29: 172-187.

Hayek, F. A. 1936. Economics and Knowledge. Economics and Knowledge. *Economica NS* 4(13): 33-54.

Hayek, F. A. 1937. Investment That Raises the Demand for Capital. *The Review of Economic Statistics* 19(4): 174-177.

Hayek, F. A. [1944] 2006. *The Road to Serfdom*. New York: Routledge.

Hayek, F. A. 1945. The Use of Knowledge in Society. *The American Economic Review* 35(4): 519-530.

Hayek, F. A. 1946. The Meaning of Competition, reprinted in *Individualism and Economic Order*. Chicago: The University of Chicago Press.

Hayek, F. A. 1952. *The Counter-Revolution of Science*. Glencoe, IL: The Free Press.

Hayek, F. A. 1960. *The Constitution of Liberty*. London: Routledge & Kegan Paul Ltd.

Hayek, F. A. 1973. *Law, Legislation and Liberty, Volume I: Rules and Order*. Chicago: The University of Chicago Press.

Hayek, F. A. 1976a. *Law, Legislation and Liberty, Volume II: The Mirage of Social Justice*. Chicago: The University of Chicago Press.

Hayek, F. A. [1952] 1976b. *The Sensory Order*. Chicago: The University of Chicago Press.

Hayek, F. A. 1994. *Hayek on Hayek: An Autobiographical Dialogue*. Stephen Kresge and Leif Wenar, (*eds*.). Chicago: The University of Chicago Press.

Hayek, F. A. 2002. *Competition as a Discovery Procedure*, (*trans*.) Marcellus S. Snow. *The Quarterly Journal of Austrian Economics* 5(3): 9-23.

Hazlitt, Henry. 1974. Review of Competition and Entrepreneurship. *The Freeman* (December): 759.

Hess, Charlotte, and Elinor Ostrom. 2003. Ideas, Artifacts, and Facilities: Information as a Common-Pool Resource. *Law and Contemporary Problems* 66(1/2): 111-145.

Hicks, J. R. 1935. A Suggestion for Simplifying the Theory of Money. *Economica*, *NS* 2(5): 1-19.

Hicks, J. R. 1939. The Foundations of Welfare Economics. *Economic Journal* 49(196): 696-712.

Hicks, J. R. 1962. Liquidity. *The Economic Journal* 72(328): 787-802.

High, Jack. 1982. Alertness and Judgment: Comment on Kirzner. In Israel M. Kirzner (*ed*.) *Method, Process and Austrian Economics*. Lexington, MA: Lexington Books.

Hirshleifer, David. 2001. Investor Psychology and Asset Pricing. *The Journal of Finance* 56(4), Papers and Proceedings of the Sixty-First Annual Meeting of the American Finance Association, New Orleans, Louisiana, (January 5-7): 1533-1597.

Hirshleifer, Jack. 1973. Where Are We in the Theory of Information? *The American Economic Review* 63(2): 31-39.

Hirshleifer, Jack, and John G. Riley. 1995. *The Analytics of Uncertainty and Information*. Cambridge:

Cambridge University Press.

Holcombe, Randall G. 1998. Entrepreneurship and Economic Growth. *The Quarterly Journal of Austrian Economics* 1(2): 45-62.

Holcombe, Randall G. 2003. Progress and Entrepreneurship. *The Quarterly Journal of Austrian Economics* 6(3): 3-26.

Hoppe, Hans-Hermann. 1989. *I*n Defense of Extreme Rationalism: Thoughts on Donald McCloskey's The Rhetoric of Economics. *Review of Austrian Economics* 3(1): 179-214.

Hoppe, Hans-Hermann. 1995. *Economic Science and the Austrian Method*. Auburn, AL: Ludwig von Mises Institute.

Hoppe, Hans-Hermann. 1997. On Certainty and Uncertainty, or: How Rational Can Our Expectations Be? *Review of Austrian Economics* 10(1): 49-78.

Hoppe, Hans-Hermann. 2007. The Limits of Numerical Probability: Frank H. Knight and Ludwig von Mises and the Frequency Interpretation. *The Quarterly Journal of Austrian Economics* 10(1): 1-20.

Huberman, Gur, and Tomer Regev. 2001. Contagious Speculation and a Cure for Cancer: A Nonevent That Made Stock Prices Soar. *The Journal of Finance* 56(1): 387-396.

Huerta de Soto, Jesús. 2004. La Teoría de la Eficiencia Dinámica. *Procesos de Mercado: Revista Europea de Economía Política* 1(1): 11-71.

Huerta de Soto, Jesús. 2005. *Socialismo, Cálculo Económico, y Función Empresarial*. Madrid: Unión Editorial.

Huerta de Soto, Jesús. 2006. *Money, Bank Credit and Economic Cycles*. Auburn,AL: Ludwig von Mises Institute.

Huerta de Soto, Sonsoles. 2005. Eficiencia y Equilibrio en los Mercados de Capitales. *Procesos de Mercado: Revista Europea de Economía Política* 2(2): 211-219.

Hülsmann, Jörg Guido. 1997. Knowledge, Judgment and the Use of Property in Society. *Review of Austrian Economics* 10(1): 23-48.

Hülsmann, Jörg Guido. 1999. Entrepreneurship and Economic Growth: Comment on Holcombe. *The Quarterly Journal of Austrian Economics* 2(2): 63-65.

Hülsmann, Jörg Guido. 2002. A Theory of Interest. *The Quarterly Journal of Austrian Economics* 5(4): 77-110.

Hülsmann, Jörg Guido. 2003. Facts and Counterfactuals in Economic Law. *Journal of Libertarian Studies* 17(1): 57-102.

Jevons, William Stanley. 1871. *The Theory of Political Economy*. London: Macmillan and Co.

Kaldor, Nicholas. 1939a. Speculation and Economic Stability. *Review of Economic Studies* 7(1): 1-27.

Kaldor, N. 1939b. Welfare Propositions in Economics and Interpersonal Comparisons of Utility. *The Economic Journal* 49(145): 549-552.

Kenall M. G. and A. B. Hill. 1953. The Analysis of Economic Time-Series-Part-I: Prices. *Journal of the Royal Statistical Society. Series A* (General) 116(1): 11-25.

Keynes, John M. 1936. *The General Theory of Employment, Interest and Money*. London: Macmillan.

Keynes, John M.. 1939. Memorandum on the Estates Committee, Kings College, Cambridge, May 8. in *The Collected Writings of John Maynard Keynes*, (*ed*.) Donald Moggridge. New York, Cambridge University Press.

Kindleberger, Charles. 1989. *Manias, Panics and Crashes: A History of Financial Crisis*. New York: Basic Books.

Kirzner, Israel M. 1973. *Competition and Entrepreneurship*. Chicago: The University of Chicago Press.

Kirzner, Israel M. 1979. *Perception, Opportunity, and Profit: Studies in the Theory of Entrepreneurship*. Chicago: The University of Chicago Press.

Kirzner, Israel M. 1985. *Discovery and the Capitalist Process*. Chicago: The University of Chicago Press.

Kirzner, Israel M. 1992. *The Meaning of the Market Process*. London: Routledge.

Kirzner, Israel M. 1996. Reflections on the Misesian Legacy in Economics. *The Review of Austrian Economics* 9(2): 143-154.

Kirzner, Israel M. 1997. *How Markets Work: Disequilibrium, Entrepreneurship, and Discovery*. London: Institute of Economic Affairs, Hobart Paper No. 133.

Kirzner, Israel M. 1997a. Entrepreneurial Discovery and the Competitive Market Process: An Austrian Approach. *Journal of Economic Literature* 35: 60-85.

Kirzner, Israel M. 1998. Creativity an/or Alertness: A Reconsideration of the Schumpeterian Entrepreneur. *The Review of Austrian Economics* 11(1/2): 5-17.

Kirzner, Israel M. 2005. Information-Knowledge and Action-Knowledge. *Econ Journal Watch* 2(1): 75-81.

Knight, Frank H. 1921. *Risk, Uncertainty and Profit*. Boston, MA: Houghton Mifflin.

Knight, Frank H. 1954. An Appraisal of Economic Change: Discussion. *American Economic Review*

(Papers and Proceedings).

Korajczyk, Robert A. 1999. *Asset Pricing and Portfolio Performance Models, Strategy and Performance Metrics*. London: Risk Books.

Kortian, Tro. 1995. *Modern Approaches to Asset Price Formation: A Survey of Recent Theoretical Literature*. Research Discussion Paper 9501, International Department and Economic Research Department Reserve Bank of Australia, March.

Kruegar, Thomas M., and Mohammad H. Rahbar. 1995. Explanation of Industry Returns Using the Variable Beta Model and Lagged Variable Beta Model. *Journal of Financial and Strategic Decisions* 8(2): 35-45.

Kurz, Mordecai. 1992. *Asset Prices with Rational Beliefs*. Mimeo, Stanford University.

Lachmann, Ludwig M. 1973. Sir John Hicks as a Neo-Austrian. *South African Journal of Economics* 41(3): 126-132.

Lachmann, Ludwig M. 1976. From Mises to Shackle: An Essay on Austrian Economics and the Kaleidic Society. *Journal of Economic Literature* 14(1): 55-59.

Lachmann, Ludwig M. 1977. Methodological Individualism and the Market Economy, reprinted in *Capital, Expectations and the Market Proces*s (*ed*.) Walter Grinder. Kansas City: Sheed Andrews and McMeel.

Lachmann, Ludwig M. 1977a. Professor Shackle on the Economic Significance of Time, reprinted in *Capital, Expectations and the Market Proces*s, Walter Grinder (*ed*.). Kansas City: Sheed Andrews and McNeel.

Lachmann, Ludwig M. [1956] 1978. *Capital and Its Structure*. Kansas City: Sheed Andrews and McMeel, Inc.

Lakonishok, Josef, Andrei Shleifer, and Robert W. Vishny. 1994. Contrarian Investment, Extrapolation, and Risk. *Journal of Finance* 49(5): 1541-1578.

Langlois, Richard N., and Paul L. Robertson. 1995. *Firms, Markets and Economic Change*. London: Routledge.

Lavoie, Don. 1985. *National Economic Planning: What is Left?* Cambridge, MA: Ballinger.

Leavons, Dickson H. 1945. Diversification of Investments. *Trusts and Estates* 80(5): 469-473.

Leibenstein, H. M. 1966. Allocative Efficiency vs. X-Efficiency. *American Economic Review* 56(3): 392-415.

Leoni, Bruno and Eugenio Frola. 1977. On Mathematical Thinking in Economics. *Journal of Libertarian Studies* 1(2): 101-109.

Leontief, W. 1947. Postulates: Keynes' General Theory and the Classicists. In *The New Economics* (*ed.*) S. Harris. New York: Knopf.

Lintner, John. 1965a. Security Prices, Risk, and Maximal Gains from Diversification. *The Journal of Finance* 20(4): 587-615.

Lintner, John. 1965b. The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets. *The Review of Economics and Statistics* 47(1): 13-37.

Lintner, John. 1969. The Aggregation of Investors' Diverse Judgments and Preferences in Perfectly Competitive Securities Markets. *Journal of Financial and Quantitative Analysis* 4(4): 347-400.

Lintner, John. 1970. The Market Price of Risk, Size of Market and Investor's Risk Aversion. *Review of Economics and Statistics* 52(1): 87-99.

Loasby, Brian J. 1982. The Entrepreneur in Economic Theory. *The Scottish Journal of Political Economy* 29(3): 235-245.

Loasby, Brian J. 1989. *The Mind and Method of the Economist, A Critical Appraisal of Major Economists in the Twentieth Century.* Aldershot, UK: Edward Elgar.

Long, Roderick T. 2006. Realism and Abstraction in Economics: Aristotle and Mises versus Friedman. *The Quarterly Journal of Austrian Economics* 9(3): 3-23.

Lucas, Robert E. Jr. 1972. Expectations and the Neutrality of Money. *Journal of Economic Theory* 4(2): 103-124.

Machlup, Fritz. 1972. *The Production and Distribution of Knowledge in the United States*. Princeton, NJ: Princeton University Press.

Malkiel, Burton. 1973. *A Random Walk Down Wall Street*. New York: W.W. Norton & Co.

Malkiel, Burton. 2003. *The Efficient Market Hypothesis and Its Critics*. Princeton University CEPS Working Paper No. 91 (April).

Mäler, Karl-Göran. 1992. *Nobel Lectures, Economics 1981-1990*. Singapore: World Scientific Publishing Co.

Mandelbrot, Benoit. 1963. The Variation of Certain Speculative Prices. *Journal of Business* 36(4): 394-419.

Mark, N. 1995. Exchange Rates and Fluctuations: Evidence on Long-Horizon Predictability. *American Economic Review* 85(1): 201-218.

Markowitz, Harry. 1952. Portfolio Selection. *The Journal of Finance* 7(1): 77-91.

Markowitz, Harry. 1958. *Portfolio Selection: Efficient Diversification of Investments*. New York: John Wiley & Sons, Inc.

Markowitz, Harry. 1991. Foundations of Portfolio Theory. *Journal of Finance* 46(2): 469-477.

Markowitz, Harry. 1999. The Early History of Portfolio Theory: 1600-1960. *Financial Analysts Journal* (July/August): 5-16.

Marschak, Jacob. 1938. Money and the Theory of Assets. *Econometrica* 6(4): 311-325.

Marshall, Alfred. 1890. *Principles of Economics.* London: Macmillan and Co., Ltd.

Meese, R. A. 1990. Currency Fluctuations in the Post-Bretton Woods Era*. Journal of Economic Perspectives* 4(1): 117-134.

Meese, R. A. and K. Rogoff. 1983. Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample? *Journal of International Economics* 14: 3-24.

Menger, Carl. 1985. *Investigations into the Method of the Social Sciences with Special Reference to Economics*, (*trans*.) Francis J. Nock. New York: New York University Press.

Menger, Carl. [1871] 2007. *Principles of Economics*, (*trans*.) James Dingwall and Bert F. Hoselitz. Auburn, AL: Ludwig von Mises Institute.

Merton, Robert C. 1973. An Intertemporal Capital Asset Pricing Model. *Econometrica* 41(5): 867-887.

Miller, Merton H. 1991. *Financial Innovations and Market Volatility*. Cambridge, MA: Blackwell Publishers.

Miller, Merton H., and Myron Scholes. 1972. Rates of Return in Relation to Risk: A re-Examination of Some Recent Findings, in *Studies in the Theory of Capital Markets*, (*ed*.) Michael C. Jensen. New York: Praeger Publishing Co.

Mises, Ludwig von. 1936. *Socialism: An Economic and Sociological Analysis*, (*trans*.) J Kahane. London: Jonathan Cape.

Mises, Ludwig von. [1957] 1985. *Theory and History*. Auburn, AL: Ludwig von Mises Institute.

Mises, Ludwig von. [1949] 1998. *Human Action: A Treatise on Economics*. Auburn, AL: Ludwig von Mises Institute.

Mises, Ludwig von. 1975. Economic Planning in the Socialist Commonwealth, in *Collectivist Economic Planning*, F. A. Hayek (*ed*.). Clifton, NJ: Augustus M. Kelly.

Mises, Ludwig von. 1976. *Epistemological Problems of Economics.* New York: New York University Press.

Mises, Ludwig von. 1977. Comments about the Mathematical Treatment of Economic Problems. *Journal of Libertarian Studies* 1(2): 97-100.

Mises, Ludwig von. 1978. *The Ultimate Foundation of Economic Science*. Kansas City: Sheed Andrews and McNeel.

Mises, Ludwig von. 1980. Profit and Loss, in *Planning for Freedom*. Spring Mills: Libertarian Press.

Mises, Richard von. 1981. *Probability, Statistics, and Truth*. New York: Dover Publications.

Modigliani, Franco, and Merton Miller. 1958. The Cost of Capital, Corporation Finance and the Theory of Investment. *American Economic Review* 48(3): 261-297.

Mossin, Jan. 1966. Equilibrium in a Capital Asset Market. *Econometrica* 34(4): 768-783.

Mueller, Antony P. 2001. Financial Cycles, Business Activity, and the Stock Market. *The Quarterly Journal of Austrian Economics* 4(1): 3-21.

Newsweek. 1990. *A Nobel Prize for Guessing the Market*. Newsweek 116(18): 56.

Nurske, R. 1944. *International Currency Experience: Lessons of the Inter-War Period*. Geneva: League of Nations.

Ohanian, Lee. 1996. When the Bubble Bursts: Psychology or Fundamentals? *The Federal Reserve Bank of Philadelphia Business Review* (July/August).

Osborne, M. F. M. 1959. Brownian Motion in the Stock Market. *Operations Research* 7(2): 145-173.

O'Driscoll, Gerald P. Jr., and Mario J. Rizzo. 1996. *The Economics of Time and Ignorance*. New York: Basil Blackwell.

Pas, P.W. van der. 1971. The Discovery of Brownian Motion. *Science Historiae* 13(17): 181-219.

Pasour, E.C. Jr. 1982. Economic Efficiency and Inefficient Economics: Another View. *Journal of Post Keynesian Economics* 4(Spring): 454-59.

Pasour, E.C. Jr. 1989. The Efficient-Markets Hypothesis and Entrepreneurship. *The Review of Austrian Economics* 3: 95-107.

Patel, Jayendu, and Richard Zeckhauser. 1990. Shared Price Tendencies: Evidence from U.S. Cities and OECD Countries. *Journal of Business and Statistics* 8(2): 179-189.

Perold, André F. 2004. The Capital Asset Pricing Model. *The Journal of Economic Perspectives* 18(3): 3-24.

Persky, Joseph. 1995. Retrospectives: The Ethology of Homo Economicus. *The Journal of Economic Perspectives* 9(2): 221-223.

Polyani, Michael. 1958a. *Personal Knowledge: Towards a Post-Critical Philosophy.* Chicago: University of Chicago Press.

Polyani, Michael. 1958b. *The Study of Man*. London: Routledge & Kegan Paul.

Polyani, Michael. 1967. *The Tacit Dimension*. Garden City, NY: Anchor Books.

Polanyi, Michael, and Harry Prosch. 1975. *Meaning*. Chicago: The University of Chicago Press.

Pratt, Shannon P. 1967. Relationship between Viability of Past Returns and Levels of Future Returns for Common Stocks, 1926-1960. Memorandum, (Sept.): 561-575.

Rabinovitch, Ramon, and Joel Owen. 1978. Nonhomogeneous Expectations and Information in the Capital Asset Market. *The Journal of Finance* 33(2): 575-587.

Rizzo, Mario. 1979. *Time, Uncertainty, and Disequilibrium*. Lexington, MA: Lexington Books.

Rizzo, Mario. 1995. *Time and Ignorance: After 10 Years*. Economic Research Reports, New York, New York University C.V. Starr Center for Applied Economics (August).

Roberts, Harry. 1959. Stock Market 'Patterns' and Financial Analysis: Methodological Suggestions. *The Journal of Finance* 14: 11-25.

Robbins, Lionel. 1952. *An Essay on the Nature and Significance of Economic Science*. London: Macmillan.

Roll, Richard. 1977. A Critique of the Asset Pricing Theory's Tests' Part I: On Past and Potential Testability of the Theory. *Journal of Financial Economics* 4(2): 129-176.

Romer, David. 1993. Rational Asset Price Movements Without News. *American Economic Review* 83(5): 1113-1130.

Rosenberg, B., K. Reid and R. Lanstein. 1985. Persuasive Evidence of Market Inefficiency. *Journal of Portfolio Management* 11(3): 9-17.

Ross, Stephen A. 1976. The Arbitrage Theory of Capital Asset Pricing. *Journal of Finance* 13(3): 341-360.

Ross, Stephen A. 1978. The Current Status of the Capital Asset Pricing Model (CAPM). *The Journal of Finance* 33(3), Papers and Proceedings of the Thirty-Sixth Annual Meeting American Finance Association, New York City, (December 28-30): 885-901.

Ross, Stephen A. 1987. The Interrelations of Finance and Economics: Theoretical Perspectives. *The American Economic Review* 77(2), Papers and Proceedings of the Ninety-Ninth Annual Meeting of the American Economic Association, (May): 29-34.

Rothbard, Murray N. 1957. In Defense of "Extreme Apriorism." *Southern Economic Journal* 23(3): 314-320.

Rothbard, Murray N. [1962] 2004. *Man, Economy and State with Power and Market*. Auburn, AL: Ludwig von Mises Institute.

Rothbard, Murray N. [1982] 1998. *The Ethics of Liberty*. New York: New York University Press.

Rothbard, Murray N. 1985. Professor Hébert on Entrepreneurship. *The Journal of Libertarian Studies* 7(2): 281-286.

Rothbard, Murray N. 1987. Breaking Out of the Walrasian Box: The Cases of Schumpeter and Hansen. *Review of Austrian Economics* 1(1): 97-108.

Rothbard, Murray N. 1996. *Economic Thought Before Adam Smith: An Austrian Perspective on the History of Economic Thought, Volume I*. Cheltenham, UK: Edward Elgar.

Rothbard, Murray N. 1997. *The Logic of Action One: Method, Money and the Austrian School*. Cheltenham, UK: Edward Elgar.

Rothbard, Murray N. 2006. *Classical Economics: An Austrian Perspective on the History of Economic Thought, Volume II*. Cheltenham, UK: Edward Elgar.

Rothschild, Michael, and Joseph E. Stiglitz. 1970. Increasing Risk: I. A Definition. *Journal of Economic Theory* 2(3): 225-243.

Roy, A. D. 1952. Safety and the Holding of Assets. *Econometrica* 20(3): 431-449.

Rubinstein, Mark. 1975. Securities Market Efficiency in an Arrow-Debreu Economy. *The American Economic Review* 65(5): 812-824.

Rubinstein, Mark. 2002. "Portfolio Selection": A Fifty-Year Retrospective. *The Journal of Finance* 57(3): 1041-1045.

Salerno, Joseph T. 1993. Mises and Hayek Dehomogenized. *The Review of Austrian Economics* 6(2): 113-136.

Salerno, Joseph T. 1994. Reply to Leland B. Yeager on 'Mises and Hayek on Calculation and Knowledge. *Review of Austrian Economics* 7(2): 111-125.

Samuelson, Paul A. 1947. *Foundations of Economic Analysis*. Cambridge: Harvard University Press.

Samuelson, Paul A. 1965. Proof That Properly Anticipated Prices Fluctuate Randomly. *Industrial Management Review* 6(2): 41-49.

Sautet, Frédéric E. 2000. *An Entrepreneurial Theory of the Firm*. London: Routledge.

Schumpeter, Joseph A. 1942. *Capitalism, Socialism and Democracy.* New York: Harper.

Schumpeter, Joseph A. 1989. The Creative Response in Economic History, reprinted in *Essays on Entrepreneurs, Innovations, Business Cycles and the Evolution of Capital*, (*ed*.) Richard V. Clemence. New Brunswick: Transaction Publishers.

Sechrest, Larry J. 2006. Explaining Malinvestment and Overinvestment. *The Quarterly Journal of Austrian Economics* 9(4): 27-38.

Shackle, G. L. S. 1958. *Time in Economics*. Amsterdam: North-Holland Publishing Company.

Shackle, G. L. S. 1970. *Decision Order and Time in Human Affairs*. London: Cambridge University Press.

Shackle, G. L. S. 1972. *Epistemics and Economics*. Cambridge: Cambridge University Press.

Shackle, G. L. S. 1994. The Imagined Deemed Possible: The Germ of the Process of History. In *Uncertainty in Economics*, (*eds*.) David G. Dickinson, Michael J. Driscoll, and Somanth Sen. Aldershot, UK: Edward Elgar.

Sharpe, William F. 1963. A Simplified Model for Portfolio Analysis. *Management Science* 9(2): 277-293.

Sharpe, William F. 1964. Capital Asset Prices: A Theory of Market Equilibrium Under Conditions of Risk. *Journal of Finance* 19(3): 425-442.

Shiller, Robert J. 1984. Stock Prices and Social Dynamics. *Brookings Papers on Economic Activity* 2: 457-498.

Shiller, Robert J. 1989. *Market Volatility.* Cambridge, MA: M.I.T. Press.

Shleifer, Andrei. and Larry Summers. 1990. The Noise Trader Approach to Finance. *Journal of Economic Perspectives* 4(2): 19-33.

Shostak, Frank. 1997. In Defense of Fundamental Analysis: A Critique of the Efficient Market Hypothesis. *The Review of Austrian Economics* 10(2): 27-45.

Shostak, Frank. 1999. Economic Growth and Its Causes: Comment on Holcombe. *The Quarterly Journal of Austrian Economics* 2(2): 67-71.

Simon, Herbert A. 1957. *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*. New York: John Wiley and Sons.

Sinclair, N. A. 1987. Multifactor Asset Pricing Models. *Accounting and Finance* 27(1): 17-36.

Soros, George. 1994. *The Alchemy of Finance.* New York: Wiley.

Soros, George. 1994a. *The Theory of Reflexivity*. Speech delivered to the MIT Department of Economics World Economy Laboratory Conference. Washington, D.C. (April 26).

Soros, George. 1995. *Soros on Soros – Staying Ahead of the Curve*. New York: Wiley.

Steely, J.M. 2001. A Note on Information Seasonality and the Disappearance of the Weekend Effect in the UK Stock Market. *Journal of Banking and Finance* 25: 1941-1956.

Stigler, George J. 1961. The Economics of Information. *Journal of Political Economy* 69(3): 213-225.

Stigler, George J. 1967. The Imperfections of the Capital Market. *Journal of Political Economy* 75(3): 287-292.

Stigler, George J. 1976. The Existence of X-Efficiency. *American Economic Review* 66(1): 213-216.

Stiglitz, Joesph E. 1991. Another Century of Economic Science. *The Economic Journal* 101(404): 131-141.

Stiglitz, Joesph E. 2002. Information and the Change in the Paradigm in Economics. *The American Economic Review* 92(3): 460-501.

Strigl, Richard von.. 2000. *Capital and Production*. Auburn, AL: The Ludwig von Mises Institute.

Stringham, Edward. 2001. Kaldor-Hicks Efficiency and the Problem of Central Planning. *The Quarterly Journal of Austrian Economics* 4(2): 41-50.

Tirole, Jean. 1982. On the Possibility of Speculation Under Rational Expectations. *Econometrica* 50(5): 1163-1181.

Tirole, Jean. 1985. Asset Bubbles and Overlapping Generations. *Econometrica* 53(6): 1499-1528.

Thomsen, Esteban F. 1992. *Prices and Knowledge: A Market-Process Perspective*. London: Routledge.

Tobin, J. 1958. Liquidity Preference as Behavior Towards Risk. *The Review of Economic Studies* 25(2): 65-86.

Treynor, Jack L. 1961. *Market Value, Time, and Risk*. Unpublished manuscript.

Treynor, Jack L. 1962. *Toward a Theory of Market Value of Risky Assets*. Unpublished manuscript.

Vaughn, Karen I. 1994. *Austrian Economics in America: the Migration of a Tradition.* Cambridge: Cambridge University Press.

Walras, Leon. 2003. *Elements of Pure Economics,* (*trans*.) William Jaffe. New York: Routledge.

Williams, John Burr. 1938. *The Theory of Investment Value*. Burlington, VT: Fraser Publishing.

Williamson, Oliver E. 1975. *Markets and Hierarchies*. New York: Free Press.

Wood, Christopher. 1992. *The Bubble Economy*. London: Sidgwick and Jackson.

Working, H. 1934. A Random-Difference Series for Use in Analysis of Time Series. *Journal of the American Statistical Association* 29(185): 11-24.

Yeager, Leland B. 1994. Mises and Hayek on Calculation and Knowledge. *The Review of Austrian Economics* 7(2): 93-109.

VOLUME II

PRODUCTION AND VALUE: MICROFOUNDATIONS OF FINANCE

[T]he success or failure of the investment in preferred stock, bonds, debentures, mortgages, and other loans depends ultimately also on the same factors that determine success or failure of the venture capital invested. There is no such thing as independence of the vicissitudes of the market.

- Ludwig von Mises (1881 - 1973)

Book I

Action and Its World

I. INTRODUCTION

The essence of economics may be summed up as: action through choice. However, contrary to what some have posited in the past, choice is not only the determinant of our actions (see, for instance, Mises 1949: 3). Instead we find that choice *is* also the result of our decisions. With this chosen end in mind, it is the decision-making process which makes use of all our capacities as humans, and drives us towards these ultimate ends – the choices which shape our actions.

There is, of course, some degree of circular reasoning with our conception of choice. A confusion abounds between the physical choices that we see manifested through action, and mental choices which remain shrouded, forever hidden in our minds. We can see that mental choice usually precedes the physical. Yet the physical choices we make also result as new determinants in future choices which were heretofore not known to exist. If choice is the incorporation of our preferences and means, is it not also choice that sends into motion, and shapes the process that results in these entities – these intermediate steps of action?

In fact, it may be said that to act is to choose. And by the same circular reasoning, every choice is a conscious action. A feedback loop of infinite regress occurs, with no clear way of determining which came first, the action or the choice. Due to the finite span of an individual's life, it is evident that the regress in question is not of infinite character, but must stop (or as the case would be, it must begin) at a certain point – the point where the first action occurs. It is at this first moment of choice that we can determine what it is that causes our physical actions; the active realization of our choice. The origin of life for the physical individual may be contestable – paying heed to our ethics, morals, intuition, or science – but for the economist life commences definitively; the *acting human begins life with a choice*.[149]

The distinction between choice and reaction becomes crucial. Choice involves the active taking and renouncing on the behalf of the actor. It is a subjective weighing of the benefits expected to be received from the enactment of the choice, with the costs to be incurred in order that these benefits can be received. Reaction involves nothing more than an unconscious response to stimuli; an

[149] See Mises (1933: 15): "As an a priori category the principle of action is on a par with the principle of causality. It is present in all knowledge of any conduct that goes beyond an unconscious reaction. 'In the beginning was the deed.'" This deed is, to Mises, the recognition that action stemming from a choice (the two being inseparable) is the start of all analysis concerning the focus of our attention – the decisions of humans. In fact, Mises likely became aware of this connection through Adam Ferguson's (1767) *An Essay on the History of Civil Society* where he compares vegetables growing from their tender shoots to animals destined to act.

unpremeditated reflection to a felt catalyst. Reactions and choices can, and do quite often, happen concurrently. Mises gives his example of a person sickened by a virus. As the virus attacks their body, a reaction occurs whereby the body instinctively creates antibodies to combat the virus' advance. This reaction cannot be confused with the purposeful action at the hand of the individual who realizes their health is at stake, and goes to a physician to have this attended to. Two very similar events occur – the action and the reaction – simultaneously, but they are merely similar in an objective way. To the observer conscious of the essence of choice, the difference between these two very similar events becomes manifest.

If choice becomes a question of gaining and renouncing (or receiving and giving), we find that this first step occurs quite early in the lifetime of an actor. In fact, it happens almost immediately at birth in most cases. But it also, then, becomes apparent that at this early stage in an individual's life resources with which to renounce are negligible; in some cases nonexistent. But regardless of how extreme the lack of resources may seem to be, the defining characteristic of humans' remains – reason. It is through this exercise of reason that the young actor renounces and takes through the weighing of the costs and benefits, unconscious at this time of the complex process which seems to be progressing so naturally. The reason inherent in the mind weighs the cost of breathing and crying, and so, even at this young age, makes possible the choice that is to demarcate the long process of choice that will continue as long as the life continues (economic life, not physical life, for the two are separable concepts).

We find that the concept of choice has two central themes which are evident at every moment of choice: knowledge and rationality. Our continually and forever imperfect knowledge provides the reason for our choices.

Knowledge is that which we use to structure our choices. A total lack of knowledge would leave no basis for a choice to be undertaken. Two explanations account for this. First is the lack of influence our reason would have if there was no knowledge to use it upon. Reason provides the tool, and knowledge the fodder, for individuals to structure their choices. Above we saw that a young infant lacking all resources external to their character was still able to chose and act. Knowledge was an essential part of this action, and although difficult to see what type of knowledge this relatively undeveloped mind could have, it provides an inextricable link to the choice they made. Consider a situation where the child had no knowledge that what they needed at that moment was air to fill their lungs to continue life. They need not know the process by which a continual, fresh supply of oxygen sustains their life, but they do need knowledge of two things. First, knowledge of the fact that they feel an uneasiness surrounding their current state of affairs. Second, they need to know that there is something attainable to them that removes or reduces this uneasiness. It matters not that whether the child had knowledge of this *a priori*, or whether they discovered it *empirically* through repeated trials. The simple fact that the child had the reason within them to transform the felt knowledge of experience (the discomfort of a lack of oxygen) into a chosen act to forgo one action, and undertake another. The child's reason would be of no avail if they could find no knowledge with which to place under its command, and continue the actions necessary for life.

The second need for knowledge stems from the requirement of known alternatives, necessary for choice. Choice is defined as the renunciation of one state of affairs, in exchange for another. Inexorably, we see that knowledge of two states of affairs are requisite for choice. The trade-off between the experienced, and the yet to be experienced, shape the decision that results in choice. The infant above could not choose if they lacked knowledge of an alternative state of affairs to the one they were born in. If they could not *imagine* a state of affairs better than the one they existed in there could be no possibility for a choice to alter that same state of affairs. And yet, we know through the very act that the great majority of newborns continue their lives on their own that this essential choice is made.

The source of this new knowledge of an alternative is of no consequence for the infant. Whether it was gained accidentally, or through a mere reaction, or whether it was at the hands of a physician forcing air into their lungs, the knowledge is taken by their reason and allows to be compared with their current knowledge. The act of breathing may seem so small and natural for a more mature individual that it is difficult to conceptualize the motions that were involved in our own first breath. However, difficult as it is to remember the process, the initial steps remain. The mixture of knowledge with reason resulted in that first choice which marked the start of our lives as *homo agens*.

Choice remains forever the defining characteristic of economic life. The subject that concerns us is the actions of individuals, and how these actions are shaped and formed. The determinants of choice are a seeming infinite regress of feedback loops – choices determining knowledge determining alternatives determining choices – but we have seen that the seemingly infinite regress halts, or commences, at one specific point: the first choice. And we have seen that at this first choice, humans have at their disposal but one capacity to make their first choice and commence life: reason. This logic of reason is where we will have to embark from to determine from where our choices originate.

II. THE ABSOLUTE TRINITY OF BEING

1. Logic

The French Thomist philosopher Jacques Maritain (1946: 1) defines logic as the means of acquiring and possessing the truth. It is the "art" which directs all acts of reason, and enables humans to advance upon the act of reason itself.[150] For Maritain, logic not only conforms to our reason, it "bears upon the act of reason itself." As Mises (1949: 34-35) would similarly write:

> "[T]he problem of the a priori . . . refers to the essential and necessary character of the logical structure of the human mind. The fundamental logical relations are not subject to proof or disproof. Every attempt to prove them must presuppose their validity. It is impossible to explain them to a being who would not possess them on his own account... The human mind is not a tabula rasa on which the external events write their own history. It is equipped with a set of tools for grasping reality.

Logic is that which is bestowed in us by right of our position as humans. We are born with it, and it shapes our decisions and actions throughout life. Logic is the absolute which defines our choices. Despite what some have tried to adduce throughout history, there can be no such concept as separate logics. The polylogists of the past have been amply put to rest by Mises (1949: 75-89). In fact, as he (1933: 102-103) laid to rest the issue and clarified the argument:

> The first point to be established . . . is that none of the sources of historical information accessible to us contains anything that could shake the assumption of the immutability of reason. Never has even an attempt been made to state concretely in what respects the logical structure of reason could have changed in the course of the ages. The champions of historicism would be greatly embarrassed if one were to require of them that they illustrate their thesis by pointing out an example... [I]n what way [is] the logic of primitive peoples ... structurally

[150] We can find agreement with *The America Heritage Dictionary* (4^{th} ed.) as it defines logic as "the principles of reasoning, distinguished from their content." Hence, we find logic as an absolute measure. With no basis on content, there can be no place where logic can be said to apply only to one situation, but not to another. For how could any situation differ from another if not by the content of its occurrence?

> different from our logic.[151]

The logical structure of the mind is not an arbitrary construct, existing independent of place and time. Instead, it is the one constant within all humans, the common bond that links us together.

The whole of the realm of logic, however, relies on a presupposition: a complete knowledge of the applicable facts of a situation. Many logical results need no immediate concern for this fact, as they are themselves grounded in the pure logic of our minds. Take a concept such as the Pythagorean theorem; we know it to be pure logic by definition. The results of this need not be proved, they are themselves based on truths. The reason for this is that the inputs – the building blocks of the theorem – are themselves able to be viewed as mental constructs. We need no knowledge of an external world to prove the relationship between the hypotenuse of a right-angled triangle and its two adjacent sides to be true. In fact, we need no information of this shape other than that it has three sides, two of which connect at ninety degrees; it makes no matter what the physical representation of such a construct is, provided our mental construct incorporates these two points.

It must be understood, however, that these truths (logic) are not mere categories *of* our minds, they *are* our minds. As logic is what defines our mind, we see that the absolute nature of logic, its irrepressible authority, must not only apply equally to all, but be endowed equally *in* all.[152] The child at birth has as much capacity for logic as an adult near death. In fact, the implication that the *same* individual could have differing degrees or capacity for logic throughout their life can be summarily dismissed on the same basis as the arguments against polylogists. Indeed, as Hayek (1952a: 135) writes:

> To speak of a mind with a structure fundamentally different from our own, or to claim that we can observe changes in the basic structure of the human mind is not only to claim what is impossible: it is a meaningless statement. Whether the human mind is in this sense constant can never become a problem—because to recognize mind cannot mean anything but to

[151] Interestingly, Mises was not always immune from this trap of polylogism himself. As Yates (2005: 52) brings to light, Mises (1949:33-34) also notes that logic can be of an "historical phenomena" that can be "transitory." That Mises confounds two very similar concepts here, reason and logic, will soon be shown.

[152] See, for example, Hoppe (1995: 20) as he mentions that: "We must recognize that such necessary truths are not simply categories of our mind, but that our mind is one of acting persons. Our mental categories have to be understood as ultimately grounded in categories of action." The common connection between humans – action – becomes a common connection due to the constancy of logic.

recognize something as operating in the same way as our own thinking.[153]

It must be remembered that logic as a concept is applicable only towards means, never to their ultimate ends. In fact, lacking any internal content of its own, we can see that the relevance of logic toward ends is completely irrelevant. Take the previous example of the Pythagorean theorem, what meaning is there to say that the theorem is a logical end in itself? In fact, although the internal structure of the theorem (the means by which the end is achieved) are logically consistent and derived accordingly, no such statement could ever be made about the end itself. Was the derivation of the theorem a logical choice lacks any answer except "we don't know."

We see logic as the absolute; the defining characteristic all humans share. Any notion of differing degrees of logic are completely unfounded. Not only do all individuals share the same logic base, but that base remains constant throughout their life. An increase in logic, or one person using more logic than another can have no significant meaning. However, the extent to which logic can be used is not absolute (even though its results will be). The extent of logic in many situations is limited by the amount of knowledge that the mind can apply the logic to. Some instances, as we saw with the example deriving the Pythagorean theorem, require no external input or information to complete; our minds internally hold all the knowledge we need to deduce these conclusions. However, many other situations in which we wish to use our logic require an input from the outside world. Previously, we saw how an infant exercises their logic to deduce two options that unfold before them at birth – to breath or not breath. They lack any prior knowledge as to which option is preferable – that option which will allow them to continue life. However, with the knowledge gained through the external world, they soon discover that the sensation of breathing is preferable to that of not breathing, and so are allowed to exercise their logic and choose to continue life with a breath.

The mind, then, plays a significant role in how this process is to continue. For it is within the confines of the mind that we see the boundaries of our logic's use are set. It is only within the limitations of information that the mind can internalize and know that we have a restriction set on the extent that our logic may be used. How this limit is created is yet to be seen.

[153] Although it must be pointed out that Hayek failed to remain fully consistent himself in this viewpoint. Later in his career, he (1988: 54) would comment: "Like other traditions, the tradition of reason is learned, not innate. It too lies between instinct and reason; and *the question of the real reasonableness and truth of this tradition of proclaimed reason and truth must now also be scrupulously examined*." This controversial book, aiming to show that Socialism represented a *logical mistake* seems to imply a type of logical duality, much like the Socialists he tried to refute themselves believed.

2. The Mind

The mind is the tool we have at our disposal to connect the outside world with our logic. It is this connection that makes possible the perceived choices which we set into action. This connection is made possible through a series of steps, each having grave importance on the following, and each needed for the logical continuation of action. Stimuli exist in the external world that are required to be mixed with out internal logic. These stimuli become the starting point, hence, of every action we undertake. The process by which the mind incorporates these stimuli with our logic is quite complex (in fact, so complex that, as we shall soon see, evades our complete understanding), but a basic road map of the process is available to guide us sufficiently along its path.

Hayek specifically wrote *The Sensory Order*, to find an "explanation of mental phenomena" which relate the problem of bridging the divide between the physical (external) and phenomenal (mental) orders (1952b: viii). More exactly, his search was to "know the kind of process by which a given physical situation is transformed into a certain phenomenal picture" (1952b: 7). This physical situation, what we may refer to as stimuli, go through two, two-step, processes that we must attend to.

First are the situations where stimuli are processed through the mind, and transmitted back as stimuli in the external world through a particular outcome. The process which the mind evolves these external stimuli into a subjective interpretation within the mind is of as great an importance as the process through which this is transmitted to the external world (which is to say, the reverse process by which had earlier occurred). The second situation of interest is the forming of memory. Memory is that achievement which eliminates the necessity to learn anew from the external world. It is the storing of stimuli interpreted by the mind *within* the mind. These two processes – interpretation and storing of stimuli – are the central roles the mind accomplishes with only logic at its immediate disposal. The outcomes of this phenomenal process are the manifestations which individuals act upon.

Lastly, the consciousness of this mental activity must be explained. For we know that each true "action" that an individual undertakes must also be a conscious action; like Mises' "rational action" we can also find this a pleonasm; no action can be meaningfully understood to occur without a conscious intent on the part of the actor.

Although the process may proceed from stimuli to outcomes or memory, it may be easier to conceptualize this process from the reverse direction. It is, after all, stimuli in our expectations (a state of mind) or in our memories that we think of if asked a simple question such as "tell me one external factor which has influenced your thoughts on this subject.” The memory, being the final resting place for this process, provides the first place to look when describing this logical development.

Memory

Until Hayek (1952b) there was one predominant view of the mind. The “storage” theory viewed every new mental entity as being held in one place of the memory, awaiting for a need to bring relevance to its use to arise. Hayek's alternative to this theory can be viewed as a linked-system. Memory is not a place in the mind, but instead is represented by the paths that interconnect and join the otherwise disjointed pieces of memory.

Stimuli exist in our minds as pieces of knowledge “linked” together to form a web of distinct, yet interrelated knowledge. Memory is that system of interconnecting webs that links together these pieces. Think for a moment of a piece of knowledge held within the mind, but that held no common link with any other piece of knowledge. In fact, we cannot even think of such a piece of knowledge, as lacking a link to any other piece of knowledge, there could be no way to recall its existence. This is the essence of Hayek's “linked-system” of memory. Memory and experience can only be used in the terms that imply their “linkage” together (Hayek 1952b: 105). As Polanyi (1958b: 28) stressed, the structure of our tacit knowledge is also our act of understanding. Placing these disjointed pieces together into a greater whole is the process by which we remember, and hence open them up to recall and use, disjoint pieces of knowledge.

Of course, there have been other attempts at classifying the essence of memory since Hayek. Tulving (1972), for example, gives the dichotomy between episodic memory – that which requires experiment and a temporal existence – and semantic – that which is knowledge based and necessarily only available to the conscious thinker. However, the assertion that a type of memory exists for the unconscious individual seems to provide a paradox. For how can we remember memory, which seems to be the only action we *can* undertake with this concept, without being conscious of doing so?

Memory derives its strength through the path dependency that links these disjointed pieces of knowledge together. As more and more paths connect a piece of knowledge with other pieces, the *stronger* (which is to say, the easier to recall), a certain piece of knowledge will be in the future. Furthermore, as more and more connections link a piece of knowledge to other pieces, the character of that knowledge becomes less assailable (Hayek 1952b: 105). A piece of knowledge held in memory by only one connection to another piece of knowledge may be changed if a new piece of knowledge enters the linkage not fully coherent with the existing link. For instance, suppose that the only link that the memory of a flower has for the individual is with another memory, that of a wedding. The meaning to this individual of this flower will be largely shaped by its path dependency on this sole linked piece of knowledge. Now suppose that an individual attends a funeral and also sees flowers. This provides the second link that flowers have within their memory relations. Prior to this occurrence the memory of a flower was largely shaped by one happy memory, now it is known that, through remembered experience, that this happy association need not always be the case.

Assume, however, that prior to seeing a flower at a funeral, this individual had a memory of a flower that had ten existing links, all similar happy thoughts (such as weddings, anniversaries, birthdays, and a trip to a botanical garden). These ten links are how the individual is able to recall the concept of a flower (conversely, these events, such as the wedding, are recalled through association with the flower, among other potential links). If one new event occurs, for example, the funeral where a flower is apparent, one new link enters the flower-memory. However, the existing links are quite numerous and strong, they outnumber the new linkage in quantity, and potentially quality (that is, how strong they themselves appear in the individual's mind). The new linked memory of a funeral will have considerably less weight on the meaning of the existing memory due to these factors.

Hence, we see memory strength is the result of two inter-related factors. One is the number of links to other memories; the more links gives rise to increased events that will occur and revive the memory in question. Second, and equally important, is the relative strength of each link associated with a memory. This depends on an inter-related loop of strength whereby these linked memories' strengths rely, in part, on the strength of the memory in question. The strength that we refer to regarding memory is two-fold as well. Strength in the inflexible character of a memory, and strength in its ability to be recalled at ease. As Hayek (1988: 87) recognized: "Each bit of knowledge contributed by one person will tend to lead some other to recall yet other facts of whose circumstances of which he did not know."

How is it then that a memory is altered within our mind? The question we have, in part, already answered. Hayek (1952b: 114) makes note that there are two distinct types of memory which are susceptible to alteration in two different ways.

One is the change that occurs to the structure of connections, that determines the path in which memories become recalled. The alteration of connections leads to changes in the path dependencies, which may result in alterations concerning the meaning, and the ultimate strength, of a memory. The other source of change is through the pattern of new impulses received at any given moment owing to new stimuli received. This continual flow of stimuli provides a perpetual alteration in the flow of impulses which recall, and determine, memories. This endless flow will never cease to exist as long as external stimuli continue to be induced. In fact, as Edelman and Tononi (2000: 93) describe the process, "memory is more like the melting and refreezing of a glacier than it is like the inscription on a rock... memory is not a representation; it is a reflection of how the brain has changed its dynamics in a way that allows the repetition of a performance."

Memory is the interconnected web of memories; a conundrum of sorts. For it is through the alteration of this preexisting web of connections that a memory becomes, itself, altered. But to determine what it means to say the network of connections is altered, it seems that satisfaction can be reached only in determining what this network is. Hence, it is actually the mind which requires description.

The Mind as a Network

Memories represent the actual linkages that connect the individual datum – those pieces of knowledge which exist in a dispersed form, but are united by their common network. The actual network, however, is what the mind actually is – that physical array of neurons which form together as a complex nexus of memories. This web of connections is the link that provides external stimuli the opportunity to gain internalization within the individual's mind.[154]

The existence of memories will effect the newly arriving stimuli and determine how they will be evaluated, and the responses that will result. Hence, as the system of connections – the mind – shapes the influx of new stimuli, Hayek (1952b: 145) points out a quandary of the memory, for: "The 'model' of the physical world which is thus formed will give only a very distorted reproduction of the relationships existing in that world; and the classification of these events by our senses will often prove to be false, that is, give rise to expectations which will not be borne out by events." Our memory aims to reproduce the stimuli as they are faithfully received. However, this portrayal forever remains clouded by the existing structure of our mind. In fact, this portrayal of a stimulant provides for an even

[154] At this point it should be stressed that experiences that seem to occur pre-sensory, and those provided through tacit knowledge belong to two different realms. Pre-sensory knowledge, or logic, is that which allows the mind to make qualitative discriminations. In opposition, tacit knowledge is that which is a form of qualitative discrimination – a product of the logic of the mind. The difference between the local knowledge of "time and place" (tacit) and universal knowledge of the internal universe (logic) becomes apparent (Hayek 1945: 80).

greater range of interpretations when we allow for the uniqueness of not only the neural map of the mind, but more importantly, the resulting memories that are created.

The mind of an individual is plagued with a continual shift as new stimuli bombard it with ever changing links and pathways. This very fact gives rise to the important point that no two people will ever have a similar pathway for a memory; the uniqueness of experiences precludes this possibility. More importantly, no *single individual* could ever have the same neural map from one moment to the next. As long the influx of new stimuli never ceases, a continual remapping will occur. This in itself does not have any specific detrimental consequences for human interactions. In fact, much like Hayek (1952b: 110) explains it: "Two people discussing the same walk, with different maps of the region before them, will in general encounter no difficulty in understanding each other, although particular points on their route may have different significance for them." Mises (1960) would consider the understanding of these two individuals' choices despite their different maps as the key role of the economist. For instance, if one were to watch the busy movements of workers throughout Grand Central Station, one may be predisposed to follow where they go, time their travels, and try to establish some sort of law or formula to describe where they will go next. But the true economist would have to ask why it is that each is going to the specific place, at the specific time, that they are. Only through a closer inspection of their intent, their logical desire to act in such a way, can the economist make a meaningful statement about their choices.

Furthermore, this uniqueness of the mind can be looked at from an even deeper vantage point. The semi-permanent structure that it comprises allows for what types of *classes* or *qualities* of stimuli may be stored. However, this does not say anything about the specific events that will be recorded at any particular moment.

We find, however, that the mind, like any living entity, has a shifting value scale over time. Turning again to the insight of Hayek (1952b: 118), we find that:

> Once such a continuous reproduction of the environment is maintained in the highest centres, it becomes the main function of the sensory impulses to keep this apparatus of orientation up to date and capable of determining the responses to particular stimuli in the light of the whole situation. The classification of impulses is then no longer specific to particular functions, but has become general in the sense that any one of them may, by its position in the comprehensive pattern, exercise some influence on practically any response.

Up to a point, stimuli are used and added to grow the mind, to expand its reaches of understanding through the use of its logic. This involves assessing each new stimuli in light of the individual pieces of memory already in use in the mind. A point becomes realized where the cost of doing so, of comparing each new stimuli with each already received, becomes prohibitively high. The comparison shifts between growing the structure and breadth of the mind, to maintaining its cohesiveness. This is done by comparing new stimuli to the greater whole, or the general apparatus of the mind. The mind has become so intertwined with linkages, that any new stimuli may illicit an alteration from almost any other existing memory.

We see, as a result, that as stimuli increase in numbers, with the resultant increase in memories, that each new stimuli will reach its network across an ever expanding array of memories. Previously we saw how the more linked memories exist, the more durable any given memory is to change. When viewed on the more general scale of the mind, we see a mind less susceptible to *quick* change. The path interdependencies that increase the range and strength of knowledge come at a price; the loss of mobility for the mind to restructure in the face of new incoming stimuli. Novel stimuli play an increasingly smaller role in the shaping of the larger mind (Hayek 1952b: 112).

It is these stimuli that ultimately determine the range of the mind's ability to use its logic, a fact which should now be given a critical note.

Stimuli

The central point that should be stressed with the sensory order of the mind is the duality of worlds that exist. The phenomenal order of our mental world (the mind) and the physical order of stimuli (external information). As Mises (1933: 75) points out, "[w]hat matters is not the data, but the mind that deals with them." Specifically, he was referring to the acts of pure logic that Galileo, Newton, Ricardo, Menger, and Freud had earlier used. However, we also see that although a purely phenomenological realm exists – that of pure logic – much of our decisions are based on an external datum; a stimuli's effect on the logic of the mind.

Several problems with external stimuli become apparent. Previously it was showed that even praxeologically, the indeterminateness of a stimuli's reception and linkages had consequences for an individual. However, we see this problem compounded when viewed within catallactic exchange.

While transmitting a memory, the knowledge in question can only be transmitted by regarding the effect it has previously had on other pieces of knowledge (other memories). A confusion arises between where the source of the knowledge came from. What we are transmitting could be mistakenly referred to as one of two things. It may be that what we are transmitting is the original source stimuli; the original object or information that we learned. Alternatively, it could be that what we are transmitting is the mental manifestation, with all the effects of path dependency, is not the original stimuli we received, but instead is the resultant memory under the influence of many other memories. It becomes clear that what we transmit to others is never an original stimulus, but rather a *proximal stimulus* (Hayek 1952b: 8). In some cases this proximal stimulus will be so far removed from the original stimulus that the transmission may fully lack in being a faithful representation of its original character.

This would be of no immediate consequence if catallactics were themselves of no immediate concern to us. However, we see this is not a reasonable stance in a developed society with its plethora of interrelations. Even in the most primitive of societies it would be a questionable assumption. Of course, as Polanyi (1958a: 54) reminds us, some knowledge can only be transmitted through a physical interpersonal method:

> Connoisseurship, like skill, can be communicated only be example, not by precept. To become an expert wine-taster, to acquire a knowledge of innumerable different blends of tea or to be trained as a medical diagnostician, you must go through long courses of experience under the guidance of a master.

In fact, the problem potentially compounds itself exponentially the more times this knowledge must be transmitted from person to person.

However, the problem of stimuli and their transmission is limited and mitigated somewhat. It is not that the whole realm of stimuli is of concern to the individual. Instead, we see that the consciousness of logic allows for a defined set of stimuli to be the focus of our attention, and hence, decisions.

Consciousness

Polanyi (1958a: 103) maintains that the mind is not the aggregate of its fully known memories, instead, it is that which we focus our attention upon while having only a *subsidiary* awareness of its potential manifestations. Consciousness becomes an important facet of the human experience, affecting our mind in two ways. First is the awareness we attribute towards stimuli; those external events which we must sense prior to internalizing them within our mind. Second is the consciousness we must accord to the memories within our mind, if we wish to make use of them. It is not possible within our capacity to focus what we commonly refer to as "attention" upon all such internal memories and external stimuli.[155] We must, instead, employ some level of discretion to this process, one which allows for greater emphasis to be placed on certain pieces of knowledge, at the expense of others.

Thierry (2008: 32) notes that consciousness is the ability of an individual to sense opportunities around them: "[T]he more conscious the phenomena, the greater the opportunities for action and the more open the possibility to formulate plans. Consciousness is the means of identifying opportunities." This awareness of opportunities, or stimuli, corresponds to the entrepreneurial function stressed in Kirzner (1973) of "alertness." This attention given to external stimuli is that which allows greater range to the possible memories that may be internalized within the mind, and hence, expand our scope for action.[156]

The consciousness heeded to our existing memories may be of even greater importance however. Polanyi (1958a: 55) points out that consciousness is necessarily focused. This implies that something which our consciousness is directed towards is offset by another which is neglected by this same attention. Only *distinct sets* of memories may be evaluated at any given time. Remember that memory is that linkage created through the network of the mind. Although a thoroughly developed mind may have linkages bringing a deep interrelation between all memories, *this need not be the* case. In fact, memories can exist independent of other memories, provided that the linkages connecting them with the conscious arousing stimuli does not cross. The result is that while we can see these phenomenological processes taking place simultaneously, they need not always affect each other. Multiple consciousnesses may occur, but only to the extent that one memory is isolated from another.[157]

---

[155] Boring (1933) notes that as consciousness is attention, and our attention is selective, that our consciousness must necessarily be selective also. As Hayek (1952b: 231) points out, as the two concepts – attention and consciousness – are "synonymous" we find the same fundamental principles in both.

[156] Or, as Hayek (1952b: 139) writes: "The experiences to which our attention is directed are more fully discriminated and perceived in greater detail than others of which we are also aware... We notice more in them and are more fully prepared to respond adequately to their occurrence."

[157] This, in contrast to Hayek (1958b:136) who posits "more than one 'unconsciousness' (or coherent system of unconsciousness mental events) while there exists (normally) in any individual only one consciousness." While we may be unconscious of any number of phenomena, we may also be conscious of a distinct set of phenomena, *provided this*

With this outline, we can see the dichotomy between subsidiary and focal awareness Polanyi (1958a: 56) writes:

> Subsidiary awareness and focal awareness are mutually exclusive. If a pianist shifts his attention from the piece he is playing to the observation of what he is doing with his fingers while playing it, he gets confused and may have to stop.[footnote omitted] This happens generally if we switch our focal attention to particulars of which we had previously been aware of only in their subsidiary role.
>
> This kind of clumsiness which is due to the fact that focal attention is directed to the subsidiary elements of an action is commonly known as 'self-consciousness.'"[158]

Why is it that the pianist would stumble upon changing their focus from one aspect of their actions to another? It stems from the confusion of attempting to focus on two interrelated memories simultaneously. For as we have seen, the mind can focus attention only to the extent that a memory belongs to a unique neural pathway. By changing the focus of our attention we have changed the memory in question from one that is disconnected from the pathway of another current thought, with the confusing effect brought on by the new memory dependency. Polanyi's focal awareness is that which we actively think about; the memory we actually recall at any given time. However, subsidiary attention is drawn on any number of unrelated (neurally) memories.[159]

Hayek (1952b: 132-154) gives several "additional characteristics" of consciousness that should be mentioned: 1) consciousness must be able to be formalized (i.e., communicated to others), 2) an

---

distinctness of path relations remains. Additionally, Hayek (1978: 45) notes that "what we consciously experience is... the result... of the processes of which we cannot be conscious." It becomes clear that consciousness is not an objective quality that can necessarily be increased, but instead stems from the depths of the mind in our quest for further understanding.

158 Polanyi (1958a: 55) provides the following clear illustration:

> When we use a hammer to drive in a nail, we attend to both nail and hammer, *but in a different way*. We *watch* the effect of our strokes on the nail and try to wield the hammer so as to hit the nail most effectively. When we bring down the hammer we do not feel that its handle has struck our palm but that its head has struck the nail. Yet in a sense we are certainly alert to the feelings in our palm and the fingers that we hold the hammer. They guide us in handling it effectively, and the degree of attention that give to the nail is given to the same extent but in a different way to these feelings. The difference may be stated by saying that the latter are not, like the nail, the objects of our attention but instruments of it. They are not watched in themselves we watch something else while keeping intensely aware of them. I have a *subsidiary awareness* of the feeling in the palm of my hand which is merged into my *focal awareness* of my driving the nail.

159 Likewise, Edelman (1990) and Edelman and Tononi (2000: 102) stress the higher-order and the primary consciouses. The primary is that which we share with many other species, it is the ability to have this higher-order (or Polanyi's subsidiary) consciousness that allows humans the ability to act upon many distinct memories coherently and simultaneously.

individual must be able to take account of other conscious experiences when making a representation of another conscious phenomenon (i.e., all conscious phenomena are to be considered homogeneous in this sense that they are all linked together), and 3) plans and actions are necessarily guided by the consciousness of the present situation.[160] These can be contrasted with *mere* tacit phenomenon which cannot be communicated to others, may not be directly taken into account when planning an action and remain distinctly heterogeneous *by definition* as they forever remain inarticulable, thus escaping any type of mental activity that could bring homogeneity to their existence.

We thus see the important role that the conscious plays in our mind. The focus of our attention is the memory with which we make plans upon, or the stimuli with which we create a memory from. However, consciousness comes with a pitfall. For what does it mean to have greater consciousness of an event? Theirry (2008: 32) answers by noting that the more conscious we are of something, the more "centralized" our responses become. Consequently, the increased centralization of the classification process – the internalization of stimuli to memory – results in a marginalization of peripheral foci (the subsidiary awareness of Polanyi). The result are responses that are more homogeneous and of increased unity. Our subsidiary awareness is marginalized as a result of increased attention on our primary focus.

Up to this point is has become clear that the logic within our mind, combines with external stimuli to create memories. These memories, or what may be more commonly referred to as knowledge, exist within the mind – an interrelated web of neurons. The strength of these memories depends on the number of path dependencies each memory has, and the related strength of these respective paths. The conscious plays the role of defining what stimuli are absorbed by the mind, and what memories (knowledge) are focused on when we create plans, and enact action. Our focus should shift, then, to look at the actual outcome of this process, and see what issues can arise in these mental outcomes leading to our choices.

Outcomes

We have briefly touched upon the indeterminableness of outcomes caused by stimuli filtered through the mind. The dynamic network that comprises the mind forever shifts and changes the focus of our attention, and the resultant meaning of our memories. Moving backward from the resultant outcomes, Hayek (1952b: 194) notes the, at times, foreign and unclear origins of these outcomes:

---

[160] See Thierry (2008: 33) for a further summary and explanation of these three points.

> [*T*]*o us* the mind must remain forever a realm of its own which we can know only thorough directly experiencing it, but which we shall never be able to fully explain or to 'reduce' to something else. Even though we may know that mental events of the kind which we experience can be produced by the same forces which operate in the rest of nature, we shall never be able to say which are the particular forces which 'correspond' to a particular mental event.

In fact, there are three unique cases that we can look at when we analyze the uniqueness of results: the actual results, and the stimulant and individual in question. These three factors give rise to eight unique cases where stimuli affect outcomes in different manners, summarized in table 1.[161]

Determinants of Outcomes		
Stimulant	**Person**	**Result**
Same	Same	Same
Same	Same	Different
Same	Different	Same
Same	Different	Different
Different	Same	Same
Different	Same	Different
Different	Different	Same
Different	Different	Different
Table 1		

As we have outlined, most of these possibilities seem entirely possible when viewed in light of the aspects of individual memory maps and changing values, preferences and foci. However, several may cause unease. DS-SP-SR suggests that a different stimulant may cause the same result with the same person at different times. This outcome seems highly unlikely, but cannot be excluded from the possible set. Second, the event DS-DP-SR seems equally unlikely, but cannot be excluded. These are imaginable events, although admittedly hardly so. Even for individuals obtaining the outcomes personally, there is an element of uncertainty as to what process was underwent. As Polanyi (1958a: 49) makes clear, "*the aim of a skillful performance is achieved by the observation of a set of rules which are not known as such to the person following them*." It seems as though our ignorance as to the actual

[161] See Hayek (1952b: 9) for a summary of some characteristics that lead to the indeterminateness of outcomes.

process may be the source of our success in, at least some forms of, action.

It seems as though the indeterminate character of outcomes leaves us with little opportunity for exploring the thoughts that shape our decisions. We must focus our attention on two factors that shape these outcomes, each equally important. The first is the character of the stimulus that is influencing the mind of the decision-maker. Second, we need to accord attention to the existing mind that an individual has had shaped by previous stimuli. This second part may possibly only be known through the resultant outcomes, much like Polanyi (1958b: 33) posits: "A man's mind can be known *only comprehensibly, by dwelling within the unspecifiable particulars of its external manifestations*." It is not only enough for us to know how the mind works, indeed this may not even be a possible eventuality.[162] However, it is important to try to gain understanding of this through the external demonstrations of this phenomenological process.

Appendix A: A Model of Humans?

Much debate surrounds the topic of *if* humans can be systematically modeled as machines. In this era of increasing computational power available at our disposal, it seems that it is only a matter of time before an algorithm comparable to a functioning human mind can be created. This idea overlooks several of the key aspects of the mind.

Polanyi (1958a: 263) tells us:

> ... neither a machine, nor a neurological model, nor an equivalent robot, can be said to think, feel, imagine, desire, mean, believe or judge something. They may conceivably simulate these propensities to such an extent as to deceive us altogether. But a deception, however compelling does not qualify thereby as truth: no amount of subsequent experience can justify us in accepting as identical two things known from the start to be different in nature.

Later (1958a: 369), he adds:

> Any machine that is to represent learning presupposes a theory of acquiring knowledge and a

[162] Hayek, while writing *The Sensory Order*, was influenced by Gödel's theorem whereby, as Dennett (1995, 429) sums up, "there are truths that "we can see" to be true that can never be formally proved to be true." Hence, in no consistent system can something prove its own consistency. In other words, no system has the capacity for complexity to describe itself.

theory of knowledge itself.

To think that a machine could be created commensurate to a human mind assumes that this same machine could have the ability to know the *meaning* of knowledge. However, the complex personal structure of the mind we have just laid out should make it clear that this possibility is excluded. For what would it mean for a machine to have only subsidiary knowledge of one action, or to make the full focus of their attention merely one stimulus or memory? An algorithm could never succeed to model a mind, not because it fails to be perfect enough, but because it cannot be *im*perfect enough. The incapacity to distinguish between the relevant, urgent, or unnecessary knowledge makes the possibility for shifting foci an impossibility.

As Polanyi (1958a: 342) would also note the complication of merely identifying objects, or stimuli, without previously knowing somehow what it was they were identifying:

> Physical and chemical knowledge can form part of biology only in its bearing *on previously established biological shapes and functions*: a complete physical and chemical topography would tell us nothing about [a frog] *as a frog*, unless we knew it previously as a frog.

We find that humans' most distinguishable act is that of understanding knowledge, paradoxically *without the need to previously understand the knowledge in question*. This seeming unparalleled talent allows for the the understanding of stimuli, without the need to have previously been told what it is we should understand from them. Our learning process, given by our innate logic structure (which is duplicable) and our unique network of mind and memory (which is not duplicable) precludes the possibility of ever having an individual gain understanding of what exactly it means to understand. The mind, complicated as it is, has been attempted to been mapped before us in a theoretical manner. The fact of the matter is, however, that this rough estimation of *what* the process is is very much different than *how* the process works. Indeed, as Polanyi and Prosch (1975: 62) remind us, "[w]e cannot spell this process out in explicit steps, and it is for this reason, as we have noted, that no 'thinking' machine can ever be adequate as a substitute, or even as a model, for the human mind."[163] Maturana and Varela (1980) argue that the mind cannot undergo passive changes to external stimuli; they are affected by the

[163] Much like Steele (2002: 137) reminds us, the brain is a *biologically* evolving instrument of an adaptive system. As the brain comprises the framework within which the mind functions, we see the difficulty in mechanically predicting its growth and advance. Consider, could we predict the growth and advance of a human's other physical characters in advance (i.e., height, weight, tone of their voice, speed of their speech, sensitivity to taste, etc)?

endogenous workings of the mind. The cause of changes in the mind cannot be sought by searching solely in the external realm, as is the case with any product of the mind (Hayek 1952: 122-127). Instead, the mind continually *reprograms* itself as novelties are experiences through its environment, but which are *known* as novelties due to the internal structure that comprises the mind.

3. Information

Up to this point, we have made much use of the concept of "stimuli" with little regard for identifying what it is exactly that is meant by this concept. What has heretofore been called stimuli will appear, in most other instances, as what we consider "information." It appears, then, that one focal point that was reviewed earlier was the issue of the indeterminateness of the relaying of a memory from one individual to another. The issue arises as the memory in question that is being transmitted is potentially significantly different in character than the stimuli (information) which is the basis of it.

Information we may define as being the objective facts and events that exist in the external world. In this way, it becomes clear that any stimuli subjected to our senses – touch, sound, smell, taste, or sight – are defined as information. They objectively exist in the world, and are available to be introduced to our minds. Remember, however, that this need not be a direct process. X-rays, for instance, are very much to be considered information, yet they are directly invisible to all of our senses. It is only through a peripheral method that we can obtain knowledge of this stimuli. It is common to all stimuli that the final method with which they become memories in our mind will be of our senses.

It might then be brought to our attention that perhaps information should be viewed as having a dual origin to us. Logic, as we looked at earlier, also bears many characteristics of information. Logic is also an objective reality; it knows no shades of existence. There is one crucial difference between logic and information that must be brought to light that will make the distinction between the two more clear. Logic is that which we already possess within us. It is objective not only in message (for that part if immutable), but also in extent to which we possess it. There is no concept of having more or less logic available; it is binary – either you have it, or you don't. Information is very much different in character for there are an infinite number of degrees of information available in the external world. Furthermore, the amount of information is in a constant state of flux. The continual process of our own actions, let alone that of other objects, continually creates and changes the information available.

Four important points concerning information become apparent: 1) information exists external of us, 2) information is finite, 3) information is waiting to be discovered, and 4) information can be compromised, hidden, perverted, destroyed, or exposed. We can then see that information is a fundamentally controllable element. However, this is only so if it is in our knowledge set, that is to say, we know it exists.

This external source of stimuli is wildly different than our internal source of logic. In fact, while this logic may be viewed as a distinctly human quality, there is nothing inherently human about information. We have as much access to it as any other conscience being. The fact that humans have, throughout history, been able to develop continually by using the available logic while other species have not points to the fact that a deeper quality separates us from other beings. The distinction cannot be found in information, instead, it manifests through logic.

4. The Absolute Trinity – Logic, Mind, and Information

We have seen that prior to any other occurrence three simultaneous entities exist in our world. We are born with only two capacities with which to grow – our logic and mind. Logic represents that tool we use to interpret that information which the mind allows to enter its realm. The mind is a framework which defines what stimuli are subject to our logic, and saves such past allowances through the network of memory. Information are those stimuli that exist in an objective form in the external world to us. Through the mind they gain admittance to the individual, and become subject to their logic and subjective interpretations.

It should be clear, nonetheless, that this trinity will enjoy a unique existence for only the briefest of moments for the individual. As was noted earlier, the moment of choice commences a remarkable series of events which continue throughout the individual's life, stopping only with death. One of the most important of these events, one which has been briefly touched upon, is the creation of knowledge. Until now this point has been used under the guise of "memory." While it is easy to consider all knowledge that we have as memories, more conventional usage of the term insists that we allow this nomenclature pass, and adopt a more modern approach. The problem with using the term *memory*, is that that which we wish to refer to as *memory*, is sometimes not *remembered*. Knowledge holds no such connotation with being active in our mind, although it must always be contained within it. It is only through a deeper look into knowledge that we may see where further complications of choice and action lie.

III. THE SUBJECTIVE DUALITY

The last chapter viewed the world in a somewhat objective way. Information exists in an objective state external to us. Logic is an absolute concept that leaves no room for differing magnitude of availability, shades, or existences. We have pointed to the indeterminateness of outcomes that result from the phenomenological process that changes stimuli (information) into memory (knowledge). Two very related topics, one stemming from the other, result from this subjective process.

The first is the topic of knowledge. When dealing with information, there is no way to distinguish between one quantity of information over another. Information, it is hoped has been made clear, exists in an absolute state, and in *a static sense*, can not have any differing extents. At moment *y* there is an *m* amount of information existing in the world. At a differing moment *z*, there is an *n* amount of information potentially in existence. Things, as we will see, change drastically with the concept of knowledge, which can exist in a form solely in the individual's mind. As such, it is plain to see that this existence must necessarily be limited, and hence, we see that the total body of knowledge must necessarily be smaller than the body of information it aims to represent.

Logic was viewed earlier as an absolute. However, we can see that it is limited in a sense by the amount of knowledge that is available to apply this logic to. With the new realization that knowledge is limited to the individual, we find a new limitation on the amount of logic at our disposal. However, it has already been shown that logic can have no differing degrees of existence; it exists or it does not. What is needed is a new concept to apply to this state of being, the state of being that defines our human lives of incomplete knowledge. The concept that we search for is *rationality.* Viewing rationality as the means with which we apply our logic, we see that rationality is also a concept endowed in all. This makes it not a meaningless concept, however, as we can easily see the reason that this is so by viewing the differing extents of personal knowledge that act to bound this rationality.

1. Knowledge

The Creation of Knowledge

A great amount of confusion exists between the concepts of information and knowledge. Typically the two words have been used interchangeably, however, in reality nothing could be further from the truth.[164] A large divide separates the two, with serious consequences regarding the role they play in the market process.

If, as we have already stated, information is the available body of facts that can become known, then knowledge, in contrast, is the information that we personally possess. Polanyi (1958b: 35) differentiates between the two stating "invention ... does not produce something that was not there before; but actually, it is only the knowledge of the invention that is new, its possibility [or information] was there before." To Polanyi the distinction is slight, but critical. The new knowledge given was only the collaboration of information that was already in existence.[165]

We find that information will never be absolutely private to us, the process of creating knowledge is the method we use to privatize information.[166] We do so by subjectively interpreting new information through the use of our existing knowledge; the use of our logic and mental structures.[167] As Hoppe (1997: 58) delineates the learning process: "Knowledge is always the knowledge *of* something: the knowledge of hands and fingers, for instance, and it cannot possibly be conceived of as anything but sequentially (in time) *acquired* knowledge (as something based upon and learned about some logically and temporarily prior facts)." The result is a fundamentally unique, and personal, knowledge which only the holder is in a position to fully comprehend.

The way in which information becomes knowledge is crucial to the discussion. When we assimilate a piece of knowledge, we do so through our own personal senses, our logic structure and the existing array of networks comprising our mind. This new input is influenced by these three elements.

---

[164] The American Heritage Dictionary of the English Language (2006 ed.), for example, defines information as "[k]nowledge derived from study..." and knowledge as "specific information about something." This circular reasoning in no way to aids the discussion.

[165] See Kirzner (2005: 76) where he states, "[i]nformation is an input that may be used in a process ... that results in the possession of knowledge." See also Sautet (2000: 14fn16) or Mises (1957: 109) where he views inventions to be the product not of something material, but of the mental process reasoning and collaborating information together.

[166] See Hoppe (2007: 14) whereby each human enjoys a privileged access to their personal, and necessarily private, knowledge.

[167] See, for example, Lachmann (1959), (1976a), and (1977) for knowledge being the subjective interpretation of information – the mixture of experience and new information.

The result is that although objective information exists in our world, knowledge may only exist as its subjective embodiment, internalized through our personal experiences. Hence, as Polanyi and Prosch (1975: 61) comment:

> We should now be able to see that all our knowledge is inescapably indeterminate. First of all, as we have seen, the bearing that empirical knowledge has upon reality is un-specifiable. There is nothing in any concept that points *objectively* or automatically to any sort of reality.

Tacit Knowledge

In volume I we distinguished between two types of knowledge: information and action (Kirzner 2005: 77). Action-knowledge is that which propels us to act. In contrast, we can hold information in our minds as knowledge, but it may not necessarily cause us to take any further course of action. Hence, action-knowledge can only be formed through the joining of two or more pieces of previously existing pieces of information-knowledge. An existing piece of information-knowledge requires an additional piece of information to increase its relevance for action. These pieces need not originate of the same person, but the resultant action-knowledge will necessarily always be of a single person.

However, there is also a third type which must be explored. In fact, Polanyi (1958) pointed out that much knowledge is not available to be easily disseminated through the conventional means. Instead, this tacit component remains in an inarticulable form in the human mind, awaiting the proper learning process to make it available to the world.

Much perversion has occurred since Polanyi conception of the tacit component of knowledge. Collins (1992: 56) describes tacit knowledge as “the name given by Michael Polanyi (1958) to our ability to perform skills without being able to articulate how we do them. The standard example is the skill involved in riding a bicycle…” If we accept this definition, it seems that this knowledge rests primarily on an existing, static state. What is more interesting, and more pertinent, is the *process* by which this form of knowledge is learned.

In fact, this distinction between tacit *knowledge* and tacit *knowing* was clearly delineated by Polanyi from the start (see Pozzali 2008). Knowledge was, after all, “best described as a process of knowing,” according to Polanyi (1969a, 132).[168] This shift from viewing tacit knowledge as a process

[168] Earlier evidence of this distinction may be found in Polanyi (1967, 7) where he remarks: “I shall always speak of ‘knowing’, therefore, to cover both practical and theoretical knowledge. Practical knowledge largely remains in the tacit

involving learning, to a state requiring understanding has had the consequence that the manner with which we acquire this specific type of knowledge is overlooked, or diminished to a footnote.

Collins (2001) attempts at rectifying this discrepancy with his five-fold categorization of tacit knowledge. The reason behind this categorization was not to "deepen our understanding at a philosophical level but to explicate the idea clearly and draw out its implication for scientific practice" (*ibid:* 71). Hence, by more accurately defining different *types* of this tacit dimension, greater knowledge of its use it hoped to be gleaned. These five categories are: 1. concealed knowledge, 2. mismatched salience, 3. ostensive knowledge, 4. unrecognized knowledge, and 5. uncognized/uncognizable knowledge. Unfortunately, almost all tacit knowledge able to be conceptualized by a third party belongs to Collins' third group – ostensive knowledge. This category treats the knowledge as an alternative to 'skill-like' knowledge (Pozzali 2008: 230). The true problem with this classificatory system is, however, that is fails to make the distinction between the *known*, and the *knowing*, a point which had already been well developed by Polanyi. The focus on the product instead of the process breeds further detrimental results for the study of tacit knowledge that we wish to focus on.

Within the framework of modern knowledge theory there exists a trichotomy of knowledge types which may provide a sound basis for a proper classification of this tacit component.

Knowledge based on true (and justified) beliefs provides propositions that are recognized as being true. It is well known that the Pythagorean theorem – $c^2 = a^2 + b^2$ – is true, and that we consider it to be true. This is because we know why it is correct; it is an easily deducible truism. Not all knowledge based on justified beliefs need be this easy to prove, although many certainly can be.

Knowledge as a competence is that which pertains to the performance of a certain task. This may explain an unconscious instinct (i.e., breathing), or through lengthy, *learned* apprenticeship. In other words, it demarcates the line between know-how and know-that.[169]

Last, we may have knowledge as acquaintance. This knowledge through experience is empirically learned and remains in our memory through continual strong linkages within our memories, as has previously been developed.

realm. See also Polanyi (1969b, 140): "This act of integration [of focal and subsidiary awareness]… I shall call …. tacit knowing." Hence, we see that tacit knowledge is very much a residual to Polanyi's original work which instead focuses on tacit *knowing* (Pozzali 2008: 231).

[169] Knowing how, which corresponds to tacit knowledge, consists in using habits which need not (indeed, can not) be explained by the individual. Conversely, knowing that (i.e., conscious knowledge), pertains to decision making. It describes the knowledge used in the formation of strategies and the objectives of the action. See Hayek (1952b: 39; 1962: 43).

With this pre-existing trichotomy, where is it that tacit knowledge may best fit in? Most literature concerning tacit knowledge would be inclined to place this element in the 'acquaintance' or 'competence' categories. However, we find that what is needed is rather a distinction between tacit knowledge and knowing. The manner in which each category of knowledge is learned becomes central to its character.

Knowledge based on beliefs cannot be learned through *mere* social interaction that tacit knowledge requires. Instead, an *understanding* of its internal logic system must be formally acquired. An understudy could sit by a great mathematician's side every day for years, but unless the logic of the deductive reasoning unfolding before their eyes was learned, the knowledge would unravel before blind eyes. Knowledge as competence can be either formally, or tacitly, learned. There is no question that riding a bicycle is a special form of knowledge which must be learned without it being *fully* articulated to us. However, certain *elements* of this process can be easily dictated that will augment our tacit knowledge as to how to best apply this knowledge.[170] Knowledge as acquaintance completes our spectrum. This may only be tacitly held, *as well as learned*. As an example, many texts could explain that while riding a bicycle at speed, to make a successful turn the front tire must be turned in an opposite direction to that which the rider is traveling. However, this could not become a knowledge as acquaintance (or second-nature in other words) without a rider repeatedly attempting this action to see that it does, in fact, hold true. It is this exact process which Polanyi emphasized when first outlining tacit *knowing*. Language is an example, commonly given by Chomsky (see 1976: 161) whereby social interaction is necessary to learn the knowledge. The rules of grammar and all definitions may be read in a text, but without the interaction between individuals, the *true* meaning of language cannot be known.

Tacit knowledge, as has been shown, makes up one end of an extreme spectrum of both knowledge and knowing. It is the antithesis of deductive logic, and must be learned through direct action and not through mental theorizing. Imperative to a thorough understanding of this tacit component is, however, knowledge of its inherent limitations.

Limits to Our Knowledge

The use of knowledge is of the greatest concern to the individual. It would seem instructional to assess

[170] For example, telling a bicyclist that a low center of gravity will aid in their stability will add to the indemonstrable knowledge of how to keep balance. Even further, it can be deductively shown that standing on the pedals will lower the center of gravity, despite raising the physical body – a counterintuitive act which may be difficult to learn through self-taught or observed methods.

what some of the limits on obtaining knowledge are. Essentially there are three types of knowledge: information knowledge, action knowledge, and tacit. Fundamentally though, there is no difference between information knowledge and action knowledge. They both involve the same manner of acquisition, the process by which all objective information becomes subjective knowledge. However, tacit knowledge involves an entirely different transmission method.

As we have seen tacit knowledge is that which we cannot articulate. It must be transmitted through a form different from more conventional ways – reading or listening – as examples of ways which more conventional knowledge can be transmitted. Tacit knowledge can be transmitted then in one of two main ways. The first is through personal contact with an individual. A great painter cannot write a text book and expect their student to learn how to paint by reading it (although some specific concepts of painting can be shared in this way). Instead they must sit down, and teach through interaction to transmit their knowledge of what makes for great painting. This type of transmission is admittedly very limited. There are only so many interpersonal interactions that can occur, and the time necessary to get the full import of the knowledge across may be prohibitively long. A great painter may *never* be able to teach another what it is that makes for a great painting; it could be a type of knowledge that is *impossible* to transmit.

The second method through which tacit knowledge may be spread is through the pricing system, as stressed by Hayek. Assume an individual wants to purchase a share of a company. They may be fully unable to articulate to anyone why it is they wish this end, it could just be a gut feeling on their part. This reason is, to some extent, not necessary to be known to the other market participants. For once the individual actually purchases their share, the increased demand will place upward pressure on the price. The resulting increase in price (assuming fixed supply) will relay knowledge to others who have access to this pricing system that there has been a factor somewhere that has increased the price of the share. However, it is important to stress that by viewing the result (the price increase) of the use of the tacit knowledge in question (the gut feeling), an individual is told absolutely nothing about the cause of the price movement in question. They are given *peripheral* knowledge of the activity, but are still unable to learn the underlying element.

Secondary limits to our knowledge can be found in the very structure of our mind. For even if the knowledge reaches our mind as a stimuli, it is still not guaranteed that it will be internalized as a memory. This is so due to the limited attention (or consciousness as was previously stressed) that an individual may have at their disposal at any given time. As consciousness is not an endless resource, that is to say, we must economize on it accordingly, it may be that an individual is unable to focus attention on the knowledge, *even though it is presented clearly before them*. Second is that the network of memory may be ill-structured to bring importance to the novel knowledge, thus also limiting the amount of attention that will be heeded to its obtainment. If the path dependency is insufficiently strong, the knowledge already contained within the individual's mind will not be strong enough to be aware of the necessity or importance of the stimulus that is presenting itself for consideration.

These limits make it obvious that knowledge, although not necessarily having a price-tag associated with it, can, in many cases, be quite costly. There are several crucial elements which must come together in proper sequence for information to be gained, and internalized as personal knowledge. The process with which this happens is as equally important as the possibility that it should actually occur. For what good would it do for us to make the assumption that *no* information is ever available to become part of our personal knowledge (despite obvious evidence to the contrary)? This complicated knowledge component provides a limit on the actions that we can undertake, and hence, places a constraint on action which cannot be conventionally quantified.

Appendix A: Imagination

Imagination is a source of knowledge which can only exist in our minds; it is endogenous to us. It is wholly created free of external (or what we would typically define as informational) elements, although it can be influenced by them. Imagination and information share some special qualities, linking them in some ways, but creating a stark divide in others.

A revisit to information will prove instructive. Information exists, but we do not know it to exist. Once information becomes known to us, it becomes a type of knowledge. Therefore, although it exists, we can never hold information, only knowledge of it. Likewise, imagination exists, but only for a fleeting moment. For once an imagined thought becomes known, it is part of our knowledge. Hence, both sources of information can only exist for a fleeting moment. Much of G. L. S. Shackle's work focused on the importance of imagination to economic activity, and as he (1994: 1) would define the topic:

> Imagination is the gift of mind whose work must precede that of any other. It is this gift, drawing suggestions from the environing scene and from stored knowledge, which provides

the entities of reason and experience can employ to make plans of action and of enterprise. Enterprise indeed is imagination actualized.

The discovery of knowledge is limited by the extent of information available. Although this is a massive amount, it still represents a physical limit. Imagination, on the other hand, has no physical limit. It can contribute to knowledge unendingly, forever expanding on the possibilities for further knowledge. The only limit to our imaginations is our imaginations.

2. Rationality

The preceding discussion should have made clear the indeterminate nature of knowledge. The subjective interpretation of all knowledge implies that all decisions that are undertaken based on this element will be conditioned separately than they would be for another. However, due to not only this interpersonal inconsistency regarding knowledge, but also the inconsistency in what external information will become knowledge, we can see that the limits to having a homogeneous knowledge set with another are quite high. In fact, we can state that, due to these two factors, having an identical general knowledge set as another becomes an impossibility.

What is it exactly that we mean by using a term such as rationality? Shackle (1972: 246-247) points out that most definitions of rationality have a paradoxical element to them. The trouble is that rationality is commonly viewed as being able to choose against things that are fully known, however, given the passage of time and our continual knowledge expansion, we see the impossibility of any element coming to be regarded as "fully known." Hence, social scientists viewing rationality and knowledge in this manner are faced with a "stark choice," that is they can "reject either time or rationality" (Shackle 1972: xi).

Simon points out that there seems to be a trend toward defining rational action very narrowly, and irrational action in quite broad terms. However, this dichotomy is unsatisfactory, for, "[i]f there are differences in viewpoint, they must lie in conceptions of what constitutes rationality, not in the fact of rationality itself" (Simon 1986: S210). Rationality is thus an absolute concept, it is only our own interpretation of it that differs. We know that humans are rational, at least to some degree, for if they were not, to ask the question would have no meaning (Grampp 1948). Once we have established, or conceded, that humans are rational in some manner, two ancillary questions arise. First, in what regard are actions rational regarding the ends sought. Second, to what degree are the changing values of means evaluated under rationality, and in what derivative ways are they appraised, and changed when necessary?

Viewing the question within this shifting means-ends framework, in fact, seems to be the crux of the issue. Mises (1949: 17-19) demonstrates that the ends of action are necessarily "ultimate givens." They are tempered by the unique wants and desires of each individual, which are forever unknowable to outside observers. Given this input, we see that rationality must be of concern only to the means of action.[171]

However, if we are then to focus on rationality regarding means, we must also understand the subjective nature of the concept. For the means will forever be subject to our incomplete and changing knowledge, and hence, as we search for a definition of rationality, it must be grounded in this unmistakable personal aspect – the uniqueness of knowledge. In this way, we can find full agreement with Mises (1949: 19) that action must always be rational from the viewpoint of the individual; to call action "rational" is pleonastic.

What happens of an individual's view of their own rationales with the passage of time? Changing factors will almost guarantee that sought ends will shift, means will be altered, and value scales will re-order themselves. Baudin (1954) clarifies this example, by showing that the end of hunger may be satiated by seeking a loaf of bread to eat. The goal of hunger removal does not vanish when the individual leaves his house to buy the loaf, nor when he actually purchases it. Furthermore, what would happen if, after purchase, the loaf was stolen prior to being eaten? The ultimate satisfaction in all cases is not satisfied, and hence, all actions directed toward it will continue to be rational until the ultimate end is finally gained.

It should become clear that, while each action is individually rational, this is so due to the differing degrees of personal knowledge each could possess. For what good would it be for one individual to pass judgment on another's rationality if their respective knowledge levels are distinct? Rationality is, then, a concept applicable only within the bounds of an individual's knowledge set.

Bounded Rationality

Simon (1957) proposed that actors are fully rational, but within a constraining sphere. His theory of "bounded rationality" posits that individuals do not reach fully rational decisions, however, they reach rational decisions based upon the limited knowledge they may have concerning an event. As Kirzner (1973: 159) points out, choices are made not only on what is known, as some options exist that will not be known (see also Hülsmann 2003). Hence, full-rationality, at least in a neo-classical sense, is impossible to achieve due to the incomplete information available for any decision.[172] This comes

[171] As action is always that of the individual, we see Lachmann's (1976b: 131) definition of rational action as a movement towards "*individual equilibrium*," to be appropriate in this regard. As individuals act in direct concern of their own want satisfaction, we can see that rationality will always be directed toward this concept of a forever shifting individual equilibrium.

[172] Shackle (1965: 12), for example, takes the stance that rational choice is the outcome of "a perfect knowledge of all circumstances which will affect its relevant outcome." Any choice made in the absence of this full and perfect

mainly as the result that:

> [T]he capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behavior in the real world – or even for a reasonable approximation to such an objective rationality. (Simon 1957: 198)

For Simon, the main limitations this incomplete knowledge may serve can manifest in two primary ways: 1) through the limits on the utility functions an individual can formulate, and 2) an inability to properly assess the costs of gathering and using information.[173] Simon (1978) viewed the more orthodox economic opinions towards rationality as constraining and detrimental for the science. The preoccupation with results rather than choices had neglected the larger issues of decision-making. For example, by focusing on the a small number of discrete situations, the larger concerns are no longer important. The question to be asked is not necessarily "how much flood insurance to buy?" but "what conditions make flood insurance attractive or rational?"

Under bounded rationality, a decision is achieved in light of the fact that an individual may not be able to assess all the possibilities in an environment. Therefore they cannot make an optimal decision in light of all the facts, hence, in Simon's view, they maximize the part of knowledge they do know, or "satisfice" upon it. As Sautet (2000: 29) summarizes, "[t]he implication of bounded rationality is that individuals will economize on their (bounded) rationality in their decision processes on the one hand, and setting up corporate governance structures on the other hand." Hence, Sautet views the limitations that bind us as being responsible for the creation of social structure in society (particularly firms), instead of allowing all processes to be performed through market processes. Hence, rationality becomes a scarce resource. This is true, in the sense that knowledge surrounding rationality is scarce. We all have access to all of our supply of it at any time if we wish, but it can be developed and expanded over time, and hence, used more efficiently and effectively.[174]

---

knowledge set becomes part of an irrational realm, or an arbitrary choice. We can see that this definition which Shackle wishes to use for rationality is actually that which we use for logic. A failure to have full information of a situation does not lead to arbitrary choice (the complete opposite of rational action), instead, we see that it leads to a rational choice within the confines of this limited information set.

173 Knight (1921), in distinction, sees true uncertainty and novelty as constraining the entrepreneur, for Simon, we see that complexity in the decision making process constrains their ability (see Sautet 2000: 51n64).

174 As Mises (1920: 102) viewed economic calculation as the method we economize, or satisfice, on this complexity. "[N]o single man can ever master all the possibilities of production, innumerable as they are, as to be in a position to make straight-away evident judgments of value without the aid of some system of calculation [i.e., economic calculation]."

We see that what Simon refers to as bounded rationality is really just an exercise in logic. Rationality always must assume that we know only what we know. Dequech (2000, 171) views rationality and knowledge as inextricably linked; rational is that action based upon knowledge pertaining to that action. All action is rational as a result, from the perspective of the actor explicitly, and from the perspective of the observer implicitly. It is, in fact, only this actor who can determine if an action is rational or not, and by the act of performing the action, we know it must be rational in their eyes. This view of self-rationality may only apply to the moment at time in which the action was undertaken, for as we have seen, the passage of time shifts the knowledge base and hence, creates a new boundary for individual rationality. Logic, on the other hand, concerns all the options available. In an absolute manner, we see that logic is necessarily always limited by the amount of knowledge an actor has at their disposal. As it is absolute, an outsider has the ability to state if an action is logical or not, as they may have access to subsidiary information that is essential for the decision making process. Simon (1978: 14) posits that rationality can be increasing over time, due to the increased technology or problem-solving techniques available. However, our logic base is constant, it is only the increase or decrease in knowledge (both quantitatively and qualitatively) that can effect its use.[175]

In fact, as Williamson (1981: 553) points out, there are four key areas that individuals are bounded by: receiving, storing, retrieving, and transmitting of knowledge. These four points are what work in a concerted effort to bound the pure logic that we have in our minds, with the result that an individual's mind becomes bounded in its rationality. Much of these four points has already seen considerable work based on Hayek (1952b). For instance, the trouble of retrieving knowledge is a problem manifested through the path dependencies that exist in the mind at any given time. The transmission of knowledge becomes of dubious quality given Polanyi's (1967) stress on tacit knowledge – that which is inarticulable and hence, unable to be easily or faithfully relayed to other individuals. Storing knowledge becomes a problem of the mind's ability to restructure its path dependencies to welcome new knowledge. Remember that as paths became more solidified, or reinforced, through agreeable pieces of corresponding knowledge, it becomes more difficult to accept fresh ideas. Lastly, we can see that it is only the problem of receiving information that receives much

See also Mises (1949: 698).

175 It is somewhat ironic, or unfortunate, that the definition of bounded rationality has slowly been perverted within the profession. From its beginnings as limited knowledge bounding an absolute logic set, we find more intrusive definitions today. For example, Arthur (1991) informs us that "[m]ost economists accept that there are limits to the reasoning abilities of human beings – that human rationality is bounded." However, the focus should be on the incomplete *knowledge* set bounding this knowledge, not necessarily on an incomplete *logic* set that operates on this knowledge. See also Arthur (1994) or Lipman (1999) for similar conclusions.

concern in modern economic analysis. Problems concerning the distribution of information take a central role, at the neglect of the more subjective, or hidden, problems that occur inside the mind concerning the storing and faithful transmission of this same knowledge.

As Conlisk (1996) sums up, there are four reasons why we should begin viewing action in terms of bounded rationality. First are the abundant empirical models that incorporate the concept correctly. Second is the broad range of applications available to be reassessed in light of this constraint. Third is that the alternative, *un*bounded rationality (i.e., logic), is an unconvincing assumption. Last, we find that economics searches to entertain all costs of decision-making. Why should the cost of our own rationality be any different? Hence, in four words, bounded rationality becomes the framework to use for decision-making due to: "evidence, success, methodology and scarcity."

Knowledge and Action

As we have determined the manner with which a specific knowledge level defines the rationality that may be a applied to a given situation, the next necessary step is to show how it is that out knowledge becomes transformed into action. The Hayekian distinction between "know how" and know that" becomes important when we view the process with which plan formation takes place.

"*Knowing that*" represents the body of knowledge that an individual not only has, but is also able to articulate and hence act upon with some degree of definite direction in a catallactic setting. In contrast, "*knowing how*" – that tacit element of our knowledge set – remains in the mind of the individual, unable to be directly articulated and disseminated throughout the market. In a fully autarkic market, one defined by an individual solely responsible for fulfilling their own needs without exchange with others, it becomes clear that tacit knowledge becomes a moot point. As there are no other individuals with which to exchange, tacit knowledge poses no significant problem to the individual. As the development of the market ultimately depends on increasing levels of knowledge dispersal – a continued refinement of the division of knowledge – a market defined by one person would have no issue with the process with which this tacit knowledge is transmitted throughout the market; the knowledge is already spread throughout the entire market within the mind of the individual.

Plan coordination in a collaborative setting (i.e., catallactic) becomes much more difficult when presented with tacit knowledge. However, it may also be seen that this special knowledge of "knowing how" is instrumental for the sustained development of the market. Tacit knowledge is not something which is created by individuals consciously and added to their knowledge realm. Instead, it is a natural creation; a product of the limitations the mind has in elaborating the processes through which it undertakes. The transmission of this tacit component must be transferred to other individuals so that plan co-ordination may be increased, with a corresponding improvement in the general want satisfaction of individuals.

It should at this point become increasingly clear that what is required to transmit "know how" is a clear understanding of its existence – we must "know that" a "know how" exists. This remains true for both the teacher and learner of tacit knowledge. For instance, if one wishes to learn how to ride a bicycle successfully, before they can begin to attempt at learning the implicit requisites of this task, they must gain knowledge of the more concrete concepts such as what a bicycle is, or that one is available to be ridden, for example. A bicycle may not exist if none knew that it could be ridden – a task which may only be learned through a tacit learning process. The knowledge that such tacit knowledge exists *is* something which may be transmitted through concrete actions delivered by other individuals. The awareness of these tacit possibilities becomes a primary role to be undertaken for the development and cohesiveness of a market.

Important as tacit knowledge is *viz* action, plans can only be constructed under the pretense of knowing that it exists. As Thierry (2008: 30) makes clear:

> Thus, sensory perceptions are not spontaneously exploitable to formulate our plans because they are not immediately translated in the conscious. While certain knowledge used in carrying out a plan may be non-explicit and unexplainable, the plan itself must necessarily be explicit.

Hence, in order to create a plan that results in action, the knowledge that we are articulating the plan with must exist in a defined, concrete form. This is not to say, however, that tacit knowledge may not be used in the plan. A student may have formulated a plan that they will travel to school on their bicycle. That the knowledge they have concerning the riding of the bicycle is necessarily tacit in nature does not mean that the planned action can never be undertaken. It does mean that the knowledge that the student has concerning their own ability to use a bicycle must exist in a concrete form. This form will exist to them in a form which makes it possible to use for plan formation, and hence, will also exist in a form that is transferable to others.

As such, action necessarily results from the planned co-ordination of distinct pieces of knowledge that one or more individuals may have access to. It also follows that action must be a conscious consequence of combinations of knowledge. These combinations must be explicitly formed in the mind of the acting individual, and hence, must also be able to be explicitly relayed to others. Tacit knowledge thus finds its delivery method through actions which are formed with a concrete knowledge that such tacit knowledge exists. For, if we consider for a moment, the lack of knowledge that tacit knowledge exists, we see that there is a lack in the method with which it may be transferred.

Suppose an individual has opened a school to teach others how to ride a bicycle. The teacher must necessarily have knowledge of their own tacit ability to ride a bicycle. For if they did not, how would they have the necessary knowledge to open a school to show others how to do so? It could be, however, that they hire others who do have knowledge of how to ride a bicycle to teach others this act. However, two pieces of concrete knowledge must result first before this is possible. First is that the owner *must* have knowledge that the tacit knowledge regarding the bicycle riding process exists. Second is that the teachers hired must have knowledge that they personally have the necessary knowledge to ride a bicycle, and be able to signal this to the school-owner. The signaling process must be conveyed explicitly through knowledge which is able to transmit the tacit component's existence.

If all action must be directed in a conscious and purposeful way, can there exist a different set of "actions" which may be directed in a non-purposeful manner? Not according to the definition that we have given to action, which is the necessary result of coordinated pieces of knowledge. There are similar movements individuals may undertake which, although by all appearances being quite similar to action, fail to include its most vital component – purpose.

Action and Reaction

Action must necessarily be directed purposefully with a clear goal sought. It may be easy to think of the opposite of this concept of directed action as random behavior; this would be erroneous however. Even *seemingly* random actions cannot be random in the true sense of the word we wish to ascribe to it. For even an individual attempting to commit a random act would still be acting in accordance with two things. First, they would be acting within the confines of their knowledge set. Action would still be tempered to the extent that an individual has knowledge of their world, and by the alternative actions that could be undertaken at that moment. Second is that they would still be acting perfectly cohesively within their ranked value scale. For instance, an individual would only walk in a zig-zagged pattern rather than a straight one, if that action had an expected greater value than its alternatives. The conundrum that develops is that no action could be random, provided that it occurs given these two caveats.

Where then might the true corollary to action lie then? *Reaction* provides the antithesis of action. If the first of the above caveats becomes disrupted, a motion undertaken stops being a purposeful action and becomes a reaction. Consider that a motion undertaken by an individual with no knowledge of the means used could not *by definition* be purposeful. This is so as no linkage could be made between the means to be used, and the goal sought. In some instances, however, we may find that a motion occurs under these instances, regardless of the fact that the individual can make no causal connection between the means and expected end. If a hand touches a hot stove accidentally, or without prior knowledge that it will be experiencing a searing heat, a motion will result which cannot be said to have been based on a premeditated thought. For once the heat is touched, the hand will *react* and pull itself away from the source of the heat. Now, it could be argued that it is known that the heat is the source of discomfort, and that to move away from this known source would be an entirely rational action in response to a felt discomfort. This criticism misses the mark however. For, if we consider the first time that a hand is exposed to the searing heat of a stove accidentally, there can be no basis of prior knowledge for which any purposeful reply to this felt uneasiness to be undertaken. Additionally, it could have been that the individual had been told prior to the incident that if they touch a hot stove, they will feel an intense heat, and that the action that would rightly remove this discomfort would be a removal of the hand. This response would find itself relying on the type of knowledge that this warning embodies. If it takes the form of tacit knowledge, than there can be little way that the individual can know exactly what the heat and discomfort they are about to react to *prior to touching the heat and feeling the discomfort*. A reaction is born.

A distinction between the physical and social realms becomes evident. To use our above example, the heat of a stove is an objectively definable measure. A stove top of one hundred degrees centigrade today will be identical to on emitting one hundred degrees centigrade in the future. Even under the possibility that the nature of the heat on an individual's hand is tacit knowledge – only to be learned through direct experience – we may see that the learning process may only occur one time before it is ingrained in the individual's mind forever. The physical world, marked with constancy allows this knowledge to be learned and directly applicable to instances separated temporally to achieve the same effect. In the social realm, marked by human interactions, what are the effects that may manifest from this process?

In a social realm lacking constancy, we find a source of fresh novelty through our interactions. We may experience emotions through our interactions with others, and add these to our knowledge set, but there can be no way that we can say that each time we feel the same emotion, for example, it has any sort of sameness to categorically similar ones experienced previously. How many times can different women tell a young man that they love him and yet the subject in question may be driven to categorically define each instance differently? Indeed, each instance whose existence remains solely in a purely mental state can provide a fresh source of novel knowledge to an individual, quite unique from others felt prior.[176]

Reaction actually results in a manner clearly delineated through Hayek's connectionist theory of the mind. Up to this point, we have considered only the first of two preconditions for rational action to occur – the reliance on logic constrained within an individual's knowledge set. This step, although of the utmost importance, is only a necessary but not sufficient criteria for a reaction's occurrence. It is, however, instrumental to the second stage's occurrence – the disjointing of a means from the existing value scale of the individual.

Hayek showed that knowledge exists as memories interconnected in the mind through a web of path dependencies. The *strength* of knowledge, that is, its ability to be recalled and the subjective importance placed upon it, is reliant not on its absolute existence, but only that in relation to other pieces of knowledge within the mental system. The knowledge that we have been talking about as a causal condition for reaction to occur is necessarily *new* to the individual. Its existence is marked by a lack of connections within the web of existing knowledge. As a result, it lacks a connectivity with the ends that it can be a means to create. Action takes place as a result of the active pursuit of the ends most highly valued on an individual's value scale. As such, there is no way that this new knowledge could be acted upon in any conscious manner. It lacks the connections to the causal source of action – the value scale.

[176] Take an example of a son who just discovers the untimely death of his mother and falls to his knees in anguish as a result. Would any forward that this action of falling was anything less than a reactive response to the news of his mother's passing? In fact, despite having experienced death of loved ones prior, including the associated pains and tribulations that are entailed with them, this particular instance may elicit a reaction not only wholly different than in other similar instances that the son has experienced, but the reaction may be wholly out of his control. Further, would anyone forward that a grief stricken individual who is crying uncontrollably is making a conscious decision to do so? Can we not consider that an event has occurred far beyond the possible control of the situation and a reactionary response, far beyond the reasoning of the individual has resulted.

There is little that we may say about reactions. As fresh knowledge enticed novel reactions there is no way to compare the reactions of one moment to that of another. The distinction between action and reaction, it is hoped has been shown, creates a limitation on any theory based upon an assumption of purposeful behavior. Luckily, there are a significant number of things that may be said about actions *a priori*. It remains important, however, to forever keep the distinction between action and reaction, and to realize the limitations that each places upon the other.

Rational Action

The presumed assumption that action is wholly rational is not an assumption, but has been demonstrated in the very way with which we formulate our actions. It turns into a tautology whereby action could be nothing other than fully rational. Action results as we seek our most highly valued ends given the means that we have at our disposal. Of overriding concern when reviewing these means is the knowledge set available to be used. The implications of this knowledge are three-fold.

First is that the ends we seek will necessarily be limited by the knowledge set we have at any given time. Only opportunities known to us can be valued accordingly.

Second is that perceived opportunities may only be ranked relative to the means available to obtain them. I may have knowledge that the Hope diamond would make a very nice birthday gift for someone, but the knowledge I have of my own means necessarily places the relative value of obtaining this end quite low on my value scale. It is not enough to only be limited in means to attain ends. It must also be *known* that an individual has limited means for attaining an end. Action, even that directed towards highly desired ends, will be withheld if a binding resource constraint is expected to be met inhibiting the final attainment of said end.

Last, we see that rationality is ultimately limited by the degree of knowledge that is available in the individual's mind. Logic, we have seen, provides an absolute method to determine outcomes given finite means. However, the issue that has been raised is the assumption implicit with the conception of logic is that the knowledge set an individual has is necessarily complete – in another word, perfect. In reality, rationality is always necessarily fully applied to any given situation, but always with the caveat that it is constrained by the knowledge set it is, in turn, being used to act upon.

Knowledge and rationality provide two ultimately subjective elements within each actor. The two also become inextricably linked as rationality results from the reality that knowledge will forever exist in an incomplete form. Rationality and knowledge both stem from seemingly complete beginnings – logic and information – however, given the mental limitations of individuals, action must always be undertaken within a constrained framework, as has been previously outlined.

Some Rational Consequences of Action

The continual subjective nature of two inter-related concepts – knowledge and rationality – gives rise to some interesting conclusions we can make about actions arising from them. One set of conclusions fall under the temporal category of *ex ante*. Those actions will concern the state of affairs existing prior to the realization of an expected event. The second set of conclusions naturally exists as the corollary; those aspects which obtain *ex post* an expected event's occurrence.

Ex Ante Action

Ex ante two aspects come to light for an action to occur. First is that the knowledge of an action to be undertaken at some point in the future must be known. This may seem moot to some, but it is quite imperative to the action process. For, as we have seen, the individual *can* only have a limited amount of knowledge at any given time. Furthermore, owing the connectionist cerebral system, knowledge may be held, but it may exist in a 'weak' form, unable to be utilized. Furthermore, the causal connections of the knowledge may not exist, or may exist in a weak state, delaying or eliminating any chance of action pertaining to this knowledge from occurring. For instance, it could be that a person has knowledge that their stove has caught on fire and that the water is turned off due to municipal work on the pipes. Simultaneously they may also have knowledge that in their fridge is a box of baking soda. However, lacking the knowledge that banking soda is an effective tool for eliminating fire, the individual's house may burn to the ground. Knowledge of the *causal relationship* between the two means is as important as knowledge *of* the two means.

In fact, with this specific example, we are able to see the difference between an absolute logic applicable to the end (eliminating the stove-top fire) and the relative rationality that is employed to combine means to enact an end, as is evident in the reality of *homo agens*. With a knowledge of all available means, as well as their causal relationships when used together, and toward a given end, a system of logic can be used to determine the most *optimal* method to obtain an end. However, as knowledge is necessarily limited to the individual, we see that a system of subjective rationality is used. In this instance, as knowledge of the causal relationships between the available means was limited, rationality directed toward the attainment of the sought end (extinguishing the fire) was also of a limited nature.

As a consequence of action, we can see through the principle of *demonstrated preference* that the action undertaken was the one valued by the individual most highly at the point it was enacted. By employing the means available to attain a sought after end, an individual demonstrates concretely that they have valued this end higher than others *available* to them. However, it has been argued that this principle may not say anything about true preferences, only those which are available through the given means. For instance, a woman may really prefer to have a ring with a larger diamond, but if at the time she purchases the item her monetary means are only sufficient to purchase one carat (and the ring ranks highest on her value scale) that is all she will leave the store with. The fact that she has left the store with a one carat diamond ring on her finger as a sign that she valued this end the greatest is not invalidated by the fact that we may say that she really valued a larger stone. This line of reasoning ignores the subjective opportunity cost of action, as well as the marginal utility of means. If the woman really did value the larger ring more than her monetary capital would allow for, she would search her other available means to find something which may also be employed for its attainment. Suppose, for example, that the woman had the following means, ranked from highest to lowest utility to the individual.

1.	Duplex	$150, 000
2.	Chevrolet car	$30, 000
3.	Sony television	$1, 500
4.	$ 1, 000	$1, 000

For illustration purposes, we may also place the expected cash values that will be obtained if the available goods are sold.[177] Further suppose that a one carat ring has a price of $1,000 and that a two carat ring has a value of $2,500. Now, if the woman valued the one carat ring at more than $1,000 (for instance, at any value greater than or equal to $1,001), she would exchange her cash holdings for the

[177] This is not to imply that the individual values the television 50% more than the $1,000 they are in possession of, or that they value their duplex five times more than their car. It is merely to illustrate the expected cash flows that would be received if they were to renounce that end.

ring. However, it may very well be said that she would really prefer to have a two carat ring more than a one carat. Also, we may say that this is a valid statement. *However*, what becomes evident is not whether she would prefer the two to the one carat ring, but rather she would have the two carat ring to the next foregone alternative. In this case, she could expect to sell her Sony television, and use her cash reserve to come up with the $2,500 necessary to buy the larger diamond. If she instead elects to purchase the smaller stone, she has demonstrated through her preferences that her value scale is ranked as follows:

1. Duplex
2. Chevrolet car
3. Sony television
4. **One carat diamond ring**
5. $1,000

Alternatively, if she demonstrates through her actions that she would rather sell her television, and use the cash proceeds, plus her existing reserve to purchase the larger two carat diamond, her preference scale would be demonstrated to be as follows:

1. Duplex
2. Chevrolet car
3. **Two carat diamond ring**
4. $1, 000
5. Sony television

We may only say which means was preferred after the choice has been made. That is to say, only once we may see through demonstrated preference what means would be renounced for the attainment of an end may we say which means were preferred to the individual more than other *available* means. A person may *say* that they would prefer some end over an available means, but until they actually commit through the renunciation of a costly means nothing may be said about the true preferences of the individual.

It may be countered that this example assumes that there are means available to be used for the attainment of the sought after end. However, what would happen to our example if the means available were insufficient to attain an end? If we now consider the original value scale of the individual, coupled with a new 10 carat diamond that has a price of $200,000, what are the implications for the concept of demonstrated preference?

The woman will be unable to renounce enough of her possessions to have the expectation that she will be able to attain the new, larger diamond. Is it still possible that she prefers to have this larger ring more than any other means she has available? We may phrase the question slightly differently by stating: “Is it possible that her ranked preference scale is really as follows?”:

1.	10 carat diamond ring	$200, 000
2.	Duplex	$150, 000
3.	Chevrolet car	$30, 000
4.	Sony television	$1, 500
5.	$ 1, 000	$1, 000

Before asking if this is possible, we should explore whether there are any useful implications for assuming such a possibility. For it may be possible that we can entirely frame the problem this way, but at the same time lose more than we gain.

First, by assuming some wants to be higher ranked than were demonstrated through an individual's actions, we create a problem of continuation. For if a 10 carat diamond ring is now the most highly preferred end, what about a theoretical 11 carat diamond ring? Or 12 carat ring? The list could go on *ad infinitum*. In fact, by now including wants that an individual may not have the means to obtain through exchange, we open up the possibility to not only any other object that they know about but lack these means will be included in the preference rank, but also that *we* may exogenously place wants on the ranking list as well. What difference would it make to an individual's preference ranking list if we include not only objects that they know exist, but lack the means to attain them *and* also add wants that they don't know to exist, and may or may not have the means to attain them? The process becomes one of absurdity where more is lost than is gained.

Second, our personal knowledge necessarily limits the extent of the wants we know to exist. For this reason, we see that not only is the previous exercise fraught with peril concerning the value we stand to gain from the extension, but it is logically impossible to state that someone desires a want satisfied that they don't know to exist. A preference ranking including want satisfactions that are outside the realm of our personal knowledge would not have any logical implication for the decision-maker.

Lastly, we see that value only exists as a relationship between two objects or events. In the previous example, when the individual traded $1,000 for a one carat ring, we were demonstrated that they *value* the ring more highly than $1,000. Value in exchange comes from the foregone alternative (in this case, the sacrificed want satisfaction from the $1,000 which the individual would not be able to experience). However, with nothing to sacrifice for a desired end, the problem that is created is that there is no basis for comparing the value. For instance, in our previous example with the 10 carat diamond ring, with a monetary value worth more than all the expected proceeds that would be obtained by an individual selling all their possessions. They could never offer anything in exchange for this 10 carat ring. We could say, of course, that they value it more highly than they do all their possessions aggregated. However, they would at the same time never act upon this fact as even if they were to sell all their possessions, the income would still be insufficient to purchase the ring. Their action is not affected by this end, even if it is desired, as they lack the foregone alternatives that would give it value.

Finally, we find that not only is knowledge of the specific means to be used to attain an end necessary for action to occur, but knowledge is needed of the causal relationship between means and ends. In fact, this causal relationship is the very essence that we wish to speak about when we use the term "rationality." It is the fact that an individual has limited and specific knowledge of given means and ends that they have a subjective rationality. The limitations of knowledge that each individual has creates the constraint on their logic system, which is given to them in an absolute form, and hence creates the need for action to be viewed in light of rationality. Additionally, as this knowledge set will be unique for each individual, rationality will likewise be a unique quality.

Hence, we find that ex ante, action will take place under two conditions. The first is the necessary knowledge of the ends attainable within the constraints of the means which an individual has at their disposal. Second is knowledge of the causal relationship that exists between these means and ends. As preferences are necessarily ranked from most to least desired, any action undertaken will necessarily be the highest ranked want at that time. This demonstrated preference has itself been demonstrated to be an inalienable fact of action.

Ex Post Action

All possible outcomes of action form a praxeologic dichotomy: success or failure. Ex ante we have seen that every action is necessarily shaped by the knowledge set that the individual has, and the resultant rationality that is endowed through that limited knowledge. Also, *every action is undertaken ex ante with the expectation that it will be successful*. However, *ex post* we find that there are a dichotomy of results that may obtain. Actions may prove to be either successful – that is, to fully meet the expectations of the actor – or failures – that is, to not meet the ex ante expectation of the actor. Each of these cases will be attended to below.

Successful actions imply that our most desired want, given our known and available means, has been fulfilled. A momentary equilibrium is fulfilled as the actor has satisfied the want that they have previously sought after. However, this fleeting moment will ultimately be so short as to be inconsequential. For the one thing that exists in super-abundance is wants. The moment that a desire is fulfilled, a new one shifts in to take its place – the ranking of wants is in a continual, dynamic flux. Success at meeting an end does not imply the logical conclusion of action and the achievement of a state of rest, rather, it implies a need to commence the process over, and start afresh the process whereby wants are satisfied.

The possibility of failure remains quite real, however, when undertaking action. An *ex post* failure implies that means (resources) have been used for the attainment of an end which has not materialized in the form that was previously expected by the individual. What will happen next will follow one of two paths, both dependent on the knowledge learned during the previous failed attempt at meeting the end.

At the planning stage when expectations are being made concerning the future end, it is not known with certainty that the end will obtain, and also it is unknown the true opportunity cost of the resources that must be dedicated to the attainment of the end. This is so owing to the dynamic flux of the individual's want ranking, coupled with the fact that means may not be dedicated to a given end until some point in the future, after the plan for action has been formulated. As these means will shift in importance to the individual, their relative value will also be altered. During the course of attaining an end, we see that a considerable amount of learning will occur as the actor is only able to find out with the passage of time what resources are truly necessary for the attainment of an end. The true cost of an end may only be determined *ex post* after the opportunity costs of the foregone alternatives may be seen.

When a failure in meeting an end is encountered, the end remains on the individual's value scale. This implies that it will still be sought after, in some capacity, the only question that remains is whether it is highly ranked enough to continue being sought at that moment.

If in the course of learning through the failed previous attempt an individual still values the end not yet met more than other alternatives, the failure will not be the end of this process. Instead, a second attempt will be applied, which, coupled with the new knowledge learned during the previous failure will result in new means being applied toward its attainment. The end will remain at a rank on the value scale superior to other ends, and hence, additional resources will be dedicated towards its attainment.

However, if in the learning process it is determined that the foregone alternatives are too great to warrant continual attempts at attaining the end, a halt to the process will result. Ex ante expectations of the true costs entailed in the end's attainment have proven to be incorrect, with the result that no further resources will be dedicated towards its achievement. In other words, due to the new knowledge of the opportunity cost involved with the end, it has shifted to a relatively less important place on the individual's value scale, with the result that other ends are deemed more pertinent and are sought after instead.

All actions are undertaken in the present with the expectation that *ex post* the end sought will obtain – there is a general equivalence between the ex ant and ex post expectations. However, it should become clear that ex ante and ex post states of affairs need not coincide (in fact, often times they will not). This divergence stems from the relationship that exists between certainty and uncertainty, or, the fact that we cannot exist in a state of full certainty at all times. In fact, action stems from the very fact that we are uncertain about an event's occurrence; it is a precondition for action's appearance. The true nature of this uncertain element must now be looked at in further detail before additional conclusions for its effect on action can be ascertained.

IV. THE KNOWN AND THE UNKNOWN

Previously we have seen the necessarily subjective nature of knowledge, in contrast to its more objectively definable background information. The limitations this partial knowledge set places on our rationality also has other consequences. Chief among these is that there must exit a realm of certainty (i.e., one defined by perfect knowledge and acted upon by logic) and another one of fundamental *un*certainty.

Partial knowledge of an event gives rise to a situation which we may best define as "risky." We know something about an event, but not everything. For that portion that we do not know we must develop a system to manage it so that we may act upon this partial knowledge. In contrast, there are some events which we know nothing about, that may best be categorized as "uncertain" situations. Although events may exist in the world, we have no existing knowledge of them, and hence their effects fall into a category which is unknown to us.

However, there are also events which we have some partial knowledge of, and yet it seems are difficult to classify as risky in our classical sense. As we shall see, Shackle's emphasis on "non-divisible, non-repeatable" events give rise to fundamentally uncertain outcomes. These correspond loosely with what Mises at the same time (circa 1949) identified as "case-probabilities." The implications for both classifications – non-divisible, non-repeatable events and case probabilities – is that the management system used for risky situations – probability theory – becomes wholly irrelevant.

Uncertainty, due to its fundamentally unknown elements, could be said to exist in a state of infinitude. We cannot speak of differing degrees of uncertainty, as by definition, we know that this element exists but not what it is comprised of. A gain in knowledge, which we typically associate with a gain in certainty pertaining to an outcome's attainment, does not have an offsetting decrease in the true uncertainty surrounding this outcome. However, what is changed is the degree of *felt* uncertainty. It is this felt uncertainty that affects the decision-making process – that which we are are ultimately concerned with.

Furthermore, the method used to deal with these two separate categories of knowledge – risk and uncertainty – are entirely different mental constructs. Risk exists within a definable set – a closed-end system – which allows for the use of probability theory. In contrast, there is no similar finitude with uncertainty, and hence the method used is one that also incorporates the open-endedness of the system. Such a method we refer to as *possibility theory*. When we refer to an event as possible to obtain, we are referring to the felt uncertainty that we perceive inherent in it. This makes no concern of the *un*felt uncertainty, which, as we will see, may only effect our decision-making process ex post.

Finally, we will look at a framework for which we can view these two concepts together. Both risk and uncertainty frame every decision we make, in some regard, and as such a method for viewing decision-making must take into account both of these factors. In fact, by viewing the decision-making process under conditions of both risk and uncertainty, we find that our expectations as to future events (and hence, our resultant decisions) are affected by new knowledge learned. More importantly however, we find that the way in which new knowledge fits in with our existing structure of expectations determines how our changes in decision-making will result.

Risk and uncertainty are inherent features of all actions. By viewing them within the an applicable framework, we see how changes in new knowledge affect our existing decisions, and hence, alter our future course of action.

1. Risk and Uncertainty

In volume 1 we looked at some differences between risk and uncertainty. Risk was shown to be a quantifiable measure which must objectively exist for our senses to gauge its parameters. In distinction, uncertainty represented those unknown elements whose existence we lack knowledge of. Knight (1921) would refer to uncertainty as a "fog" which engulfed the world we act in, and created opportunities to profit from accordingly. Mises (1949) expanded on this concept to define what he called case and class probabilities. Class probabilities are those which exist in a defined collective, and hence a frequency probability may be extrapolated to apply to future events within the collective. In distinction, case probabilities are those for which no collective exists, and hence, rely on pure entrepreneurial forecasting to determine what their projected outcome will be.

While providing some difficulties for our actions, risk and uncertainty also provide the motive for us to act. For if the future were determined, that is to say that there was no element of risk or uncertainty, would require no action as no change could result. In what follows, we will assess the current methods of dealing with the risky element, and offer an alternative framework for modeling decision-making under conditions of true Knightian uncertainty.

Uncertainty Axiomized

von Neumann and Morgenstern (1944) set out to properly deal with the world decision-makers face, as they try to direct their actions towards unknown ends. In their classical theory of decision-making, a set of axioms is provided whereby expected utility is maximized. These axioms are phrased in terms of probability however, and hence, from the start, open themselves only to describing situations characterized by risk. Since no explanation can be provided by this approach as to where probabilities emerge from when they are not assumed to be given, there is no method to support the search for a utility maximizing outcome if uncertainty cannot be quantified.[178]

Savage (1954) extended this paradigm to situations where objective probabilities fail to exist. This is achieved through the use of several axioms, most notably: that preferences are complete and transitive,[179] and that tastes and beliefs can be separated accordingly. Effectively, Savage transformed

[178] The influence of von Neumann and Morgenstern is apparent today as it provides the foundation of all risk and uncertainty literature. Christian Gollier's widely heralded book, *The Economics of Risk and Time*, explicitly builds off the expected utility framework to forward new theories of decision-making under risk and uncertainty. However, we find that these are considered synonyms throughout the text. As he (2001: xv) opens the book: "Uncertainty is everywhere. There is no field in economics in which risk is not a dimension of the decision-making environment." Later, he (*ibid*.: 3) continues the bifurcation: "Before addressing any decision problem under uncertainty, it is necessary to build a preference functional that evaluates the level of satisfaction of the decision maker who bears the risk."

[179] Previous attacks on the validity of transitivity by Loomes et al. (1989, 1991) and Blavatskyy (2006) have been recently

problems surrounding decision-making under uncertainty to decision-making under conditions of risk, *disguised as uncertainty*. By assuming that an uncertain event has a set of possible outcomes attached to it, Savage shows that choice can be explained due to the subjective probabilities attached to this expected uncertainty. However, we see the concept of an expected uncertainty is the antithesis of a true uncertainty.

Gilboa, Postwaite and Schmeidler (2008) outline the two main limitations of the modern utility maximizing approach to probability as the lack of as well-formed probability beliefs that are assumed to exist, and the silence surrounding the literature as to the origin of these beliefs. Both of these are dealt with statistically by defined *sets* of probabilities or distributions, however, these exclude uncertainty from occurring or being a factor in the decision-making process.

The assumption that a preference set is complete has also been called into question. This assumption is standard in consumer theory, and innocuous enough when outcomes are known in advance. However, given cases of true uncertainty, the assumption becomes decidedly less compelling, as Shackle (1972: 20-21), Shafer (1986) and Brewley (2002) have demonstrated. Likewise, these approaches also exclude qualitative effects from becoming apparent in probability beliefs. Schmeidler (1989) argues that if we have two probability distributions which are quantitatively similar, but are qualitatively different we will act upon them uniquely.[180] Much probability literature fails to distinguish between those based on knowledge, and those that have arisen by default.

Shackle (1972: 422) maintains that (von Neumann and Morgenstern's) theory of games:

> [i]s the product of a superb mathematical virtuosity... [However] by an extraordinary paradox, it *assumes away* the whole of that aspect of business, science, art and contest, which allows originative genius to exist... Surprise is that dislocation and subversion of received thoughts, which springs from an actual experience outside what has been judged fully possible, or else an experience of a character which has never been imagined and thus never assessed as either possible or impossible...

This paradox relies on the concept that a physical system is both unchanging and will vary. The variance will continue within defined boundaries and these will remain statically given over time. A

challenged, and transitivity defended by Birnbaum and Gutierrez (2007) and Birnbaum and Schmidt (2008). As Mises (1949) correctly notes, to say that preferences are transitive means only that they are transitive at any given moment. To imply that x is preferred to y which is preferred to z on this date, is not inconsistent with an individual preferring z to x at a later date. The shift in knowledge over time can change the preference ranking, and hence transitivity can be thought of existing only in strictly static terms.

180 Ellsburg (1961) shows that agents express different preferences for choices with the source of probability known, then those where it is unknown, regardless of whether the probabilities are identical. Likewise, Tversky and Khaneman (1974) find the same bias prevalent when qualitatively different probability sources are identified.

distinction must be made between games of chance, and games of skill that rely on our personal qualities – our entrepreneurial insights.

In fact, many statisticians of today if asked "What is probability?" would likely merely compute it for you. Yet if we return to the roots of probability theory, we find that they must be rooted in several crucial points. Richard von Mises (1928: 28-29) pointed out that probabilities can only be used in cases where: (1) a properly defined collective exists, and (2) an unlimited sequence of observations is available with a fixed and concrete limit. Much probability theory of today rests heavily on the law of large numbers, not on the notion of relative frequency as it was originally reckoned.

Much of this perversion occurred as the statistical methods were stretched beyond their abilities. As a result, when the mathematical method could not model the probabilistic world, the world was altered to make this possible (Gigerenzer et al. 1989: 226-227). The mind was idealized to draw samples from incoming information to produce probabilities accordingly. However, as Kahneman and Tversky (1972) have shown, there is no evidence that people act as computationally rationally as probability models suggest. To view man as a reasonable intuitive statistician will lead to inherently discouraging results (Nisbett et al. 1983: 340).

This alteration as to the nature of probability has led to much ambiguity and confusion surrounding the results which must now be addressed.

Seriable and Divisible Events

One of the greatest lessons from Cantillon's *Essai* is that history consists of making human affairs. It follows that we cannot know history until we make it. But much probability theory assumes that we can know the outcomes, at least probabilistically, before they occur. This point would mark much of G. L. S. Shackle's work, as he would remain a harsh critique of the numerical approach to probability theory his whole career. Three aspects of trials would bring Shackle (1949b) to stress the importance of the non-seriable, non-divisible nature of decision-making.[181]

First is the uniqueness of trials. The effect each action gives to an individual is different each trial and is a solitary experience for the agent. Second is the isolated nature of trials. As they are undertaken at different points in time, they must represent different experiences complicating their direct comparability. Last is the crucialness of outcomes. Each outcome will affect the future outcomes that follow. In fact, this last point is the most important to stress. It is all pervasive in the action of

[181] R. von Mises (1928: 15) earlier explained the limits of "probability" theory in these cases: "'The probability of winning a battle,' for instance, has no place in out theory of probability because we cannot think of a collective to which it belongs. The theory of probability cannot be applied to this problem any more than the physical concept of work can be applied to the calculation of the 'work' done by an actor in reciting his part in a play."

choice – "it steers my life into a path different from what it would otherwise have followed" (Shackle *ibid*.: 63). When events are defined by any one of these characteristics, we find the hypothetical outcomes non-additive. In effect, we find the outcomes are mutually exclusive; we cannot average the rival hypotheses, instead we must choose between them.[182]

In this well known example, Shackle (1949c: 8) provides a succinct example of a non-seriable experiment:

> Suppose the captains in a Test Match have agreed that instead of tossing a coin for choice of innings they will decide the matter by this next throw of the die, and that if it shows an ace Australia shall bat first, if any other number, then England shall bat first. Can we now give any meaningful answer whatever to the question 'Who will bat first?' except 'We do not know'? For a non-divisible, non-seriable experiment the concept of frequency-ratios is wholly irrelevant.

Of course, it can well be pointed out that if we were given a choice between betting on a coin toss or the throw of a dice, we would be provided better odds with the former.[183] However, it must be recognized that this is only the case if we can repeat the trial enough times to create a valid frequency distribution. Since the Test captain has only a single chance to decide, the distribution built off the past frequency is irrelevant. This point has been clarified later in Shackle (1952a: 110-111) where he phrases the choice as being:

> between drawing a card from a pack containing 25 red cards and a black one, or from a pack containing 25 black cards and a red one; and if the drawing of a red card were destined to transport him to eternal felicity and that a black one to consign him an everlasting woe, it would be foolish to deny that he ought to prefer the pack containing the larger proportion of red cards, although from the nature of the risk, it could not be repeated... [S]uppose he should choose the red pack and draw the black card. What consolation would he have? He might say that he had acted in accordance with reason, but that would only show that his reason was absolutely worthless. And if he should choose the red card, how could he regard it as anything but a happy accident? He could not say that if he had drawn from the other pack he might have

[182] Likewise, Egerton (1960: 4) cites two limitations of statistical probability under conditions of uncertainty. First is that an *expected* outcome can only be significant to an investor if it is contained in a group of similar outcomes (Shackle's uniqueness). Second, as many trials *are* unique a lack of data exists to derive *expected* outcomes. It should be emphasized that it is not a *mere* lack of data that excludes a probabilistic approach, but an *impossibility* of gathering uncertain data.

[183] This is a point raised by Ellsburg (1961: 664fn) whereby he defends the use of the frequency distribution, disregarding the relevance (or rather lack thereof) it has for a unique experiment.

drawn the wrong one, because an hypothetical proposition such as 'If *A*, then *B*' means nothing with reference to a single case.'[184]

We can summarize these objections to the frequency-approach to probability for unique events as being three-fold. First, entrepreneurial decisions seldom form a series of repeatable events. Even if they were to allow for this unlikely repetition external conditions would never allow for the same experiment to be repeated under identical circumstances. Second, some choices eliminate the opportunity for continued trials. A gambler, for example, who loses their money will have to retire from the table and cease any further trials. Last, possible outcomes are therefore to be regarded, not as related occurrences, but as wholly independent and mutually exclusive trials. The emphasis should be laid on a single event occurring in the future; instead we find frequency theory purporting to extrapolate this singular future event, when in fact it refers to a *large class of events having occurred in the past* (O'Connor 1954: 14).[185]

In fact, by underscoring the uniqueness of decisions, we see that the uncertainty agents live through give rise to Mises' case probabilities – those lacking any form of historical distribution relevant for their future probabilities. As Lachmann (1956: 26-27) points out, however, the emphasis would be better removed from the uniqueness of decisions, and instead should be rooted in the "heterogeneity of situations."[186] Few decisions are truly unique in the sense that they are solitary events. In fact, many decisions are routine, repeated events throughout life (see Hoppe 1997). Lachmann highlights the fact that there are almost certainly more situations than decisions in the world, as each decision may breed a different result. However, we find that both categories are bounded by the uniqueness of results – the end sought. Lachmann's heterogeneous situation creates the environment that Shackle's unique decision is born within. Both these events arise due to the bounded nature of results that are expected to occur. In fact, Mises (1949: 105) points out the devastating effects that an unbounded nature of results would have for our actions – a world with a known future would incite no action, a world with an unknowable future would leave no recourse to purposeful action.[187]

---

[184] Similar examples are provided throughout Shackle's work, notably (1952a: 25 and 1961: 59).

[185] Taleb (2004: 56) explains the difficulty in viewing the future in a less pre-defined (if only conditionally) way: "When you look at the past, the past will always be deterministic, since only one single observation took place. Our mind will interpret most events not with the preceding ones in mind, but the following ones. Imagine taking a test knowing the answer. While we know that history flows forward, it is difficult to realize that we envision it backward."

[186] Mises (1954: 88) also agrees that probability concerns that which is "produced" by an individual. Hence, his own emphasis was on the outcome, and not specifically the decision. Knight (1921: 204) explicitly mentions Locke as an influence in sharing this same view. As primary qualities are combined, secondary qualities are created. Primary qualities (typically more concrete elements) can be combined in an infinite number of combinations (typically outcomes). Hence, we see that decisions and situations may not be necessarily unlimited in number, but outcomes from these decisions are.

[187] See also Shackle (1958: 48) as he notes that uncertainty must be bounded to some degree. For if there were no bounds on the consequences of action, there could be no purpose driving our actions.

Empirical evidence that trials concerning human created events are non-seriable, non-divisible, has been accumulating since Shackle's work. In fact, some research points to agents' inability to use the historical distribution, even if it exists! Khaneman and Tversky (1973; 1979b) and Tversky and Khaneman (1980: chap. 10) find that agents tend to rely on singular information, even if it is scanty or unavailable, or give a low weight to the distributional data available. Furthermore, contra the law of large numbers, a law of *small numbers* has been forwarded. As agents act to compensate for events opposite of what the previous average "predicts" we find that large outliers are attempted to be compensated for by treating them not as part of a larger set, but as unique events to be averaged out by the next action (Tversky and Khaneman 1971). Shackle's (1952a: 6) own view on "large numbers" was that it served only to "turn our ignorance into knowledge, or doubt and fear into assurance."[188] As we are assured that our results will average out, we lose sight of the assured outcome that if we are wrong on the first attempt, we can attempt no further trials.

Dynamic Uncertainty and Static Risk

If we take the distinction between risk and uncertainty as one of measurable and immeasurable elements (as per Knight) we see that the source of the difference lies in a temporal element. Risk exists in a static state. There is knowledge which we do not know about, yet we know that it exists. Taking Mises' definition of class probabilities, we find that even though we may not have knowledge about any given object or event within a certain class, we do know that it exists as a part of the class, and also that it's behavior will be cohesive with the rest of that class. Hence, we know that a risk exists within the class, and that in a static state it will continue to exist.

We also see that two definitions of uncertainty have emerged. One is the Shackleian defintion, whereby uncertainty results from the actions undertaken that serve as isolated events changing the future state of affairs. Mises hold the other view that uncertainty exists exogenous human action; some events are uncertainty by their very nature (case probabilities).[189] In fact, we see that both definitions are two sides of the same coin. Returning to Knight's original conception, the gist of uncertainty was not about probabilities – case versus class for Mises – but between possible and probable events, á la Shackle. However, this implies – contra Shackle – that the passage of time need not only be the sole source of uncertainty as individuals are placed in situation in which there is latent uncertainty in existence (case-probabilities). At the same time uncertainty is a product of time as the possibilities we

[188] More recently, Taleb (2007: 138) points out that increases in knowledge can lead to corresponding increases in "confusion, ignorance, and conceit." Keynes (1921, 336) insightfully points out that it matters not that the sample size is large, but rather that its mean is stable.

[189] In fact, Mises seems at times unclear which definition of uncertainty he favors. See Mises (1949: 105) for examples of this bifurcation of definition occurring on the same page.

can be a significant source of uncertainty, but as Garrison (1984) points out that the opposite need not necessarily be true. Uncertain situations need not be caused solely by time, hence, although time breeds uncertainty, uncertainty is a poor proxy for time

However, in a dynamic setting the situation changes markedly and uncertainty does gain a new complexity. Uncertainty becomes an issue only due to the appearance of new temporal viewpoints, which were unable to come into being prior to the temporal passage. As knowledge evolves through the passage of time, we see that there will exist future possibilities that will not be able to be quantitatively reduced or foreseen in the present. These immeasurable uncertainties (again to use Knight's distinction) correspond to what Mises preferred to refer to as case probabilities. An event *may* occur, although we may not have knowledge that it occurs, and also we do not know anything about its nature – it exists as a class to itself.

However, it can be seen that risk can also be applied to dynamic settings. For example, we know that x% of glasses on an assembly line break, and that to produce an amount y suitable for sale we must produce $y + x$%. We have ex ante knowledge of what must be undertaken to obtain our desired production level ex post. This stems from the distinction between the natural and physical worlds. The production of glasses represents an element of constancy. There is no learning process involved, time passes for the production process but does not instill change in the event itself. Risk may be applied to a dynamic setting where the process itself does not endogenously enact change upon its own future course of affairs.

However, processes that endogenously enact change upon their future course of action cannot be treated in terms of risk. As Shackle pointed out, this stems from three points: the uniqueness of trials, the elimination of future trials through present attempts, and independent and mutually exclusive outcomes that result from trials. Human actions fall into all three of these categories, with our decisions and the outcomes of our actions representing Shackleian "non-seriable, non-divisible" events. It then becomes clear that traditional risk-theory cannot properly manage these outcomes, owing to the nature of the outcomes themselves. The risk-based approach cannot be fine-tuned to answer these questions more clearly, as the question they seek to answer is of a different category than the one that it is capable of answering.

Instead, these events that fall into this category of uncertainty must be dealt with in a drastically different manner than the commonly assumed risk-based approach. In fact, Shackle's distinction between the possible and the probable now becomes clearly the distinction that must be utilized when viewing decision-making under uncertain conditions.

2. Probability and Possibility

The distinction between risk and uncertainty creates a dichotomy of concepts that the individual uses to manage these different obstructions to plan attainment. If, as we have seen, risk and uncertainty remain two fundamentally different elements that must be dealt with through action, the methods that we use to manage these elements may also require different approaches.

The Infinitude of Uncertainty

Uncertainty, by definition, represents an unknown element, void of measurement. As a result, to speak of an *amount* of uncertainty existing at any one time becomes a moot point. We cannot speak of a definite, or indefinite, quantity of uncertainty. In fact, we cannot even say that we "feel" more uncertain about a given situation; uncertainty must remain as a continual state of existence, but we cannot say that its existence has expanded or narrowed concerning our actions.

Our limited knowledge gives rise to uncertainty, which we now know to exist owing the partiality of our knowledge. However, we also know that uncertainty is not an absolute state, engulfing all of our future actions in an indeterminable realm, and hence making all action impossible through an uncertainty concerning the causal connections between means that will strive towards ends. Lachmann (1976a) may have reckoned that the future is forever "unknowable." However, as Hoppe (1997) correctly points out, we know that the future is uncertain, and hence, there exists some degree of certainty concerning the concept. We are not engulfed in total uncertainty, but there exist elements that we are also certain about – *even in a dynamic setting*.

However, as uncertainty represents the realm of unknown unknowns, we may never speak of differing degrees of this concept. When an individual says "I have less uncertainty surrounding event *x* than event *y* obtaining," what they are really saying is "I feel more risk surrounding event *y* than I do about event *x*." The corollary of uncertainty is not certainty. To make a statement such as "I have less uncertainty about a future event, therefore I have more certainty about it" is to confuse these two concepts. For if an event is certain to occur, than we have full knowledge surrounding factors effecting it. But if we say we have uncertainty surrounding an the obtainment of an event, than we are saying that we have only some knowledge about the factors affecting it, and no knowledge of other factors. To be *fully* uncertain of an event is not even a logically valid statement! For to be fully uncertain one would have to have *no* knowledge about this event – not even of its *mere* existence.

Due to the fact that we have *no* knowledge of some uncertain events, to say that a decision has become more *certain* does not imply that it has simultaneously become less *uncertain*. It only implies that we have learned knowledge surrounding its attainment that we did not have prior, but says nothing about the knowledge that we still do not know anything about (as we cannot state anything about knowledge yet to be learned). What has changed is the amount of *perceived* uncertainty that exists, or Knightian fog that shrouds the outcome.

In our decision-making process, it is not the uncertainty that concerns us (as it could not be definition) but rather the effect this uncertainty might have on the obtainment of our ends. It then becomes clear that perceived uncertainty, that is to say, that which we think exists is what affects our actions. This amount of perceived uncertainty makes no reference to the total uncertainty outstanding at any given time. For example, when undertaking an investment in a foreign country, the fact that we perceive a lesser degree of uncertainty concerning the outcome of a war does not decrease the total amount of uncertainty surrounding the event. For there still remain an infinite number of uncertain events which we do not know about that may effect the investment.

The perception of this uncertain event does not, however, imply that the event is now a risky factor. For, to use the previous example, the outbreak of war still represents a non-seriable, non-divisible event which must fall into the category of uncertainty. What we have instead made note of is that we assign a greater subjective degree of the possibility that an uncertain event may obtain, without making reference to a measurable probability (subjective, logical, or frequency).

Is the Probable Possible?

As early as 1761, D'Almbert (1717-83) demonstrated that events with minuscule probabilities would be acted upon as if the probability was zero. Buffon (1707-1788), writing during the same period, reckoned that a person could not distinguish between a probability of 1:10,000 and zero – the odds were so small that a person would treat them as being *impossible*.[190] This subtle point has been picked up on, and lightly expanded upon, in more modern times. The implication for decision-making under cases of uncertainty is that agents only respond to hazards that they can perceive (Slovic, Fischoff, and

[190] Cohen (1964: chap. 9) outlines the early history in viewing probability in terms of possibility.

Lichenstein 1980: 463).

Lange (1944: 29) revived this viewpoint,[191] coining his "practical range" to describe the field of expected events that bound our actions:

> In most cases the entrepreneur or consumer does not consider the whole range of possible values of the expected price, but disregards the extreme values at both tails of the probability distribution. He does so because the joint probability of these extreme values is too small to bother about. The range with the tail values thus cut off will be called the '*practical range*' and will serve as our measure of the degree of uncertainty of price expectations.

Individuals do not act directly upon an outcome that they do not expect to occur with any degree of possibility. As Shackle (1958: 40-42) describes the process, decision-makers are concerned with what is possible to occur, and not what is necessarily probable. Possibility in this sense is wholly subjective – an outcome is possible only if a decision-maker deems it to be so. Hence, we find that agents can anticipate with some degree of realism both very negative and positive events occurring. There are also events which an agent would view completely possible as occurring. The result is that we can eschew the traditional probability distribution for the "potential surprise function," as pioneered by Shackle (1952a) and illustrated in figure 1.[192]

[191] Although R. von Mises (1928: 67) preceded this in hinting implicitly that possibility may be a more optimal measure than probability. Additionally, as Lachmann (1945: 125) points out, Lange, like most other writers in the field, discards the concept of a practical range for point certainty (as defined through probability) soon after outlining this discovery. Shackle's (1938) early work even stressed a point "certainty equivalent" outcome. Several earlier economists stressing the range of prices instead of a singular price include Menger (1871: 91-97) and Böhm-Bawerk (1889: 195-222).

[192] As has been pointed out to the present author by Pete Leeson, it is may be better to rename Shackle's "potential surprise function" with a term more suitable. A function is, by definition, mathematically additive, which is the complete opposite result we wish to develop. For our purposes, the reader will do well to bear in mind that the term "function" is used in a sense different than that of a mathematical logic. In the interest of maintaining a type of historical longevity this is deemed appropriate, however, it may become advantageous to find a more relevant and exact term in the future.

Potential Surprise Function

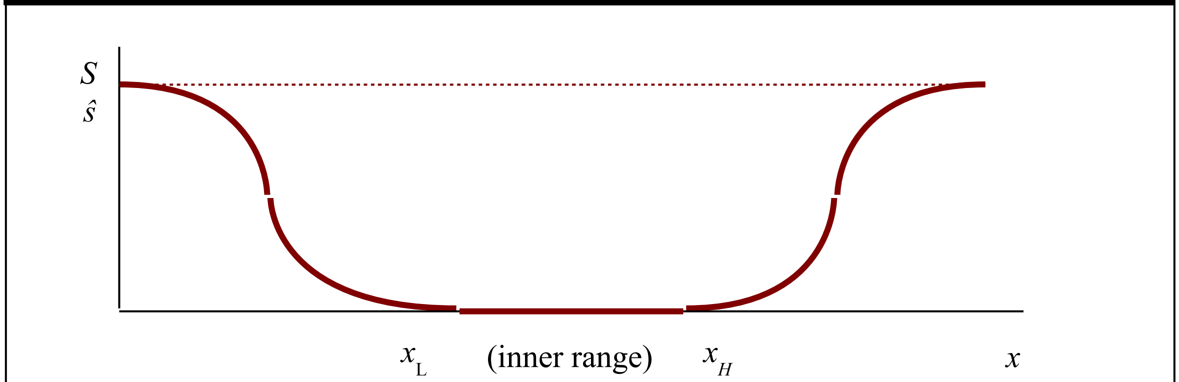

Figure 1

Points along the abscissa represent the outcomes that the actor deems possible to occur. The range between x_L and x_H represents the inner range where the agent deems outcomes to be all perfectly possible. These outcomes will result in no surprise to the agent should they occur. The y-axis represents the degree of surprise a given outcome will entice (here labeled S to denote surprise). As the outcomes move away from the two points bounding the inner range, the possibility of them occurring, that is to say, the surprise generated by their occurrence, increases until a point is reached, $\hat{s}$. At this point, the maximum surprise possible is generated by an outcome. An event more extreme than this (i.e., a value of x greater or lesser than these the x-intercepts of outcomes generating surprise $\hat{s}$), is deemed to be an impossibility by the agent – its occurrence is so remote as to note even warrant attention. As outcomes generating more surprise than $\hat{s}$ are viewed as impossibility, there can be no differing degrees as to how impossible they are thought of. They become *non sequiturs* to the agent due to this fact.

Surprise becomes an event which must occur in some degree with all actions undertaken. It is inextricably linked with the concept of uncertainty, and as not event obtains with full certainty, we find that some degree of surprise will be entailed in every action's result. In figure 1 about, when we say that zero surprise is expected by an outcome's obtainment, we are really saying that the minimum level of surprise possible will be felt by such an outcome.

The differences between the potential surprise function and a frequency distribution curve may seem subtle in many ways, but the implications drastic. Frequency distributions place outcomes into additive classes according to their probability of occurrence. It therefore follows that these distributions lend themselves to being multiplied by their probabilities, with the result that point-*un*certainties become probabilistic point-certainties (certainties used in the statistical sense of knowable unknowns). This operation excludes the possibility that surprise could be induced in an agent's mind by an occurrence. The potential surprise function, as shown in figure 1, allows personal judgment to determine what outcomes will affect the decision-making process. It is this personal judgment that compels each agent to assign a certain amount of surprise to each potential hypothesis (even if this is done vaguely or unknowingly).

A distribution tells us with near certainty that if enough trials are undertaken, we can make a statement about the total outcome. This total outcome is, of course, dependent on all the trials being undertaken and added together. The potential surprise function operates in a non-additive way. Indeed, as Shackle (1961: 73-74) sums the reasoning:

> Such-and-such a thing ... could perfectly easily happen and it could perfectly easily not happen. Thus the happening *E* and its absence, not-*E*, are each assigned zero potential surprise. But not-*E* can nearly always be split up into many or even infinitely many component particular happenings. If *E* is rain, not-*E* will have come true if there proves to be sun, or fog, or snow, or hail. How can probability deal with this need to accord equal status to *E* and to not-*E*? To give a probability of ½ to rain will leave the remaining ½ to be shared amongst sun, fog, snow and hail, and if, as may well be the case, he feels that each of them deserves an equal status with rain, he will be totally frustrated. By contrast potential surprise, because it is non-distributional, can be assigned in zero degree to an unlimited number of particularized components of not-*E* and so to not-*E* itself.[193]

In contrast, the addition of a new outcome to a probability distribution will serve to decrease, at least one or some, of each existing events' respective probability of occurrence. However, we see that the addition of a new outcome to the potential surprise function will not glean the same result. Each additional outcome may leave each existing outcome equally possible – potential surprise is not affected by the total number of possibilities (Shackle 1949c: 9-10).

A range of events that appear to be perfectly possible to the agent may, and often will, result. The substitution of probabilistic point-certainty eliminates more than we gain however (Lachmann

[193] Shackle (1952b: 31) illustrates this idea with the occurrence that he sees an acquaintance on the London tube reading a book. Although the book could be in several languages that would not cause any surprise (English or French), he would be mildly surprised to see it written in German, Italian or Spanish, and very much surprised to discover that it is a Dutch or Welsh book. Finally, a book in any other language is a possibility so remote it may be excluded from the set.

1956: 25-26). When undertaking a decision, we often must resign ourselves to the fact that we have a range of outcomes that we feel are wholly possible, and some whose occurrence we deem wholly impossible. A trip to the grocery store to purchase a liter of milk may yield the expectation that *any* price between $1.75 and $2.25 is wholly possible; the agent would not be surprised in the least to see a actual price in this range. Prices ranging from $1.25 to $1.74 and from $2.26 to$2.75 may be deemed somewhat possible; the agent would be a little surprised by their occurrence, but is able to foresee their possibility. However, the agent could deem prices lower than $1.24 and higher than $2.76 to be wholly impossible; their occurrence is thought of as being so remote that they assign no possibility of their occurrence. It follows that when the agent is making their initial decision to proceed to the store to purchase the milk, they will act upon their expected possible price ranges. Hence, one outcome may be that they will only bring $2.75 with them, as they view the occurrence of a price greater than this to be outside the realm of possibility, and relevance. Note, however, that as there is a range of prices all deemed perfectly possible, there be no sense in which an agent can feel any degree of positive confidence in any of them (Shackle 1958: 44). All occurrences must be viewed as equally possible.[194]

Expected outcomes in the inner-range of fully expected possibilities result in no pleasure or pain for the agent.[195] Their lack of surprise due to the full expectation surrounding their occurrence leads to no change in *emotion* surrounding their realization. However, outcomes in the upper outer-range, that is to say, outcomes more positive than what was deemed fully possible without surprise, will elicit a positive feeling in the agent once reached. These were only deemed relatively less possible to occur, and if realized through a chance event, our expectation will have been exceeded. The corollary occurrence in the lower outer-range elicits the opposite reaction. As this negative event passes which was viewed as only mildly possible, an agent feels a pain as their expectation is not fulfilled (Shackle 1954: 96).

It may be raised that by focusing solely on the possible outcomes, uncertainty is, by definition, excluded from the analysis. However, as Lachmann (1945: 129) points out:

---

[194] Carter (1958: 544), in his review of *Time and Economics*, seems to have missed this point. By criticizing Shackle on grounds that, although he personally admits to not being surprised should either Labour or the Conservatives win the next election, he would be more likely to bet money on one than the other. As we will show shortly, by incorporating the concept of subjective probability into these equally possible expectations, one choice may become appear to be more valid than another.

[195] Note that this applies only to the expectation, the actual physical event may cause considerable pleasure or pain. Suppose that a surgery occurs whereby the recuperation phase in hospital is estimated to last 5-7 days (that is the boundary of the inner range). If the patient stays in hospital for 6 days, their expectation will have been fulfilled as per the expectation, and no pleasure or pain will have resulted. However, the actual action of recovering for this time will have caused considerable emotion (likely pain).

The width of the range expresses the degree of our uncertainty about the exhaustiveness of the information at our disposal. If we thought we knew everything relevant to the expected event about the forces, major and minor, which shape the situation, we could predict one price with certainty. An increasing range expresses an increasing uncertainty about the completeness of our knowledge.

Uncertainty is what shapes the potential possibility function. The width of the inner-range (deemed entirely possible) will be narrowed when uncertainty is small, and widened as uncertainty increases. As can be seen in figure 2, a set of outcomes will be relatively less certain if its dispersal is broader than a different set of outcomes.[196] Expected outcomes falling on the outer, red line exhibit more uncertainty of occurrence. The inner range (x_{L1} to x_{H1}) is much broader than that of the inner, blue line, which provides an example of a set of outcomes deemed more certain of occurrence.

Multiple Scenarios of Uncertainty

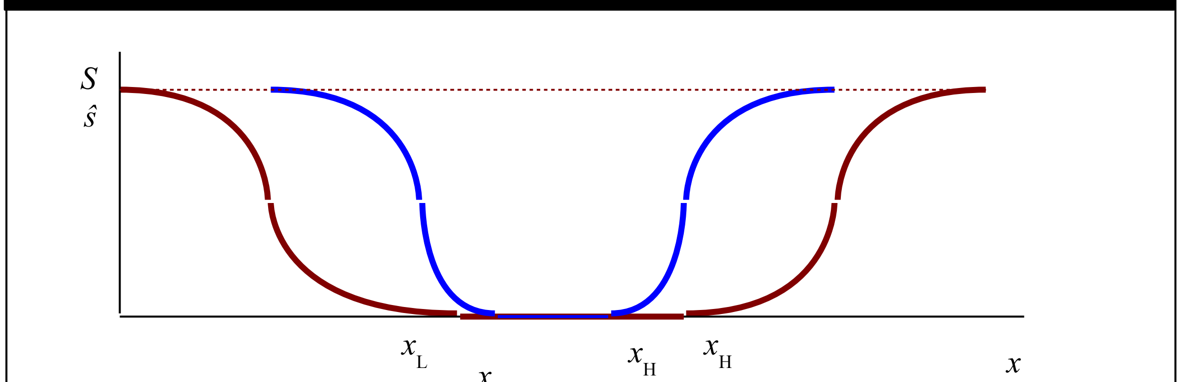

Figure 2

Watikins (1955: 77) describes Shackle's function's "*y*-axis [as being] like a string fixed at both ends, and his *x*-axis is like a straight elastic string fixed in the middle. Values on these axes will never get out of order but they can be bunched." We see that as the *x*-axis is stretched, our expected outcomes

[196] Prospect theory concerns itself with the reality that agents underweight probable outcomes compared to more certain alternatives. Kahneman and Tversky (1979a) find that reducing the odds from 0.1 to 0.0 will be more valued than reducing the odds from 0.2 to 0.1. Likewise, Ellsberg (1960) found that people avoid risks altogether when the probabilities are unknown. Becker and Brownson (1964) show that people will actually pay money to avoid making choices with unknown probabilities.

remain transitive. Likewise, the *y*-axis can be stretched vertically, illustrating our increased surprise at some outcomes. The slope of the function, combined with the width of the inner-range, can be thusly viewed as illustrating how certain we are about a set of outcomes occurrence. Uncertainty is a dominant force that defines the shape, extent, and elasticity of our expectations. Instead of being erased from the analysis, it becomes instrumental to it.

Focus Points

To say that agents make decisions based solely on the possible outcomes would oversimplify the decision-process.[197] In fact, the realization that the expected possible outcomes frame the decision is only the first step toward analyzing the process. Admittedly, structuring the problem within the confines of the possible events is a crucial, even the most crucial, part of the analysis. Phrasing the problem in terms of possible outcomes defines the uncertainty that the agent feels they face. In fact, as uncertainty changes the dynamic between their elasticity of expectations and the range of expected outcomes, we see that the overriding question of uncertainty is re-packaged and internalized within the agent's potential surprise function. As Kirzner (1982) points out, creating a defined means-ends framework must occur prior to an individual allocating effectively the means necessary for the achievement of ends. As our world is necessarily open-ended, the necessity for the definition of a closed-ended framework must come logically prior to any decision can be made about a particular end.

The claim that agents make no use of a subjective probability distribution would be too strong a claim in light of the decision-making process. If possibility is the first step used to manage the uncertainty of an outcome, a probability distribution is used to manage the risk inherent in it. The distribution used, however, it is necessarily quite different than that claimed by modern decision theorists.[198]

We have outlined previously that an inner-range of outcomes exist that are each equally possible in the eyes of the agent. What this means for their subjective probability distribution is somewhat different than is commonly assumed. What it does mean, à la D'Almbert and Buffon, is that the probabilities assigned to each chance occurrence are so similar that the agent cannot distinguish

---

[197] Even Knight (1921: 226) would admit that judgments of probability in the face of uncertainty are likely.

[198] “Entrepreneurs and consumers need not and usually do not, visualize an exact probability distribution of possible prices. For our purpose It is sufficient to assume that each person forms some idea about the most probable value and the 'practical range' of the expected price” (Lange 1944: 30). The distribution may be ill-defined, but will exist due to the structure of the potential surprise function.

between them. Assume a coin weighted in a uneven manner. If flipped, heads appears up 49% of the time, and tails appears up the other 51% of the time (on average, over a sufficient amount of trials). For a decision-maker with only one flip of the coin, and a corresponding guess at which face will appear up, the two probabilities are so similar, that they are for practical purposes indistinguishable. For this reason, the agent will express no degree of surprise in seeing either a head or a tail appear, despite knowing that the probability of success will be higher (provided enough trials are undertaken) with a guess of a head than a tail. The agent's guess may be conditioned by the probability distribution expected to prevail over the long series of trials, but this frequency distribution has no logical implication for the guess – as Shackle would say, if the agent guessed heads and was correct, what could they attribute this to but luck, and if they guessed heads and lost, what could they do but blame their faulty reasoning for providing an incorrect answer.

A subjective probability distribution will exist but only along the whole potential surprise function, *not the whole of the potential outcomes*. As outcomes with values of surprise (S) less than 0 cannot occur, and outcomes that evoke more than total surprise ($\hat{s}$) are viewed as being impossible occurrences. A probability distribution then can only appear along the values of outcomes deemed possible (Shackle 1961: 147).[199] The values of the probabilities (both the spreads and the absolute values) will be conditioned by the shape and location on the function. Outcomes all lying within the inner-range, and hence all perfectly possible, will see probabilities defined as very similar (and not necessarily identical, although they could be). Outcomes becoming less and less potential – those eliciting more and more surprise should they occur – will become progressively less probable. This will continue until the final two outcomes, those deemed the smallest or largest possible, will see the smallest probability of occurrence. As Knight (1921: 227) pointed out, the formation of this probability distribution ultimately rests on a two-step process: “the formation of an estimate and the estimation of its value.” It may be noticed that the formation of the estimate, in the Knightian sense, involves an individual creating a range of possible outcomes. Next, the estimation of this value of this estimate is

[199] Anyone questioning the validity, at this juncture, of using a subjective probability distribution in the realm of decision-making should remember the words of Mises (1957: 88):

> What the neo-indeterminist school of physics fails to see is that the proposition: A produces B in n% of the cases and C in the rest of the cases is, epistemologically, not different from the proposition: A always produces B. The former proposition differs from the latter only in combining in its notion of A two elements, X and Y, which the perfect form of causal law would have to distinguish. But no question of contingency is raised.

See also Mises (1962: 23) as he notes that this distribution will be an “ultimate given,” with no reason to directly understand its creation.

given by the probability (either subjective, historical, or logically attained) which is assigned to all possible outcomes.[200] It should be noted that the distribution along the potential surprise function in no way requires an assumption of symmetry. Figure 3 illustrates a subjective probability distribution transposed onto the potential surprise function. Note that probabilities only occur for those outcomes believed to be attainable, and that the inner range is dominated by probabilities so close in magnitude that the decision-maker can not meaningfully distinguish between them.

Subjective Probability and the Potential Surprise Function

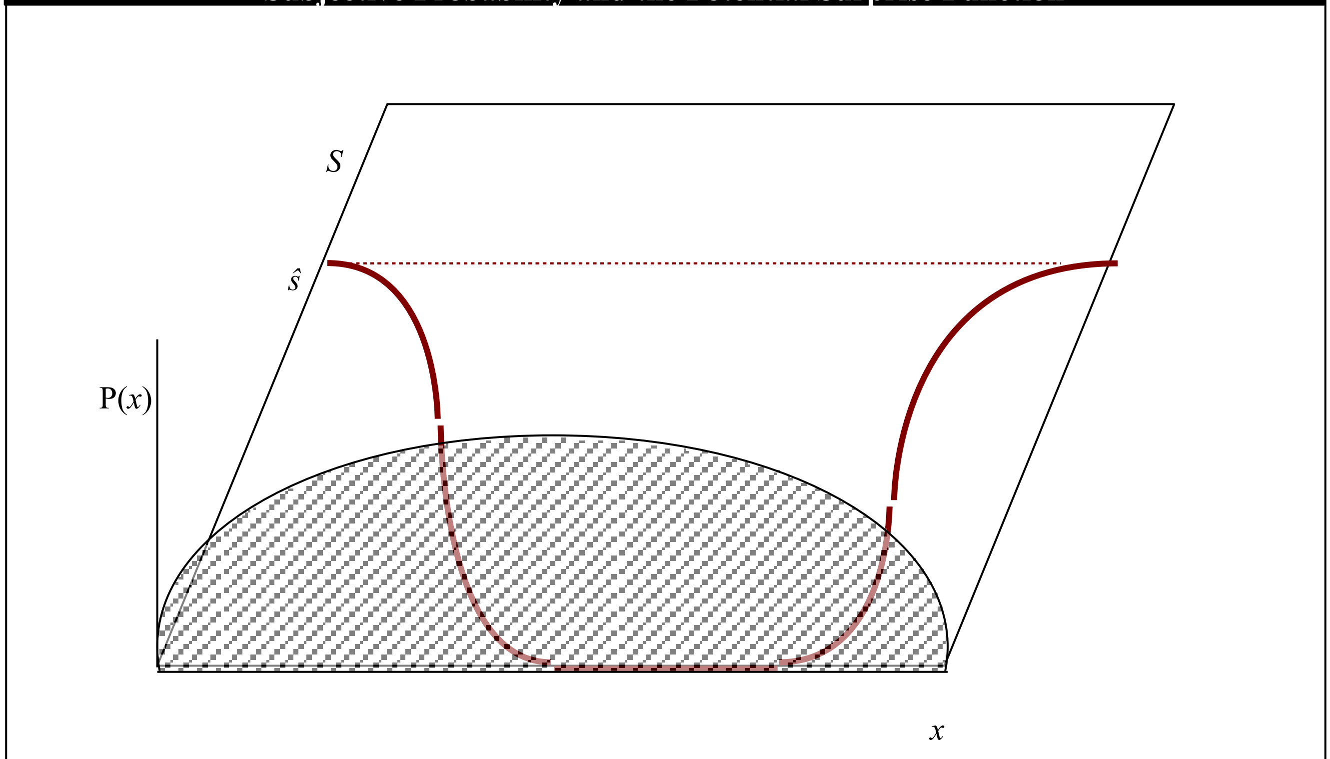

Figure 3

If we look at table 2, we can see a sample outcome with both the potential surprise, and expected probabilities, respectively.

[200] Langlois and Cosgel (1993: 460) come to the same conclusion, where they note that Knight's first stage, the formation of an estimate of the outcome, is a wholly subjective category. The second, the estimation of the value inherent in the first category, requires probabilities with which to rank the validity of expected outcomes. Even through these probabilities may, in many cases, be subjectively attained, they represent a different category of estimation – from wholly qualitative in the first instance, to numerically quantitative in the second.

Sample Outcome			
Outcome (x)	**Potential surprise (y)**	**Probability (%)**	**Expected outcome (Φ)**
0	$\hat{s}$	5	0
1	r	15	0.15
2	0	25	0.5
3	0	30	0.9
4	r	20	0.8
5	$\hat{s}$	5	0.25
		Expected cumulative outcome: 2.60	
Table 2			

Potential surprise is given by the letters r, and $\hat{s}$ (remember that some values will garner no degree, or more correctly, the minimal degree, of surprise upon occurrence). $\hat{s}$ represents a maximum amount of surprise; all outcomes greater than those with a maximum $\hat{s}$ degree of surprise are deemed impossible occurrences. Likewise, outcomes with 0 surprise values are deemed fully like, they would elicit no response if they occur. Outcomes with potential surprise of r are possible, but would be somewhat of a shock if they occurred. Letters are used instead of a more conventional numbering system to denote and stress the non-additive nature of surprise.

Note the seeming disconnect that occurs with the five outcomes, their surprise values and subjective probabilities. Outcomes with values of 2 and 3 both are fully possible to the agent. Neither of these occurrences would surprise them in the least. However, they are represented by different probabilities, at 25% and 30% respectively. The probability values are too close to effect the agent's decision-making process. In their mind, both 25% and 30% are equally possible, and hence, viewed the same way. Probabilities, however, will tend to decrease the more potential surprise will be incurred by an outcome.

Modern probability theory would tell us that the outcome we can expect from this example is 2.60, that is, the weighted probabilities of all the chance occurrences. However, there is a significant flaw in this reasoning. In fact, the only thing we know with *absolute certainty* about this (assumed) non-seriable, non-divisible case is that the outcome will *not* be 2.4! The outcome will be contained within the outcome set (values 0 to 5 inclusive).The probabilities are, however, quite useful as a cognitive tool.[201]

[201] In fact, the addition of a subjective probability component removes previous criticism Shackle's potential surprise function had received. This omission can be viewed, in part, as the reason the function has been withdrawn from risk and uncertainty literature. See Foldes (1958) for an early criticism on this ground.

If we use this example and ask the question "Will you play a game with this outcome set if the cost of doing so is nil?" the answer will almost certainly be in the affirmative. There will be no cost foregone, and an almost assured positive payoff. However, what occurs if a cost is associated with the case?

If we change the example slightly to include a $2 fee for partaking in the example, and leaving all other data the same, we see a new pay-off table created, as shown in table 3.

Cost Adjusted Sample Outcomes				
Outcome (*x*)	**Net outcome**	**Potential surprise (*y*)**	**Probability (%)**	**Expected net outcome (*Φ*)**
0	-2	$\hat{s}$	5	-0.1
1	**-1**	***r***	**15**	**-0.15**
2	0	*0*	25	0
3	1	*0*	30	0.3
4	**2**	***r***	**20**	**0.4**
5	3	$\hat{s}$	5	0.15
			Expected cumulative outcome: 0.6	
Table 3				

Net outcomes now involve the cost component, implying that some pay-offs are negative (we can assume negative payoffs are undesirable, although this need not always be the case). With the potential surprise and probability distributions unchanged, we see the final expected outcome has been reduced dramatically, from $2.40 to $0.60. This expected outcome is still of no consequence to the agent; it remains wholly inapplicable for our non-seriable, non-divisible example. However, the individual expected outcomes are very much of interest to our agent; they point to which *real net outcomes* they think will be most alluring to them.[202]

By synthesizing our individual expected outcomes onto our original potential surprise function, we can see the outcomes that the agent *knows will possibly occur with certainty* that grasp their attention the fullest. These become the values that arrest the attention of the decision-maker; as the

[202] As Taleb (2004: 183) paints the example:

> Consider a bet you make with a colleague for the amount of $1,999, which, in your opinion is exactly fair. Tomorrow night you will have zero or $2,000 in your pocket, each with a 50% probability. In purely mathematical terms, the fair value of a bet is the linear combination of the states, here called the *mathematical expectation*, i.e., the probabilities of each payoff multiplied by dollar values at stack (50% multiplied by 0 and 50% multiplied by $2,000 = $1,000). Can you *imagine* (that is visualize, not compute mathematically) the value being $1,000? We can conjure up one *and only one* state at a given time, i.e., either 0 or $2,000. Left to our own devices, we are likely to bet in an irrational way, as one of the states would dominate the picture – the fear of ending with nothing or the excitement of an extra $2,000.

highest expected values, they also represent the corresponding *real* outcomes that the decision-maker now hopes to attain. These values – expected values of possible outcomes – correspond to what Shackle (1952a) referred to as "primary focus-outcomes," or the "stimulation function," as shown in figure 4.[203]

Subjective Probability and the Potential Surprise Function

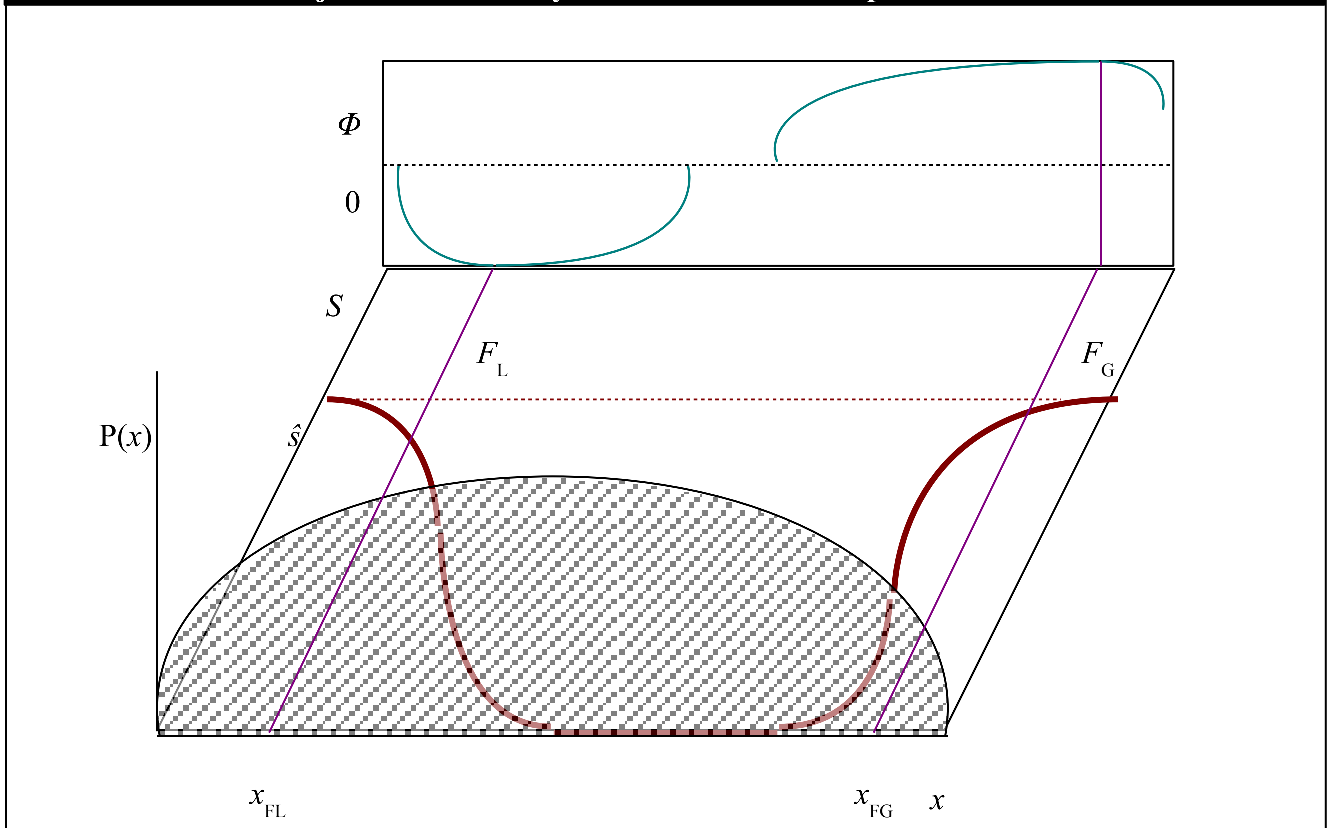

Figure 4

The decision-maker does not concentrate on all possible outcomes, only those which attract the most attention. In our example, we find that the outcome that arrests the agent is not the most probable (as probability is unsuitable for assessing this single trial), and also not a fictional expected cumulative outcome. Instead, we find two separate outcomes consuming the attention of the decision-maker, and

[203] To answer questions as to whether it is realistic to assume an agent focuses their attention on only the most attention arresting outcomes, Shackle (1961: 144) provides the following anecdote:

> A young man is attracted in various ways and degrees by various girls, but he can only experience a genuine and absorbing passion for one of them at a time. It would, perhaps, be conceivable for him to graph the intensity of attraction he felt for each girl as he became acquainted with her, but when he really fell in love, all points on the graph except one would become irrelevant.

acting to form their decision of whether to partake in the example, or not.

The outcome of a loss of $1.00 is the first that needs to be assessed. This is the outcome that is expected to lose the most money on average, but it is not the average loss concerning the decision, but the actual loss that would materialize (x_{FL} in figure 4). The opposite extreme is the outcome which the decision-maker expects to gain the most. In this example, we find one such focal-point. The net outcome of $2 arrests our attention the most on the positive end of the outcome spectrum (x_{FG} in figure 4). This is not due to the fact that it is expected to yield $0.30 over a series of trials. Instead, it is this $0.30 expectation that arrests our attention and makes the agent anticipate the *real* outcome of $2.[204] It is, therefore, the focus-outcomes that are the principle criteria a decision-maker uses to base their decision on, and the real outcome that provides the comparison of risk and reward. The probability distribution, achieved only as a derivative product of the original potential surprise function, is of no direct use in forming the decision whether to undertake a costly action or not. It does, however, reveal to us the outcomes that the decision-maker will focus their attention on the most, and therefore, reveal their decision through the interplay of these resultant *focus -outcomes*.

The notion of comparing two focus outcomes – a focus-gain and a focus-loss – may seem counterintuitive to many modern probability theorists.[205] If decision-making can be reduced to one probabilistically certain outcome the process is simplified and provides a clear answer to whether the actor expects to gain or lose from undertaking the action presented. However, by reducing the process to one number we stand to lose more than we gain. First, it tells us nothing of the risk-preferences of the individual. By removing the comparison of expected, or attention arresting, costs from the same concept as applied to benefits, we may only see *ex post* how much the agent expects to gain or loss. Viewed this way negative expected pay-off decisions seem counter-intuitive. The reduction to one final expected pay-off also bifurcates the issue between seriable and non-seriable decisions. Provided an agent could undertake the action repeatedly with no fear of reprisal (i.e., loss of capital so large as to preclude continued trials) we find a single expected measure may shed light into a long-term achievable pay-off. However, for our unique cases, no such opportunity exists.

---

[204] Meredith (1954: 39) points out that the focus-outcomes (Φ) are a measure of "neural energy," or are "instinctual. The *real* outcomes (x) are, in distinction, the value of the objects under consideration. The focus-outcome is the force that causes us to be interested in the real outcome that can be expected to occur.

[205] To some it may even entail an element of paradox. Shackle (1952a: 76) responds to a criticism of the concept that we surrender our attention to two values which we expect not to occur, or at least would elicit surprise if they did occur. Shackle counters that this is empirically what we observe everyday. We do not expect our house to catch fire and burn to the ground, yet we insure it against this chance occurrence. The reason is that it is not the possibility alone that arrests our attention by the focus-outcomes, rather, it is the expectation of the consequences of the event that will occur if these outcomes manifest – both positive and negative.

Instead the dual concepts of attention arresting costs and benefits must be weighed comparatively. As Egerton (1960: 64) summarizes the trade-off:

> The argument that an investor reduces a range of possible outcomes to two values rather than one is simple. If an investor reduced a range of uncertainty to one value this would imply that he was not concerned with balancing the prospects of gain against the prospects of loss, that he was making his investment decision as if a particular gain were certain.

If only a sole focus-point existed to occupy our decision, an agent could be presumed to invest their capital in the avenue expected to yield the greatest return. Instead, by focusing on our two points, the gain and the loss, we see this seemingly certain investment decision is balanced, or counter-weighted as it were, by the uncertainty provided by the expectation that a loss may also be incurred.[206]

Furthermore, we see that the only time an agent could be neutral regarding a decision is if the focus-gain and focus-loss are equivalent with the propensity for risk of the decision-maker. In this way, the attention afforded by the positive outcome is counterweighted by the attention drawn from the negative outcome. Shackle (1961: 125) provides two such scenarios where a neutral outcome may be obtained. First is a situation where an outcome produces neither "pleasure nor pain" to the imagination. That is to say, the focus-gain elicits no amount of positive anticipation, and the focus-loss produces no feeling of uneasiness. Second is the possibility that an outcome will not change the present situation of the agent. In fact, much like has been pointed out since Wicksell (1898: 104), the "moral expectation" of a profit must be positive for an agent to undertake a decision in the present. In more modern terminology, we may say that an agent's expected focus-gain must outweigh their focus-loss to create an expectation conducive for action.

Egerton (1960: 55) uses the spread between the focus-gain and focus-loss as a measure of the propensity for risk a decision-maker may have. The ratio between the attention-arresting ability of each of the two foci can, in fact, indicate to us *ex post* if the positive outcome was preferable enough to the decision-maker to compensate for the attention given to the negative outcome. However, it must be remembered that although the ratio between the attention-arresting ability of the focal outcomes (Φ) provides the propensity for risk that the decision-maker has, the actual outcomes (x_{FL} and x_{FG}) are the

[206] Of course, more than two points may be the focus of our attention. However, the case of one focus-gain and one focus-loss provides the most simple example from which to start. Further directions with this framework can include multiple foci.

values that they focus the decision on. In our previous example (table 3), although outcome 4 had an expected outcome of only 0.4, the decision-maker can never conceive of the non-seriable, non-divisible trial ever leading to this result. They are certain that the realized outcome will be part of the set defined by the net-outcome column. In this particular example, although 0.4 is the expected outcome of that decision that arrests the most positive attention from the decision-maker, in reality the result is that the net-outcome (x_{FG}) of 2 will be the *real* value captivating their attention.

The use of focus-gains and focus-losses melds together the seen and the unseen expected consequences of our actions. The expected seen consequence is the risk that is accounted for through our subjective probability distribution. The unseen manifests through the uncertainty that shapes our original potential surprise function. Both these elements taken together give a comprehensive insight into the true uncertainty of choice.

Qualitative Concerns in Decision-Making

Noted earlier was the difficulty in modeling qualitative factors in the decision-making process. Although in multiple instances we may perceive the expected pay-off to be probabilistically identical, factors may change the certainty with which we believe this to true. This subjective factor can also be shown to be demonstrated through the model developed herein.

We may assume two separate investments. Investment S is a security, which a decision-maker has formed a potential-surprise function, and corresponding subjective probability function for. We may assume that the expected pay-off of this investment is m%. Investment B is a bond, which the investor has also created the same two functions for, and has also resulted in an expected pay-off of m %. If no heed was given to qualitative factors of decision-making we may forward the conclusion that the investor is indifferent between the two choices; their expected result is identical in each case.

In fact, there are two qualitative factors at work in the decision-making process, factors which may only become apparent once the true uncertainty of the decision is accounted for.

First is the difference between the certainty of the result in each case. Security yields entail much more uncertainty than those of fixed-income alternatives. The result is that the initial potential-surprise function of each case will differ wildly. Fixed income yields can be ascertained with near certainty in advance, provided the investment is held to maturity.[207] Security yields, in distinction, are

[207] At least this is true in strictly nominal terms; inflation will always add uncertainty to the expected real yield. There is also the chance of default in cases of bankruptcy of the bond issuer which must always affect expected yields, and the

highly uncertain. In fact, there is no objective way to determine the future expected return of these investments, even in nominal terms, leading us to recognize the fundamentally uncertain nature surrounding decisions directed toward them. The effect this has on the respective potential surprise function is two-fold. First, fixed-income investments will have a relatively narrow *inner range* concerning the outcomes that will occur with *no* degree of surprise involved. Additionally, the total range of possible outcomes (x) will be relatively narrow. Due to the more certain nature of returns, the range of possible outcomes will be considered quite limited (at least in nominal terms). Securities, on the other hand, will differ in two main ways. First is that the inner range of fully possible outcomes will be wider than that of the fixed-income choice. Multiple outcomes can be considered as fully possible as the increased uncertainty of the decision creates a difficulty in pinpointing with any degree of exactitude a singular outcome considered more possible than another. Second is that the range of possible outcomes will be much wider than a fixed-income issuance. Note that, for example, nominal maximum returns (when held to maturity) are limited by the coupon rate of fixed income choices, while they are theoretically unlimited on securities. As both the range of fully possible outcomes, as well as the total set of merely possible outcomes will be narrower for fixed-income choices compared with those of securities, we see the true uncertainty surrounding these choices manifest in the potential surprise function.

The second qualitative factor determining decision-making under uncertainty is a derivative of the above mentioned one. As the potential surprise function is much more contained with relatively more certain choices than those of greater uncertainty, we find the subjective probability distribution is likewise affected. Remember that the probabilities contained within the inner range of fully possible outcomes will be so similar as to be indistinguishable from the viewpoint of the decision-maker. Hence, fixed income securities will have a flat, or nearly so, probability distribution within the more narrow inner range, as compared with securities. Additionally, the probabilities contained within the inner range will be quite high, reflecting: (1) the certainty that the expected result will attain, and (2) the narrow range of total possible outcomes that the probability distribution could be divided over. In distinction, securities will have a wider range of similar subjective probabilities contained within their inner range of possible results. Additionally, there will be longer tails of relatively insignificant probabilities as their will be many possible outcomes, but each will have a progressively lower

certainty of their attainment. However, even in the case of complete bankruptcy of a company, holders of the fixed-income issuance are given first priority of repayment, making the expected return an *a priori* determinable measure with a high degree of certainty.

incidence of probability attached to them.

Qualitatively, the decision-maker will prefer the focus-points that are located closer to the inner range. The closer to an inner range an outcome value resides, two related factors increase in relevance and significance. First is that the possibility of the outcome's attainment is increased. More possible outcomes will always be qualitatively preferred to less possible, even if the quantitative expected return or degree of attention arresting ability is identical. Second is that the subjective probability of attaining the outcome will be increased. As an outcome arresting our attention the most will be qualitatively preferred if we subjectively attach a greater probability to its attainment, we will be more 'sure' of this type of result.

These two factors cannot be modeled with traditional approaches to probability theory that focus solely on the quantitative aspects of the decision. Furthermore, by examining the ratio between the focus-gain and focus-loss values, we can see that quantitatively negative expected value pay-offs may still be chosen if the qualitative factors assigned to the negative focus-outcome are much weaker than those assigned to the positive focus-outcome.[208] These qualitative considerations do much to rectify the concerns of found in Schmeidler (1989) whereby certain outcomes are preferred, despite identical quantitative expected outcomes, to more uncertain outcomes as the quality of information these decisions are based upon is deemed higher in the first case.[209] Likewise, it agrees with studies such as Ellsburg (1961) and Tversky and Khaneman (1974) whereby a bias exists for decisions where the probability of occurrence is known with a higher degree of certainty, regardless of if the expected outcomes are identical.

Decision-Making and Informational Considerations

Both Hayek (1945) and Mises (1936: 115) stressed the importance of prices as transmitters of information throughout the market. By condensing and collaborating information from many distinct market agents, prices easily and at low cost act as a medium of information exchange among the relevant market participants. As Hayek demonstrated in his now famous “tin” example, information concerning the price of tin (i.e., new demand, new buyers, new sellers, new supply, etc.) became

[208] That is to say, if we qualitatively value a small positive outcome to a larger negative outcome.

[209] Keynes (1921) opened his *Treatise on Probability* by pointing out that the weight we give to a decision cannot be stated in terms of probability. In fact, by pointing out that one argument (or choice) will carry more weight in the decision-making process than another due to the amount of relevant evidence it is based upon, Keynes foreshadowed the work here based on the qualitative effects of decision-making.

extraneous information summarized effectively by the general price. Entrepreneurs need only know one piece of information – the price – instead of the many individual factors that contribute to the price.[210]

Lachmann (1956) demonstrates that entrepreneurs pay heed not to every change in prices, but only to prices which occur between different "bands." In this way, there is a distinction between informationally meaningful price movements, and informationally meaningless price movements. Suppose that there is a good of price *x*. Two bands of prices may be formed around this value, with different consequences for the information conveyed through the respective price movements through the bands' boundaries. The first, or inner band, may be set at *x* plus or minus 10%. The second, or outer band, may be set between the inner band's boundaries, and *x* plus or minus 20%. Lachmann considers price movements within the inner band to be subject to conditions inconsequential for the greater meaning of a price movement. The result is that price movements within this range conveys no useful information to the entrepreneur. Price movements that occur within the outer range may or may not be meaningful. Prices within this outer range are viewed by Lachmann as requiring "supplementary criteria" to determine if they serve an informational role or not. Only price movements beyond the outer range (*x* plus or minus 20% in our example) are deemed to always convey meaning to the entrepreneur. An example of one such conception of a series of price bands is provided in figure 5.

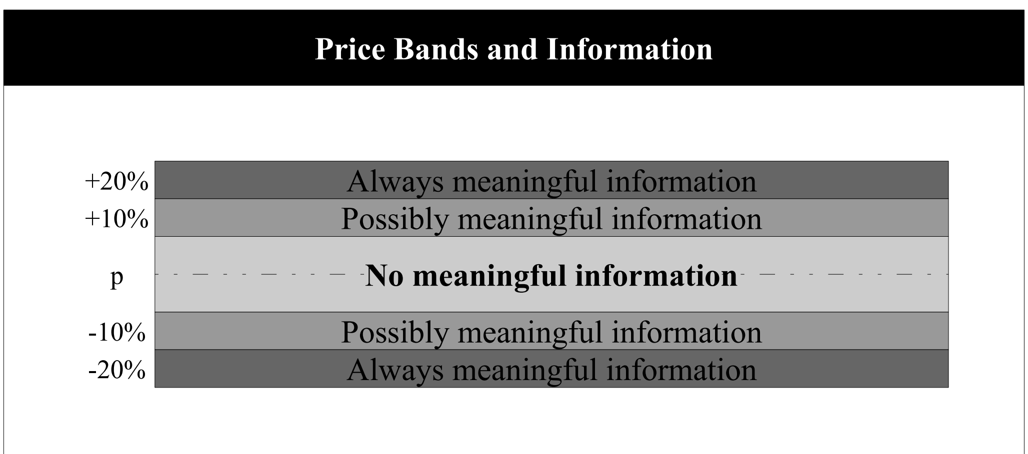

Figure 5

[210] The indeterminacy of knowledge becomes an issue due to the subjective interpretations originating with each market participant. Hayek's early work stressed that there was no way to be certain that subjective data held by a person, or between people, could be identical unless it was sourced from the same objective facts (for example, 1937: 44). However, in light of his later work, we can see that, although it is possible that the same interpretation may result, it is still not predetermined merely by utilizing the same objective source. Hayek (1976: 11) comments that the only thing we share with our knowledge holding counterparts is a general and very abstract concept of the knowledge they have. Lachmann (1956: 70), in distinction, posits that knowledge and expectations can be homogenized through the market – the stock market for example – creating a more definite meaning for transmitted knowledge.

Lachmann provides the basis of informationally meaningful price movements that we can apply to the potential surprise function. The inner range given as wholly possible outcomes are, by definition, informationally indistinguishable from one to another. No new information can be provided from a price movement from the low end of the inner range (x_L), to the high end (x_H), *as long as the end boundaries of the range are not assailed*.

However, contra Lachmann, the outer ranges of the potential surprise function will offer great amounts of information provided prices move into that range. As these outer boundaries include all outcomes deemed to be possible, but not fully possible, that is to say, they are outcomes which entail some degree of surprise once achieved, we gain information when they are achieved as to their true possibility.

Suppose an individual is going to the store to purchase a liter of milk. They are not sure the exact price of the milk at the store, but their reckoning is that values between $1.50 and $2.00 are all equally possible. Values between $1.00 and $1.49, and those between $2.01 and $2.50 all involve some degree of possibility. All values less than $0.99 and greater than $2.51 are viewed as being impossibilities. Further suppose the individual needs to decide how much money to carry with them to the store to purchase the milk. As $2.50 is the maximum value possible in this agent's viewpoint, that is the amount that they will bring.

Suppose once they get to the store that it is found out that a liter of milk costs $2.25. This information is within the possible range of prices, hence the individual had previously expected that the price would occur in reality. But has any meaningful information been learnt here? Lachmann would say not necessarily, that in this range of possible outcomes individual cases and additional information are needed to determine if any meaningful information is provided by a price. In reality we can see a drastic change will result from the information. First will be a shift in the inner range of fully possible outcomes. A new value larger than previously thought entirely possible has occurred. A revision of the inner range of fully possible outcomes will result. This will necessarily transpire in all similar cases where results that are in the outer range of the potential surprise function occur.

A secondary result *may* occur whereby the total set of potential outcomes will shift due to the new information provided by a possible price. In our example, the realized price was in the upper range of possible outcomes. As the inner range will necessarily shift right, it is possible that this may coincide with a shift to the right in the total range of possible outcomes. The result of such an occurrence would be that the lowest possible outcome would be reset higher than originally thought, and the maximum possible outcome may also make a corresponding shift to the right. Of course, it is also entirely possible that no such change in the total range of possible outcomes will occur, or that it will occur on only one end of the spectrum. The existence of a legally mandated price ceiling, for instance, would provide a barrier against expecting higher possible prices despite the occurrence of an event in the upper outer range. Using a price ceiling as an example, we can see that this will not preclude the possibility that the possible range will shift on the bottom end of the range, or the range originally unaffected by the legal imposition.

The aftermath of a price movement beyond the outer range provides an interesting case example that requires deeper attention. To continue our previous example, what would happen to an individual if they went to a store and a liter of milk cost $3.00, much more than the $2.50 previously thought possible (besides the fact that they would find themselves with insufficient money to purchase their milk)? Such an occurrence would far surpass anything the individual had previously thought within the realm of possibility, and will serve to enact the maximum level of surprise possible; it has shocked the agent.

The result will be the loss of an informational reference point to the agent. The information they previously had, and the conclusions drawn from it have proven to be fully erroneous. The arrival of an event previously thought of as being *impossible* has the effect of removing the individual from their existing reference point and into a previously unknown (or at least, unconsidered) territory. If the interpretation of the existing information led to a result that was deemed impossible, we may make two conclusions.

First is that the previous information obtained was incomplete. Of course, this will forever be the case absent conditions of perfect information, as Hayek (1973: 83) noted, "[w]e never act, and could never act, in full consideration of all the facts of a particular situation, but always by singling out as relevant only some aspects of it.". However, it may be said that we need not all the information to make a decision, only a relevant enough amount to approximate the whole information that is possibly available.[211] Some degree of information is needed to approximate a value of an expected outcome. Taking our example of an individual going to a store to purchase a liter of milk, it is not necessary that they know the price of milk at every store nearby to be able to reason an expectation of price. It is, however, necessary to know the price of milk in at least one store, or the price of milk substitutes, or

[211] We see the parallel drawn with the EMH, whereby Fama originally posited that it was not that every agent need agree on future price outcomes, only that a "sufficient number" of them do so (1970: 388). We find that it is not generally necessary to have all the information to determine its relevance, magnitude and importance, but only that a vague *sufficient* amount will suffice in most instances.

any other piece of information that may assist in establishing an expectation of price. An expectation as to the possible outcomes of milk prices based solely upon limited information may lead to an apparent impossibility when the true reality of milk prices becomes known. For example, lacking any information about milk price, but intimately knowing every price of bottled sparkling water, we may foresee our individual approximating their expectation based upon this apparent refreshment substitute. However, if it is determined that the lowest possible price that the agent can expect based upon known sparkling water prices is $3.00 a liter, they will definitely be shocked to go to the store an discover a liter of milk will be less than this, and retail for only $1.75. The previously deemed impossible has become possible with the result that the individual must now reject the previous belief, and update their expectations as per the new information.

A second explanation for an impossible outcome occurring is that information previously attained has been misinterpreted with the result that the range of expectations was incomplete.[212] Bikhchandani, Hirshleifer and Welch's (1992) concept of information cascades comes to mind as an explanation for the erroneous interpretation of information's relevance, meaning or importance. As individuals stop interpreting information and accept instead the interpretation of the crowd, information's quality degrades over time. Prices based on these cascades are quite fragile in nature, as they can potentially be based on very little *real* information.[213] As a result, we see that outcomes beyond an agent's range of possibilities may occur that upset the previous set of expectations.[214] In fact, as Williamson (1975: 5) emphasized that “[t]he 'marvel' of the economic system is that prices serve as sufficient statistics, thereby economizing on bounded rationality.” However, the issue with this viewpoint is that it tells us nothing of how these prices were arrived at (Thomsen 1992: 48). As a result, prices may tell us how prior actors valued the information; they may show us their demonstrated preferences that resulted in prices. However, the prices themselves can tell us nothing of the relative value between the previous agent and the present one regarding the same information. Hence, although

[212] In this way, we are reminded of Hayek (1937a: 33) as he would comment that before we ask why anyone should commit mistakes, we should ask why they would be correct. Likewise, High (1982: 165) notes that it is not relevant for progress whether an individual acts upon information, but only that they act upon the *relevant* information with the *correct judgment* for a beneficial result.

[213] However, there is perhaps little reason why individuals would be mere “information takers.” As Lavoie (1985,:83) alludes to, “market participants are not and could not be price takers any more than scientists could be theory takers... Entrepreneurs (or scientists) actively disagree with existing prices (or theories) and commit themselves to their own projects (or ideas) by bidding prices up or down (or by criticizing or elaborating existing theories.” The corollary that agents become price takers is that they must also be information takers.

[214] As Knight (1921: 201-202) summed the four error sources surrounding the use of information: “We do not perceive the present as it is and in its totality, nor do we infer the future from the present with any high degree of dependability, nor yet do we accurately know the consequences of our own actions. In addition, there is a fourth source of error to be taken into account, for we do not execute actions in the precise form in which they are imaged and willed.”

prices tell us much of what others believe the value of past information to be, it is confined (or bounded) by the ancillary information they may have. This may or may not agree with how the present actor values the information based upon their own knowledge.

In any case, incomplete, asymmetric or highly dispersed information, although creating uncertainty and numerous derivative problems for decision-making, is an important element for the acting individual. As Salerno (1994b: 114) tells us, “dispersed knowledge is not a bane but a boon to the human race; without it, there would be no scope for the intellectual division of labor, and social cooperation under division of labor would, consequently, prove impossible.” It is interesting to revisit Hayek (1946: 95) where he emphasized “the paralyzing effect really perfect knowledge and foresight would have on all action.” By his own reckoning, we see that imperfect knowledge is essential to action itself – action's existence could not be possible without this lack. As Kirzner (1992: 117) points out, in a disequilibrium state the informational role of prices is quite distinct than under equilibrium circumstances. Hence, in disequilibrium, prices spur on entrepreneurial discovery, and therefore, create new information.[215] As the world we act in is necessarily one of disequilibrium, we see the importance of this viewpoint.

Risk and Uncertainty Under Dynamic Conditions

Uncertainty becomes a moot point under static conditions. As no changes will occur in available information, with the result that preferences will remain unchanged, we find that only risk (calculable risk) will exist. However, adding a dimension of dynamism greatly increases the complexity (and realism) of our world. Up until now the model of risk and uncertainty we have been working with has assumed static expectations. We have not explicitly considered the change in expectations that will modify the expected risk and uncertainty inherent in a decision.

In fact, expectations become central to the problems of risk and uncertainty as they can only arise within a dynamic temporal setting. As O'Driscoll and Rizzo (1985:59) state, if the temporal passage is not “real,” no changes will occur and hence, time will not “renew.” This implies that expectations will remain, by definition, static. In fact, a static expectation is the antithesis of what an expectation aims to achieve – an estimation of what the changed future will hold. In fact, as Garrison

[215] For both Mises (1949; 1980), and Hayek (1968), this disequilibrium state would represent a circumstance of *false* prices permeating throughout the economy. It is Hayek's “competitive discovery process” that seeks the necessary knowledge to replace these false prices with somewhat less false prices. As we can never reach a state of full equilibrium, we see that prices will forever have to be considered as *technically false*, or in a state of disequilibrium.

(1997) sees the issue:

> If we think in terms of market solutions to economic problems, we must accord expectations a crucial role. But that role is overplayed if it is assumed that expectations come ready made on the basis of information that is actually revealed only as the market process unfolds; it is underplayed if it is assumed that expectations are and forever remain at odds with economic realities despite the unfolding of the market process. Either assumption would detract from the equally crucial role played by the market process itself, which alone can continuously inform expectations.

We find that to give expectations their due worth, they must be solidly grounded in information and aimed towards definite ends. How does this reconcile creation and transmission of information with the potential surprise function.

When an individual is formulating a place then, we see that they base their expectations upon its eventual obtainment on the perceived uncertainty which in turn frames the subjective risk. However, it should be noted that not just one set of expectations occurs, but a *structure* of expectations extending in time from the present to the date when the outcome is expected to obtain. It becomes clear that as the individual moves temporally through this structure, new knowledge considerations will affect not the original expectations, but the ones expected to prevail on that date where the individual finds themselves at.

Hence, we find values that lie outside the inner-range of expected outcomes, as well as those that are wholly unexpected (fully surprising) alter not only the possibilities that structure the possible surprise function in the present, but also all future structure of expectations that are shaped by the possible surprise functions of the future.

Mars (1951: 2) notes that there are six critical dates concerning the expectation forming process: (T1) the planning date when plans concerning the future are considered, (T2) the decision making date when decisions are made based upon the prior planning date, (T3) the planned action date, when the plan will actually be acted upon, (T4) the outcome of consequences date, when the outcomes will become certain, (T5) the fresh outcome accrual dates, these represent the times where new information becomes available to the agent, and (T6) the revision dates, the dates where new knowledge will lead to a revision of expectations. Hence, we see at the planning date (T1) an expectation is formed as to the future outcomes. This outcome is shrouded in risk and uncertainty until date T4 when the outcome will be revealed with full certainty. In the interim period, new information will become available at T5 which may lead to revisions of the expected outcome.

In fact, not one expectation will be produced at T1, but a whole interlocking sequence of expectations. This *structure of expectations* will provide the agent with an outline of the outcome's progress throughout time as it moves from one state of existence to another. New information in two forms will be continually provided during T5. Some information will become available to the agent as reflected in prices, and some information will not yet be reflected in the price structure.

Information contained in prices can be meaningful or meaningless, as has been previously discussed. The way in which this information affects the price structure will determine whether it gives rise to an expectational change or not.

A decision is made by comparing the relative values of the focal-points. However, once this decision has been made, there are now only two values that matter to the agent – the extremes of the inner range of fully possible outcomes (those where S=0; x_L and x_H respectively) (Shackle 1953: 42-43). Provided that through the time period where the outcome's attainment remains uncertain, that is to say the time between action and realization (T4), the present outcomes remain consistent with the structure of expectations, no changes will be provoked resulting in a change of action. At any given time, therefore, provided the price coincides with the inner range of prices deemed fully possible, no change *can* result to the expectational structure; nothing will have occurred to cause the agent to revise their expectations.

Concerning the structure of expected outcomes, an agent cannot distinguish *in any degree* between those contained within the inner range. However, as the price deviates from within to outside the inner range, the agent is *informed with certainty* that there is a new, previously un-thought of, price which is fully possible; they are seeing it occur before their very eyes. It follows that this will evoke a restructuring of the range of expectations, as well as a reassessment as to the validity and degree of success that the plan is deemed to realize.

This explains the process with which expectational changes occur, or don't occur, in response to information transmitted through the price system. But what of information not yet available or summarized in prices?

Herein we find the essential entrepreneurial role. In fact, finding and using information to make decisions concerning the future is a synthesis of two well-known theories surrounding the entrepreneurial function. Mises' (1949: 252) entrepreneur looks to the future to see what opportunities exist to exploit in the present. Kirzner's (1973: 48) entrepreneur is alert to information that exists in the present that gives rise to the knowledge of these future opportunities.[216] The information that is not yet to be found in the pricing system is attained by the entrepreneurs who decide if it is materially relevant to change their structure of expectations or not. Note that this new information need not have an isolated effect on only one part of the potential surprise function (i.e., the inner range, the value of $\hat{s}$, etc.); it may affect any or all relevant factors.

We find then that changes in the structure of expectations stem from information transmitted through the existing price system, or information discovered and transmitted *to* the price system by the entrepreneur. The former can only affect the agent's decision if the price is outside the inner range previously established for only these prices will elicit a degree of *surprise* that the agent will react to. In distinction, information discovered by the entrepreneur can always affect any of the factors comprising the structure of expectations.

Appendix A: A Closer Look at Impossible Outcomes

By far the most interesting, or at least imperative, outcome concerning the structure of expectations is that which has been considered as impossible prior to its occurrence. As these were fully expected *not* to obtain, a clear signal has been sent to the individual about their existing state, and interpretation, of knowledge. However, there are two comments which should additionally be made.

The first concerns the connectionist reality of the mind. Knowledge in the mind is structured according to its inter-related web of neurons which give strength to some "memories" (or pieces of knowledge) according to the prevalence with which they have connections with other "memories." One implication is that some pieces of knowledge will be stronger, that is to say, they are easier to recall, then others based upon the inter-connectivity with which they exist in. It becomes clear that when outcomes deemed as impossible obtain, it may not be as a result of a *lack* of knowledge concerning them. Instead, an individual may have the knowledge, but has "forgotten" it or at least placed less emphasis on it than they otherwise would.

The second implication, however, concerns the earlier distinction that was made between action and reaction. As reactions occur due to the necessity for the mind to categorize knowledge which it has

[216] Of course, there is nothing incompatible of these two separate concepts, as Sautet (2000: 61) points out. See also Huerta de Soto (2004) for this synthesis.

previously not had possession of, events occur which we may say are not the result of a planned action, at least not in the sense we traditionally like to assign to the word. If an outcome obtains that was deemed impossible previously, not only will the structure of expectations be disrupted, but the manner in which it is disrupted will not necessarily be in a way that is rational (with the definition that we typically like to assign to the word rational). The reason is that rationality is inter-twined with the concept of knowledge. As an outcome occurs outside of our set of knowledge, there is no manner with which we can initially act upon this new knowledge in a rational way.

As no links occur at the moment we acquire the new knowledge, we lack the reference points – those inter-related links connecting the new knowledge to existing knowledge – and hence, find ourselves unable to initially see the causal relationships that result. Phrased another way, no structure of expectations exists concerning new knowledge. We are, as a result, unable to form a rational plan based upon new knowledge until it is first causally linked to other pieces of knowledge in our minds. Once this is achieved, it can be factored for in our future expectations by influencing our potential surprise function.

Impossible outcomes *may*, as a result, coincide with actions which seem *irrational*. Indeed, they actually will be by definition. Lacking a place in the existing structure of expectations, and also lacking knowledge of the causal relationships between this knowledge and other pieces of knowledge which would normally exist inside the mind together, there is no way for an individual to make an action upon this event, based upon a pre-conceived plan. As Shackle (1973: 40-41) summarizes the problem, there is no general theory in economics that will dictate how individuals will respond to disappointments. Indeed, given disappointments must necessarily be not only events that negatively befall upon individuals, but also events which negatively fall upon the individuals *and* were previously deemed impossible occurrences. Events deemed impossible ex ante will fail to be included coherently into the existing expectations and plans of the individual. The prime motive for action will be replaced with that of *re*action.

V. THE ENTREPRENEUR

When we view the world of human action in terms of future uncertainty, as was shown in the previous section, we see that we live in a fundamentally open-ended system; there are endless future possibilities open to us.

As was alluded to previously, the way that actors mitigate the uncertainty of the future is through entrepreneurial action. Future states do not come into existence independently of action, but instead are the created result through the present actions with a look to the future (Kirzner 1985: 56). There are several different visions we can hold for how the entrepreneur achieves this function. Knight's vision of the future as being wholly uncertain precluded the possibility of an entrepreneur moving us into it in any purposeful way. Any action aimed at the future would necessarily be random. Robbins' entrepreneur operated in a closed-ended system, hence they were maximizers of the known scarce resources. This definition is incompatible with the open-ended world we previously described under conditions of uncertainty. A Robbinsian entrepreneur, or maximizer, could never move us to a future state, as all they could do is operate within the context of what presently exists (Vaughn 1994: 141).

Two complementary theories of the entrepreneurial function have been forwarded. Mises' entrepreneur was seen as not being a resource owner by necessity, but as finding the knowledge necessary to move into a future state. The temporal element is crucial to understanding Mises' entrepreneur, as he (1949: 252) defines the role, "[e]very actor is always an entrepreneur and speculator." Kirzner's (1973: 48) entrepreneur is an *alert* individual, one who requires no investment to perform their role. Hence, the Kirznerian entrepreneur does not need a resource to exploit for profitable gain, instead their alertness gives them the ability to realize when such opportunity presents itself to be exploited.

These two separate aspects of the entrepreneur – Robbinsian maximizing and Misesian/Kirznerian alertness – have each focused on one aspect of the problem that must be met. Robbinsian maximizers eliminate, or at least continually trend towards a reduction in, the risk of the market. Alternatively, Mises' and Kirzner's entrepreneurs are concerned with the future state of affairs that will come into being through their actions. The problem that arises when viewing these in isolation is that while the future is the driving force that the entrepreneur is striving towards always, all action that they partake must be done in the present. We may consider the fact that if the entrepreneur was unable to effectively mitigate the risks of the present, they would never arrive at the uncertain future.

The entrepreneur is an element that requires no resources to function – the resources that they use are inherently endowed in them. However, we also find that the pure entrepreneur is able to effectively manage (and hence eliminate) two types of disruptions. First, risk which occurs in the present must be accounted and planned for. Second, the discovery process which prevents inter-temporal disco-ordinations through uncertainty must be sought and exploited. To the extent that an individual can achieve each of these ends perfectly they will be considered a *pure entrepreneur.*

1. The Pure Entrepreneur

Previously, in the first volume of this work we explored some of the conventional theories of the entrepreneur. By looking backward to the history of this function, we have been able to see where the need for this function has come from, as well as limitations that the current entrepreneur places on economic theory. Indeed, the omission of this function has been a latent, if salient, source of error in much modern economic theory.

There are common themes in all these previous definitions of the entrepreneur. However, not one single definition seems to provide a consistent synthesis from which to build from. The fundamental omission of each approach is a pre-defined definition of what it is that entrepreneurship (the process) is, and why it exists.

First, we may wish to outline *why* it is that we need an entrepreneur. It is commonly mentioned that the entrepreneur is the force that drives the market forward. However, it is often glazed over *what* this forward state of the market is. For a market can improve in myriad ways, and each would be grouped together in what is commonly referred to as *moving* the market *forward*. For instance, if a company develops a new product, never offered before, is this to be considered an improvement from the previous state of affairs? Perhaps, provided that it is an improvement that the market desires. However, assume that a company develops a new product, which the market does *and has* proven to desire (i.e., demand and profitability are established), but a second company copies the product exactly and offers it for sale at a lower price than was previously possible. Is this to also be considered moving the market forward? Typically we would be inclined to say yes, although it now obvious that the reason why is of a fundamentally different nature. In the first case the market was improved by moving towards a fundamentally uncertain outcome – a product was produced that was not previously known to be desired. In the second case the market was also moved forward, however it was due to a more risky element – demand was established, and hence, there was no uncertainty as to the market desiring the product. However, there was considerable risk involved with whether the market could: 1) bear two producers of the same product, and 2) whether the second company could 'copy' the product at greater profit than the original producer.

Risk and uncertainty may both give rise to outcomes which are defined as moving the market forward. But before exploring why this is so, it must first be established what it is that we mean when we say the market moves to this "*forward*" condition. This *forward* condition is always conditioned by the fact that consumers' desires are being met in ways that they were not previously. Previous definitions of entrepreneurship have glossed over this fact, and assumed that any change in the welfare of consumers must stem from some uncertain future. However, there is one absolute to consumption that has been overlooked.

The process of consumption always entails a prior act of exchange. This can either occur catallactically – between two individuals concerning goods – or in autarky – concerning a singular person and their preference for a good now rather than later. In either case, the key point is that a renunciation must be undertaken before consumption may be initiated. It is a fundamental truth that what will be renounced will always be that with the smallest opportunity cost to the individual. It follows that consumers will *always* prefer to have a consumption good at a lower cost rather than a higher one. Hence, without knowing anything about *uncertain* future preferences of individuals, entrepreneurs can also increase consumer welfare solely by reducing the cost of the good they provide in the present. If the entrepreneurial process is the process that moves the market forward by increasing consumer want satisfaction (as we have defined it), then we see that this may take place by improving want satisfaction in the future, and by increasing the satisfaction of existing wants in the present – both involve the same end.

The process of moving a market forward towards greater consumer want satisfaction, that is to say, towards a continually retreating equilibrium, need not imply any temporal presumption. This *process* can occur in both ways which are concerned with an unknown future, and also a known to exist, but yet to be discovered, present. This dichotomy of concerns actually entails the duality of risk and uncertainty that has been previously explained. In fact, as we will see, a pure entrepreneur cannot be solely concerned with uncertainty, but must also manage the inherent risk effectively as well.

The Entrepreneur as Risk and Uncertainty Bearer

Haynes (1895) demonstrated that risk and uncertainty need not always be separated. In fact, in the course of many outcomes they become intertwined, inseparable even. Likewise, certainty and uncertainty can manifest at the same time. Haynes' particular example concerned the fact that the outcome "we must all die" is known with full certainty. However, "when we will die" is one of the greatest uncertainties that we can face. It can be argued that death falls into Mises' class-probabilities, and hence, is not a true uncertainty but a risk that may be mitigated away through aggregation (class-probability). However, while this may true from one perspective (i.e., an insurance company aggregating classes of policy holders together) for the individual it becomes a true uncertainty. To look at the issue as Shackle would, the event of death is a non-seriable, non-divisible outcome. Likewise, we have seen that it is certain that we can improve consumer satisfaction in the present, but simultaneously uncertain how this will be achieved.

Similarly, following Knight and Shackle, we see that uncertainty need not only occur due to the element of time. Some situations are ontologically unclassifiable, and hence, even in a static sense uncertainties may exist. What we define as risk may only exist in a statically-dynamic sense, as it pertains to a class which we know probabilistically everything about, while also knowing that nothing will change this probability distribution in the future. Hence, we may apply the risk concept to a temporal setting as we utilize past information (i.e., historical frequency distributions) or logical probabilities. Although an element of uncertainty will still be applicable, in most cases this will be so small as to make the difference between risk and uncertainty moot. For example, determining the side that faces up after tossing a coin once is a static uncertainty, while tossing it many times turns it into a dynamic risk. Conversely, guessing which team will win a game turns into a dynamic uncertainty – we can't say with any probabilistic certainty any time it occurs. Table 4 summarizes these outcomes.

Time Dynamics and Event Types		
	Risk	**Uncertainty**
Static (one-period)	Non-seriable/non-divisible events	
Dynamic (multi-period)	Class Probability	Case Probability

Table 4

In a static, one-period setting, the only possibility that may obtain are non-seriable, non-divisible events. These outcomes, even provided we have logical or historical frequency probabilities that are deemed accurate, still must be treated as fundamentally uncertain events. Shackle's coin toss to decide which team bats first in the Cricket test is a prime example. Alternatively, in a multi-period setting, risk may be attached to cases where stable frequency, or a priori logical probabilities, exist. These fall into Mises' class probabilities. Last, we find that in dynamic settings (i.e., multi-period outcomes) that examples of true uncertainty exist. No logical probability distribution may be obtained, and historically obtained frequency distributions are inapplicable due to changes that exist in the data set. These are classified as Mises' case probabilities.

The entrepreneurial process, once defined as moving the market towards greater consumer want satisfaction, can be achieved through mitigating any one of these three categories.

Non-seriable, non-divisible events represent the traditionally Kirznerian entrepreneurial focus. Opportunities exist in the present, which are waiting latent to be discovered and exploited. However, due to the very nature of these events, they represent a great amount of uncertainty. As there is no method to scale these opportunities, either up or down, there is no method to apply risk to them – they are purely uncertain events. Hence, presently existing disequilibria may be exploited, with the end result that consumers' want satisfactions have been increased. An example of this may be a a neighborhood lacking a coffee shop, and consumers having to travel one metro stop to find their nearest café. The introduction of a new café at the nearest metro stop is not-seriable nor divisible, however, it represents a latent opportunity that awaits exploitation. The *means* of the opportunity exist (i.e., the concept of the café, the demand, etc), but an entrepreneur has failed to capitalize on the need for a closer café than has previously been available to consumers, and hence, increase their want satisfaction.

Risk-based entrepreneurial activity may also increase consumer want satisfaction. As has been noted, it is certain that consumers will always desire their wants satisfied while having to renunciate the least amount of resources for this occurrence. Commonly, we may refer to this in more informal terms as saying that consumers wish to pay the least money for what they purchase. Risk factors can exist in a dynamic setting, one that is marked by repeatable events that give rise to stable probabilities (either through historical frequencies or a priori logic). An example of this may be a production process where the rate of flawed product that is unsuitable for consumption is x%. This risk factor may be reduced, or mitigated, and hence enable a profit opportunity for the entrepreneur through one of two ways: 1) a greater profit margin on sales, or 2) a greater volume of goods sold at a lower price than competitors with a higher cost of risk can offer. Note however that this entrepreneurial profit depends in no way on overcoming uncertainty; a product may have a previously established demand and profitability and hence, only risk is remaining as a residual "cost" of production.

Last, entrepreneurial activity focuses on dynamic uncertainties. These are those that Mises focused on, disequilibria that exist in the future, that need the foresight of the entrepreneur to envision, and hence, move towards rectifying. These are consumption wants that individuals do not even know to exist yet, and the entrepreneur, through the proper foresight of an uncertain (indeed, *unknown*) future, brings to their attention and fulfills a previously *unidentified* want. These may be seen as being owed to the Misesian foresight (i.e., entrepreneurs looking to the future with the eyes of historians) or with Shackleian imagination (i.e., imagining an unknown, but not an unimaginable, future). An example of this type of entrepreneurial discovery is a wholly new want satisfaction, such as that which the introduction of a new product creates (i.e., email superseding telephones, cars making carriages obsolete, etc).

Hence, we can see that the pure entrepreneur, by moving the market towards greater consumer want satisfaction, accomplishes this feat in two main ways – by reducing risk, and by shouldering uncertainty. However, we may also observe that in the absence of continual uncertainty, there is a finite point that risk may be eliminated to. This is so as uncertainty shifts the existing dynamics of the state of affairs, and in so doing, creates continual new, risky opportunities. Uncertainty continually gives rise to new risks, which must be learned. It becomes clear that in a world absent uncertainty, the entrepreneur would soon become eliminated as all risks are discovered, and eliminated.

It may prove instructive to look back at what gives rise to uncertainty in the first place. Previously, we have seen that incomplete knowledge limits the extent to which we may know about a given situation. This impartial knowledge set gives rise to risky situations. However, there are also situations which we know nothing concrete, and these may only be classified as uncertain events. Additionally, events that are non-repeatable cannot be acted on in any certain way, not even probabilistically certain, hence, they too become uncertain situations. Two facts become clear, both with positive implications for the continued role of the entrepreneur. First is that knowledge may never be complete. We have seen that this is due not only to the limited mental capacities that an individual may have (not only for learning knowledge, but also for interpreting tacit knowledge), but also due to the fact that the process of action continually creates new knowledge. As the continuity of action continually adds new knowledge to the realm that we may learn from. This brings the first implication that an individual can never learn *all* the knowledge available. Second is that, even if we ignore this first part and believe for a moment that an individual did know all the knowledge necessary for an outcome, this still would not eliminate endogenous uncertainty. Uncertain events also give rise to the fact that many actions we undertake are non-seriable, non-divisible events. Even given the perfect knowledge that an historical or logical probability distribution existed concerning an outcome, if we were only given one time to achieve this outcome, our knowledge set would be insufficient to remove us from a risky world, and instead we remain trapped in an endogenously uncertain realm.

As a result, as entrepreneurs continually discover new opportunities (both in the future, and in the present), shifts in the risk factors occur which also require attending to in order that consumers' want satisfactions are increased. The entrepreneurial process has no end.

The Pure Entrepreneur

If the entrepreneurial process is one which continually moves consumers towards a point of greater want satisfaction, we see that this achieved in two temporal ways – statically through unexploited opportunities, and dynamically through mitigating risk and shouldering uncertainty. Returning to table 5, we can see that there are three main areas that an entrepreneur must act.

A pure entrepreneur acting to remove all dynamic uncertainty will also remove all static uncertainties as well. As static uncertainty may only arise through unexploited opportunities, we see that an entrepreneur acting in a way that removes all future uncertainties that will also remove all static uncertainties as well. Hence, to be a pure entrepreneur, the focus need not be on exploiting undiscovered opportunities as Kirzner stressed, but rather on exploiting future opportunities that are not apparent yet, but will be eventually. We find that Mises' forward looking entrepreneur is able to eliminate all the subject matter that the Kirznerian entrepreneur typically focuses on, provided they are able to fully foresee and exploit these future opportunities (i.e., they *are* a pure entrepreneur). In practice, of course, we may see that as a pure entrepreneur need not exist – present uncertainties remaining from previously unexploited opportunities will exist, and hence there will be a need for an alert entrepreneur to discover and meet these statically uncertain events.

However, assuming that a pure entrepreneur is able to foresee and meet all future uncertainties, will there arise a need for mitigation of risk factors? In other words, if all uncertainty is removed from a system, is it possible that risk remains? As we will see, risk belongs to a wholly different epistemological category than uncertainty, and as such requires a different method to manage it. At the same time, however, we see that risk represents a fully-exploitable set of opportunities. Risk may only be seen to continually prevail in the presence of continued uncertainty. This arises as risk exists only within a closed system – events of a set must be defined and not subject to change. It then becomes evident that risk may *only* exist in a dynamic sense for two reasons. First was the necessity of stable probability distributions which by definition require repeated trials for the arrival of pre-determined expected outcomes. The second becomes clear now as we find that risk will be fully exploited in any static period as the entrepreneur fully mitigates the risk. The continuance of this factor may only occur with a maintained temporal passage unveiling fresh uncertainties, and their associated risky profiles which have yet to be exploited.[217]

However, as risk need not rely on uncertainty, we find that even in a fully foreseen uncertain environment, there still remain elements of risk to be exploited (prior to being fully exploited and removed from the system). This distinction is more apparent in Mises' case/class probability distinction. Case probabilities, those we normally refer to as uncertain events, belong in an open-ended world where continuously created unique events cannot be aggregated and acted upon as a group. With class probabilities, those normally referred to as risk, relying on a pre-defined homogeneous class of outcomes, we find that this risky element can exist even in the presence of a fully certain outcome. For example, it may be fully certain that an outcome will obtain, but risky as to when this will occur.[218] In our current world, it is fully certain that everyone will face death eventually, however, this does not inhibit a developed and complex insurance industry from existing to continually mitigate the risk surrounding these certain events. However, what gives rise to the fresh risks that make the insurance industry dynamically sustainable (i.e., profitable) are the continued sources of uncertainties that temper the risk factors determining when an individual will die.

Therefore, two different scenarios arise which will result in differing entrepreneurs.

In the first case, we may look at the more long-run case where future uncertainties are fully discovered by an entrepreneur. However, in the short-run there will continue to exist risky factors requiring exploitation. These will only exist until one of two uncertainty-based events occur. The first is that risk will only continue to the point where a fresh uncertainties present themselves. Second is that risky factors will only exist until they are fully exploited by other entrepreneurs. It follows that even in the presence of perfect entrepreneurs and a dynamic setting, there will continually be risk-factors in need of mitigation as the exploited uncertainties give rise to new, previously unknown risk factors. The disappearance of an entrepreneur focused on risk mitigation could only occur lacking fresh uncertainties – it is this factor which, even provided entrepreneurs fully foresee these uncertainties – gives rise to risk-factors that were previously not accounted for.

However, much like we have previously seen (see *The Infinitude of Uncertainty*) it is not possible to speak of differing degrees of uncertainty. Prevailing at any time will be *uncertainty*, but

[217] We are reminded of Shackle's (1972: 422) criticism of game theory on the grounds that it assumes away the very element that gives rise to its sustained risky possibilities – uncertainty.

[218] The classic Haynesian example with death being a certainty for all, but he time of death being one of the most prevalent risks that we face.

never a specific or measurable (i.e., comparable) amount thereof. One implication of this is that risk factors would have to be made apparent, despite the presence of a perfect entrepreneur foreseeing all uncertainty. Even in the presence of such perfect foresight, fresh uncertain factors would continually arise which, although being perfectly foreseen by themselves, would provide a cause for new risk in other avenues.

Up to this point, we have been implicitly assuming that the risk-mitigating and the uncertainty-bearing entrepreneur are two separate people. In reality, the *pure* entrepreneur will be one who fulfills both of these roles and is able to complete both perfectly. The future will be fully known in advance to this pure-entrepreneur, thus eliminating all uncertainty surrounding these yet unknown events and eliminating static uncertainties from occurring. Additionally, the risk of the present will also be completely known, and continually adjusted for as the changing (but fully foreseen) future calls for. A pure entrepreneur will, therefore, create the future as they see it to be created, and they work toward eliminating presently occurring disquilibria through the elimination of risk (as it has been fully foreseen by the now certain future uncertainty).

Qualitative Aspects of the Entrepreneur

It follows that the three-tasks of the entrepreneur involve different skill sets which may not be endowed in any one person simultaneously. In fact, given the heterogeneity of humans, it will become apparent that some individuals will excel at some entrepreneurial aspects, while others do so in other avenues. The fact will remain that, provided an individual acts in one of these three manners – static and dynamic uncertainty bearing or risk mitigation – they will be moving the consumer towards greater want satisfaction, and hence, engaging in entrepreneurial activity.

Kirzner's stress on alertness and the exploitation of opportunities manifests as one skill-set needed. As choices made in the present which lead to non-seriable, non-divisible outcomes will not be repeatable, entrepreneurs will need two qualities to successfully undertake this task. First is the aforementioned alertness to discover these latent opportunities. More importantly though is the judgment and foresight to make *correct* choices surrounding their implementation. It would benefit none if the existing disequilibria were discovered, however, lacking the knowledge or judgment of how to *correctly* manage these events, no shift towards greater consumer want satisfaction would occur.

Risk management implies an entrepreneur qualitatively similar to a Robbinsian maximizer.

Working within a defined means-ends framework, this entrepreneur is concerned with computing the probabilities giving rise to risky situations, and act upon these accordingly to bring greater *static efficiency* to the market. At the same time, this risk management involves assessing what the present means-ends framework is. Correct risk management will depend largely on a faithful assessment of the present conditions prevailing and the resources available to be used for this end.

Last, Mises' entrepreneur with their forward looking ability to see the future comprises the last characteristic we find defining an entrepreneur. Looking into the unknown and then envisioning the future that is yet to be created, the entrepreneur acts towards building this envisioned future. Again, it is not only enough to to foresee the future, but to see the future that other individuals in the market value more highly than the present. By projecting in their mind a future condition for the market to reach, entrepreneurs must correctly judge the *possibility* that such a state will be valued once attained.

Previously it was demonstrated that the third quality of an entrepreneur, if perfectly exercised, would eliminate the need for the first quality. However, as not all are created equally, we find that some individuals will perform certain aspects of the entrepreneurial role better than others. This gives rise to the necessity that all three roles will be undertaken simultaneously by various individuals within society to move the market towards a state of greater consumer want satisfaction.

Entrepreneur as Resource Owner and Profit Earner

Little attention has been heretofore paid to entrepreneurs *vis-à-vis* the possibility of simultaneously being a resource owner. As entrepreneurial activity has been defined as moving the market towards a state of greater consumer want satisfaction, it is apparent that the entrepreneurial function cannot be a *mere* mental game, but must also involve physical actions which improve the conditions of ancillary market participants as well.

However, a pure entrepreneur can be defined in terms wholly separate from that of a resource owner. Pure entrepreneurs will have *both* perfect foresight for uncertainty *and* perfect risk-mitigation abilities. A consequence of this is that, by definition, they will not err. The issue that concerns having an entrepreneur that lacks resources is that this factor eliminates losses from befalling them. With no logical corollary to profits, the plausibility of the entrepreneur's role comes into question. However, if by definition we eliminate the possibility of losses from occurring, we find no significant issue with a *pure* entrepreneur not requiring ownership of the resources they use. Kirzner's entrepreneur, focusing on the mental qualities of perception and alertness, failed to recognize the physical necessity of resource owner to further the entrepreneurial process. Faithfulness to the entrepreneur as an *individual* has sacrificed attention on the *entrepreneur process*.

As was outlined in the previous section, there is significant reason to believe that a pure entrepreneur, while theoretically feasible, is a practical impossibility. For this reason, we find in reality that an *essential characteristic* for identifying a true entrepreneur is the use of owned resources coupled with any of the previous three qualities. Only this possibility will give rise to the chance of entrepreneurial loss, a logical necessity as a corollary to entrepreneurial profit.

Resource owners become a separate class of individuals in the market. As they are the ones who ultimately bear the loss or earn the entrepreneurial profit if their judgment proves incorrect, they are defined separately from *employees* of a company, who earn a fixed wage and have no need to be resource owners. What seems to be the normal appearance of profits to this group are, in reality, merely an expense to the true entrepreneurs. Likewise, labor in the form of hired employees cannot belong to the entrepreneur class, but is rather to be considered as an input factor in production, much like capital, land or any other number of real inputs. It follows that one resource which an entrepreneur may typically hire is labor, which is a distinctly different class of production-factors, despite having similar prima facie features (i.e., both stem from an individual).

Entrepreneurial profit, which is typically difficult to conceptualize and distinguish from wages or return on capital, is that profit which stems from the excess return made off real resources, *and* paid to resource owners, through activity which pertains to any of the previously mentioned three avenues of entrepreneurial activity. It cannot, however, be confused with such concepts as return on capital, or similar measures. The concept of entrepreneurial profit must focus on all inputs that are *combined* to increase consumer want satisfaction through a *specific* end. Therefore, not just one input factor can be used to determine the resultant profit that the entrepreneur receives, but rather, *all* inputs must be considered together for this possibility. Likewise, this profit may only accrue to those individuals who have an actual ownership stake in the used resources. To the extent that the entrepreneur *qua* resource owner earns a profit or loss on all used resources beyond that return which would have obtained regardless of this process (i.e., interest, as we will see in the next section), this may be referred to as pure entrepreneurial profit or loss.

VI. TIME

The concept of time is the gravest issue concerning the economist. From this seemingly humble topic stem all the issues that become the primary fodder for economic research – knowledge problems, allocative issues, or uncertainty, as examples. A severe misconception is prevalent as to the true nature of time as it pertains to acting humans. Economists are spry to recognize the issue, prima facie at least, but fail to recognize the true nature of the problem.

Take this well known quote from Alfred Marshall's (1890) *Principles of Economics*: "[T]he element of time [is] the source of many of the greatest difficulties in economics." However, despite enjoying a prominent role among *perceived* economic issues, the concept has been misinterpreted by much of the profession. Much of this perversion has occurred as a result of economics as a subject shifting from a basis of verbal, to mathematically formalized logic. Take for instance the conception of time as Paul Samuelson recounted it: "When a mathematician says '*y* rises as *x* falls', he is implying nothing about temporal sequence or anything different from 'when *x* is low, *y* is high'" (as quoted in Robinson 1980: 220). For many economists falling Samuelson's lead, there is no significant difference between the approach of the mathematician and the economist.

However, the true nature of time is not subject to the strict formal logic that is *de rigeur* today. Linear Newtonian time is wholly inapplicable to the realm of human action. Bergson's (1889) concept of "*la durée*" or real time is the true concept we need to apply when analyzing acting humans. It is within this less linear conception of time that we must view our actions.

A secondary issue arises with the concept of time preference, and its close market manifestation: the rate of originary interest.[219] Much has been written in defense, and contra to, the idea of time preference as being the cause of the originary interest rate. As the essence of time is widely misunderstood, much of the literature dealing with this associated topic has also been deficient.

Time preference, the bias towards the current enjoyment of wants to later enjoyment, holds true *in all circumstances*, but not for the previously given reasons. Instead, we find that time preference is a categorical part of human action due to one of the proofs that make us human: life. It is our awareness of the limited time we have in life, coupled with the unlimited nature of our wants, that gives rise to the want for present satisfaction of wants over that in the future. Limited time, and our human appreciation of this fact, brings value to our individual portions of time.

[219] Originary interest will be used as the concept of interest, and not its market manifestation – monetary interest – throughout this work.

This paper will commence with a brief overview of time, as well as a comparison between the two viewpoints of it held today – Newtonian and Bergsonian. The concept of time preference will be assessed, with differing opinions both pro and contra given. We find that the Böhm-Bawerkian conception of time preference as the root of originary interest fully supported by the present evidence, however, the assertion that this is due to the need for consumption to live (as forwarded by Mises, and accepted by many followers) is found lacking. The alternative rationale is provided to support, and indeed strengthen, the case for this idea.

The paper will conclude with an overview, and an outline for a new basis for establishing originary interest through time preference based on Bergsonian time.

1. The Nature of Time

Two Fathers of Time: Newton and Bergson

Typically, as our earlier quote from Samuelson alludes to, time is viewed in a very mechanical, *Newtonistic*, sense. The passage of time is seen as existing on a one-dimensional plane, with movements from one temporal state to the next viewed as linear progressions from one point to the next. Conceived this way, one unit of time is equivalent to any other. What is true of time at a single point in time must be true for all points on the continuum. The value, and hence, utility, of time is constant – *they are time-invariant*.

This second viewpoint, epitomized in the work of Shackle, follows Henri Bergson's idea of duration (*la durée*). Bergson (1889: 76) set out to distinguish between two similar concepts; "time as quality and time as quantity." Time represents not a measurable entity; its continual flux implies that by the time an attempt to measure its existence is undertaken, the moment of time in question has passed. Hence, instead of existing as a one-dimensional point, we can think of this real-time as existing *outside* of a measurable area. Time exists, not in a physical space but only as *the passing of time*. As Mises (1949: 100) would describe this process:[220]

> The 'now' of the present is continually shifted to the past and is retained in the memory only. Reflecting about the past, say the philosophers, man becomes aware of time... Time as we measure it by various mechanical devices is always past, and time as the philosophers use this concept is always either past or future. The present is, from these aspects, nothing but an ideal boundary line separating the past from the future. [221]

Čapek (1961: 36) points out the three main aspects of Newtonian time. The first is that, as time is independent of its content, if changes occur in time they must occur at the beginning of time. Time itself adds nothing to induce change. Second, time is infinitely divisible. Hence, points on a time-line can each be broken into smaller points, with no adverse consequences for the effect of time (it is

[220] Regardless of the fact that time has no explicit cost associated with it, it remains forever a factor of every decision undertaken. Mises (1933, 176) stressed this point as a failure of "objective theory of value" economists in recognizing the important role time serves in our all actions.

[221] We also realize the passage of time through our expectations necessarily based on an uncertain future.

scalable). Last, time is homogeneous, it elapses without anything happening. The effect is that this conception of time is the antithesis of time as experienced by humans – *it is static*.

In contrast, Čapek (1971: 90) points out the three defining characteristics of real, Bergsonian, time. The first is that the passage of time is a source of change, and hence, originality. Time is hence a powerful source of change. Second, time is heterogeneous. Our memories act as the link between the past and the present. As our continually expanding past augments our memories, our perceptions of time are relentlessly being altered. Last, time is not independent of its content, nor can points in time be viewed in isolation of others. Past, present and the anticipated future are all linked together through our memories, experiences, and expectations.

Table 6 contrasts these two conceptions of time, as summarized by O'Driscoll and Rizzo (1985).[222]

	NEWTONIAN	**BERGSONIAN**
Uniformity	Homogeneous: Can elapse without anything happening.	Heterogeneous: Memory of experience links the past to the present.
Continuity	Mathematical continuity: Continuously divisible. Instants are independent of each other.	Dynamic continuity: Time is not isolated from other points in time.
Effect	Causal inertness: Time is independent of content.	Causal efficacy: Elapse of time is a continual source of change.

Table 6

It becomes clear that Bergson's temporal conception is the one applicable to the realm of human action. We find two different worlds pertaining to time: one characterized by the movement through space, the other the consciousness of sensation and the existence in time (Bergson 1896: 267). Humans do not, and cannot, live as automatons, unaware of the passing of time. Instead, we live and act fully conscious of its presence, and the implications this holds for our actions. There are several additional comments that must be made on the topic of time before we move on to looking into how this element effects our actions.

Four Timely Comments on the Uniqueness of Time

222 In a similar vein, Shackle (1958) defines two types of time: outside and inside. Outside time is similar to our Newtonian viewpoint, and is applicable to the physical sciences. Inside time is subjectively experienced, and is based on the solitary experiences of each individual.

Was – now – will be. Humans must act in the *now*, directing our intentions towards an unknown state that *will be*. At the same time we have only knowledge of what *was*. *Now* is the joint that connects what *was*, with what *will be*. The true nature of the problem is what implication for action time has in the fleeting moment – the *now*. It is this concept of *now* which Shackle (1958: 13) referred to as the "*moment-in-being*, which is the locus of every actual sense-experience, every thought, feeling, decision and action." There are four key details of time that make it a unique element for humans to negotiate.[223]

Time, paradoxically, represents an element of simultaneous finitude and infinitude. In a general sense, time exists without end. *Now* is merely the leading edge of time, which fades into an ever lengthening past, and occurs prior to an everlasting future. Any concept of time that we hold negates the non-existence of time. However, for the acting human, we find that time is a distinctly finite element. Action takes place always within a definite (if ill-defined) temporal period. One cannot undertake an action thinking it will continue forever – *the realization of time's extent is a distinctly human element*.

Time's passage implies a degree of uncertainty.[224] A world of full certainty forever alludes us as our transition through time carries us to fresh, unknown, temporal viewpoints. This transition is what Shackle (1958: 15) referred to as the "dynamic movement in time." All actions are affected as one situation is translated to another. Our memories of the past affect the actions of the present and shape our expected future.

We can always attempt to affect our future temporal allocation by some action in the present, but can never be assured with certainty that this will be successful. Unlike physical goods, we can never purchase additional time with certainty that we will be allowed the utility of it. This uncertainty is not a wholly Knightian fog that engulfs us, instead, we mitigate this element through entrepreneurship when dealing with the future. Complete certainty is incompatible with any concept of time, but a lack of any certainty (a state of pure uncertainty) is incompatible with any concept of action (Shackle 1961, 3-7; Mises 1949). We find that actors must continually navigate their way through this foggy, uncertain future, unaware of every possible eventuality, but mitigating it as best they can.

---

223 We find some attempts at de-emphasizing the past, instead giving more importance to the Shackleian "moment-in-being." See, for example, Heidegger (1953: 376) as he distinguishes between the "now, the "then" and the "now not yet." Less emphasis is given to the the "was" and more to the "meanwhile," or the moment in passing. When the memory of an actor links the past to create their expectations of the future, we see the important role the temporal past serves in the creation of the future.

224 Garrison (1984) points out that the corollary need not be true. Uncertain situations are not solely caused by the passage of time, hence, although time may be said to breed uncertainty, uncertainty is a poor proxy for time.

The novelty, or uncertainty, of time, as Rizzo (1994: 113) reminds us, is solely due to time's relation to another time. These concepts can only occur in reference to another point or period of time. Newtonian time erases the possibility of any novel freshness occurring, as a reference point is erased due to the solitary nature in which time is viewed as. In contrast, real-time, being forever linked to every other point of time in some way, is a constant stream of fresh, and novel, experiences.

Time exists as an irreversible flow. The existence of time is a flow, moving from one point occurring temporally prior to one occurring temporally more distant. The reversion of time is an impossibility. A divide is created between the physical and the temporal realms. Physical occurrences are generally able to be reversed, but a temporal event will forever remain irreversible – what occurs in a moment of time remains in that moment forever.[225] Non-repeatability of events becomes an essential element of time (O'Driscoll and Rizzo 1985: 78).

Time remains a purely immeasurable entity, its *mere* existence negates attempts at concretely defining and comparing its component units. Actors realize time exists as their continual existence is completely engulfed by it. For the actor, time only occurs concurrent to action. As Lachmann (1977: 85) points out, "[t]ime and knowledge belong together." Furthermore, as Shackle (1972: 156) reveals, "[s]o far as men are concerned, *being* consists in continual and endless fresh *knowing*." The trinity that exists between acting humans, time and knowledge becomes apparent.[226] The apparent complication that confronts us is that each element of time will necessarily be experienced under different knowledge conditions. It is never time we wish to compare directly, but an event that transpires at a point in time. As these actions can never be viewed through equivalent eyes due to our changing knowledge base, it becomes apparent that these temporal elements also defy direct comparison, and measurement.

Time's existence is a continual one. In an abstract sense, we can see that the one element that links all actors together is the ceaseless flow of time we must all experience. Curiously, it is this endless passage of time that becomes the only constant to human action.

Hence, we see these four qualities of time all serve to make its existence particularly unique: it is simultaneously infinite and finite, it is an irreversible flow, it defies direct measurement, and it is a constant presence. The questions, and answers, of time cannot be viewed in isolation. Future events, the events which concern the acting human, occur only during a distinct temporal passing. Despite what

[225] Jevons' adage may have been "bygones are forever bygones", however, as we will see shortly, the *memory* of the past has a significant bearing on our perspective of future time. In fact, Rizzo (1994: 113) owes the "survival" of the past – memory – to the fact that each individual, while never at the same juncture of history, is never wholly separated from it.

[226] See, for example, Knight (1921: 313): "The existence of a problem of knowledge depends on the future being different than the past, while the possibility of the solution of the problem depends on the future being like the past." Or Lachmann's (1959: 73) contention that "as soon as we permit time to elapse we must permit knowledge to change."

some of Newton's followers may think, time for humans can only exist as a passage, never a static point. A constant renewal of time must occur lest “real time will cease to be” (O'Driscoll and Rizzo 1985: 59). Any meaningful analysis of time must then take into account these unique factors of the Bergsonian passing of time.

2. The Preference for Time

A Time Preference or a Time Law?

Time, while encompassing all human action, becomes truly important when the topic of interest rates arises. Many economists, starting with Böhm-Bawerk (1889), contend that a tendency exists whereby wants are preferred to be satisfied in the temporal present over the future.[227] This time preference gives rise to originary interest, and hence, the derivative monetary interest rate that we see in financial markets. This is a somewhat controversial topic, with some economists arguing against this as anything more than a tendency. More recently, Hülsmann (2002) has argued that it is not time preference which gives rise to originary interest (hereafter referred to solely as interest), but the value spread between ends and means. We will now provide a brief overview of time preference theory

The concept of time-preference essentially begins with Böhm-Bawerk (1889: 259) as he defines the "nub and kernel" of the interest problem as the fact that: "*Present goods are as a general rule worth more than future goods of equal quality and quantity*."[228] It was argued that human have a preference for goods in the temporal present over the equivalent good in the future. To explain why there exists a preference for the discounting of future value Böhm-Bawerk presented three rationales. First was due to the tendency for wealth to increase over time, thus making the marginal value of a present good higher than its respective marginal value in the future. Second was a psychological factor, whereby there is a systematic discounting of future want satisfaction. Last was a technological factor. As more time consuming "roundabout" production processes were employed, present goods possess a higher value due to this increased output potential.

In fact, we can see that for Böhm-Bawerk the true source of the value differential is the growth of future goods as they mature into present goods, as he himself realized (Böhm-Bawerk 1889: 337). However, we see several problems with his three justifications. His third reason – more roundabout productive processes – has been rejected almost unanimously by Böhm-Bawerk's followers (see Fetter 1902; Mises 1949: 485-486; Rothbard 1962: 424n27).[229] His first two reasons, one based on a *tendency*

[227] The pure time preference theory of interest has since been supported by Mises (1949), Rothbard (1962), Block (1978), Garrison (1979), Kirzner (1996), Lewin (1999), and Hoppe (1999). For interesting critiques, see Moss (1978), and Pellengahr (1996). Time preference theory also appears under other guises, such as Fisher's (1930: 66) "impatience theory of interest", or Čuhel's (1907: 307) similar concept called "egence."

[228] Although, it should be noted, Menger (1871: 156) essentially lays the foundation for the concept.

[229] Interestingly, Böhm-Bawerk (1884) was concerned with refuting the old productivity theories of interest, which is essentially the basis for his third cause of interest (see Fetter 1902 for this critique). For a defense of Böhm-Bawerk's

and the other a psychological factor also fail to apodictically provide justification for time preference. In fact, Böhm-Bawerk's stress on the "general" case for preferring present over future goods led him to the conclusion that time preference was not universally positive.

Mises built upon Böhm-Bawerk's concept of time preference while providing a more satisfactory justification for the occurrence. Hence: "Other things being equal, satisfaction in a nearer period of the future is preferred to satisfaction in a more distant period; disutility is seen in waiting" (1949: 480).[230] Satisfaction for Mises (1961: 118) could only be sourced from ends, never means. Mises grounds his root cause of time preference on *the* ultimate end – consumption. As humans require consumption in the present for continued survival, the tendency becomes necessary for the preference of present over future consumption. The praxeological fact that consumption is necessary to life ties together the consumption-theory of time preference with originary interest. For Mises, any form of consumption is indicative of time preference – only if an individual lacked any preference between now and later would there arise a case lacking consumption. As all humans are assumed to want consumption, the rate of originary interest is *always* positive – we could never want satisfaction of our consumption needs later to sooner.[231]

However, again we find that this justification is not absolute, but of a special case (Hülsmann 2002: 80). As Hülsmann elaborates, warriors, martyrs, or people pursuing suicide all provide examples where the drive for human life is less apparent, or wholly neglected, at the expense of some other overriding end. Mises anticipated these counterexamples, but downplayed their significance and dismissed them, almost sleight of hand.[232] The existence of these eventualities implies however that no absolute theory based upon human action can ignore their existence.

Of particular troubling consequence for both Böhm-Bawerk and Mises' theories of time preference is the confusion between identical *physical* products at differing time periods, and identical

third cause of interest, see Murphy (2003: chap. 1). Interestingly, Fisher was an early advocate of Böhm-Bawerk's third cause, explicitly attributing his want to write *The Rate of Interest* (1907) to expand upon these technical causes of interest (see Fisher (1913: 610)). However, his tone would change shortly thereafter, as he drops technical factors (i.e., physical productivity) as an influence of interest in favor of his "Impatience Theory of Interest" (see Fisher 1912: 371). Impatience was, for Fisher (1930: 66) "synonymous" with time preference.

230 As Lewin (1997: 146) reminds us, there is considerable ambiguity as to what exactly Mises posits is preferred in the present. *Human Action* uses many different words, almost interchangeably, to describe time preference: consumption, satisfaction, gratification, and enjoyment. Garrison (2002) reckons that Mises, with his ever-present emphasis on action, implied action is preferred in the present to the future. This is best demonstrated by Mises' section in *Human Action* entitled "Time Preference as an Essential Requisite for Action." Kirzner, whose views essentially shadow Mises on the subject, words the problem as the "preference for the achievement of goals sooner rather than later" (1993, 178).

231 Lachmann (1956: 78) agreed that originary interest would always remain positive, justifying it as the temporal flow always moves forward, therefore, capital stocks can forever be carried forward, but never backward through time.

232 See, for example, Mises (1949: 490). Also, Mises (1961: 121) realizes that time preference based on consumption could not be an absolute law, but only that "the vast majority prefer life to death and wealth to poverty."

utilities at differing time periods. Böhm-Bawerk originally realized this and set out two different definitions based upon these differing concepts (see Murphy 2003: 65-67). However, Mises seems to confound the issue, shifting between intertemporal comparisons of utility and physical product as if they are interchangeable ideas; *they are not*.[233] As Fetter (1914a: 857) made clear: "A theory of interest must be *essentially* a value-theory."

Human action is concerned with using the means available in the present, to satisfy a want at some point in the future. The comparison of physical goods at two different temporal vantage points is not possible as it compares two different, directly incomparable, concepts: means and ends. It was this realization that Hülsmann (2002) set out to discover the true cause of the interest causing value spread.

Ends, Means: Sources of Interest or Profit?

Hülsmann (2002) strikes out to phrase the choice as not between present and future satisfaction, but between means that move us closer to satisfaction, and the ends that directly provide it. When viewed this way we are inclined to view "[*o*]*riginary interest* [*as*] *the fundamental spread between the value of an end and the value of the means that serve to attain this end*" (87: emphases in original).[234] A caveat of this value spread is that means and ends are either distinct entities, or of distinct people. Two cases where the means and ends are identical but belong to, (a) the same person, and (b) different people, can be assessed. In fact, Hayek (1928: 192) would foreshadow this viewpoint of the value spread between means and ends as the cause of interest.

The first case arises from the case where means and ends coincide – a person singing and

[233] Fetter (1915: 245) clearly delineates the difference between physical and valuational change. However, Fetter focuses on valuational change that still stems from a physical change (i.e., the ripening of wine) and not with the strict valuational change that occurs with the passage of time due to altered wants. Although originally being critical of technical productivity as a source of interest, he would later recant, at least "*some* influence" on the temporal comparison of satisfaction by the a general increase in physical productivity (see Fetter 1914b: 247).

Mises also experienced a change in opinion sometime after his *Theory of Money and Credit* (1912). In his preface to the second German edition (dated 1924), he notes that interest falls outside of the scope of indirect exchange. Obviously, the attribution of this is that intertemporal comparisons of goods cannot be made, only their resultant utilities as compared through value (see esp. page 34). Unfortunately, the comparison of present and future goods has been established widely as the basis for time preference. See particularly Hoppe (2001: 2) for this error.

[234] Hülsmann takes care to note that money and originary interest are two categorically different occurrences. Monetary interest may only occur where the means and ends are physically homogeneous so that a quantitatively definable spread can be determined. Hence, "money interest is [not] something like 'originary interest become visible'" (2002: 93). Instead, it is due to the existence of originary interest that we can determine the source of the money interest. Likewise, although originary interest must be positive (ends must be valued more than means), monetary interest may be negative. For example, a businessperson investing in a hospital may invest $100 and purposefully reap only $90 of profit, thus earning a negative return. This negative monetary interest rate occurs independent of the positive originary interest rate.

hearing their song for their own pleasure, for example. In a *Hülsmannian* world, there is no value spread (originary interest) between the means and ends involved in this action – they are one and the same. It follows that the case for interest can only be made when humans engage in labor. Hence, originary interest no longer remains a category of human action, but instead is relegated to the realm of human labor.

Hülsmann (94-95) then gives an example of originary interest stemming from an example of coinciding means and ends without a separate temporal element, like in the previous example. If a person sings a song, and another hears it, a value spread occurs between the means (the person singing) and the end (the person hearing). That these two occurrences are simultaneous yet a value spread occurs regardless naturally leads Hülsmann to the assumption that originary interest is independent of time.

Hülsmann (108) summarizes his position thusly:

> Originary interest does not spring from the passing of time, but from the value relationship between means and ends. The means of action are inherently less valuable than the ends they serve. Therefore there is a value spread between means and ends – originary interest – in all human actions in which means and ends can be distinguished.

What Hülsmann has elucidated is a *mere* tautology. That ends are valued more highly than means is implied in the fact by that we use means solely for the attainment of ends. As Mises elaborated, we cannot think of an action that would occur at a psychic loss, one stemming from valuing the means used greater than the ends attained. Menger's imputation theory of value tells us this as well. That value is imputed from ends back to the means that contribute to its attainment, we find that the component means can never amount to more than the expected value of the end, in an ex ante sense.

Hülsmann's main error has occurred in the confounding of two similar, yet distinct, concepts: entrepreneurial profit, and interest. As Hayek (1941: 38) clarifies the issue: 'It would perhaps be more correct if we referred to this difference between cost [factor prices] and prices [output prices] as profits rather than interest."

Pure entrepreneurship can be seen as a strict arbitrage requiring no resources (see Kirzner 1973; Huerta de Soto 2005). Thus, when Hülsmann writes of means and ends (of two separate actors) yielding originary interest without a temporal element, the true result he is describing is the source of true Kirznerian entrepreneurial profit. It is serendipitous nature of the opportunity that allows the entrepreneur to complete a resourceless arbitrage opportunity to yield a profit for them. Indeed, *under conditions of pure competition*, we find that in the absence of a temporal passage, there is no value spread between the aggregate means and the end they construct. All factor prices would be fully imputed from their respective end and the rate of profit would be zero. We find that with the addition of a temporal element, a value spread occurs, *even in the presence of perfect competition*. According to Kirzner (1993: 168), "[c]ompetition cannot erode income." In this case the income in question is the spread caused by originary interest that is irremovable. It is not due to the productivity of waiting that causes this spread, but instead is rooted in the very nature of time preference, and the desire for humans to value satisfaction sooner rather than later.[235]

Likewise, Hülsmann correctly criticizes other time preference theorists (Böhm-Bawerk and Mises in particular) for treating heterogeneous goods as homogeneous in determining the value spread between present and future goods.[236] However, Hülsmann himself cannot claim immunity from this same error. In comparing ends and means, we see that means are not strictly additive in the formation of ends. Let us assume, for example, baking a dessert with only two ingredients: rhubarb and apple. The value placed on the end – the pie – is not strictly the aggregate of the additive and independent values of the means involved. Instead, we find that synergies occur when means are used together which create value not inherent in any individual good. Perhaps in this example, our two fruits combine for a pleasing taste, superior to that which would result from merely adding the two individual tastes.[237]

Finally, Hülsmann (2002: 86) posits that a valuational spread between means and ends can exist without a temporal element. There are two damaging comments that must be raised against this

235 Indeed, as Kirzner (1993: 173) informs us, "*no productivity considerations can possibly enter at all in the explanation offered for interest*." see also Kirzner (1996, 138) for the view that factor values can never be bid up in the aggregate to the same as the expected future value of the end.

236 Hülsmann particularly takes issue with Mises' confounding of goods and value in determining the value spread between temporally distinct objects. However, as Gunning (2005: 86) argues, Mises did in fact understand the need to compare heterogeneous objects, and worded his theory of time preference in terms of present and future values of *satisfactions*.

237 One of the better examples of non-homogeneous goods is Mises' (1949: 489) example where he tackles the problem of comparing ice in the winter and in the summer. If a person is temporally in the winter, they will value the ice more highly during the coming summer, a point further away from their temporal location. This example was solved by explaining that the two ices were different goods, and hence, not subject to pure time preference theory. The explanation lies in the fact that it is not the ice that is valued at different points in time, but a desire fulfilled by the ice, which would be different at the temporally separated points. See also Rothbard (1962: 69) and Block (1978: 128) for this viewpoint. All three authors seem to overlook the fact that what is compared in time preference is not physical means, but ultimate ends.

assertion.

The first is that, as he realizes, action and time are inseparable concepts. The concept of an action not occurring in time is unthinkable. Therefore, the idea that a value spread occurring through two distinct concepts – means and ends – fused by action without time is also unthinkable. Distinct means and ends must necessarily imply a temporal passage linking them together. Second, if we view the strict corollary of this assumption, a conception of time being so plentiful so as to require no economizing, we see that no value spread could occur between means and ends. How could this be if we know that we only value means to attain certain ends?

The reason stems from the very definition of a want. A want never exists in a timeless state, but instead the mere existence of a want implies we want it *now*, or at least, at a definite point in time. If time were an infinite continuance concerning the individual actor, there would be no concept of sooner or later. As a result, there would be no room for value – the temporal aspect is essential for value to occur. Means and ends would remain distinct in such a world with unlimited time, but our marginal utility of time would be so greatly reduced (i.e., it would be eliminated) in such a temporally endless world there would be no reason to prefer means to ends at any *point in time*.

In fact, we *know* interest to be caused by the concept of time preference and not the valuational spread between means and ends by looking to Mises' (1949: 245) evenly rotating economy (ERE). In this artificial construct, two elements are removed: uncertainty and Bergsonian time.[238] We find that time still passes, albeit in a linear, *Newtonian* manner. As the economy is in a steady state equilibrium, the profit rate has already trended to zero. However, we find a yield still exists between stages of production. The reason for this yield is not found in the valuational spread between means and ends of different stages, it is found through the scarcity of capital caused by the originary interest stemming from time preference. Since, in the ERE, uncertainty has been removed, we see that means will be valued exactly proportional to the ends they create, discounted for a time preference determined originary interest factor. As not all capital will be available to be used in the present, there will be a relative scarcity in the present that will be enticed to be made available through an appropriate yield. Hence, we can see that the ERE yields a valuational spread between stages of production not due to a disparity of means and ends, but through the concept of time preference manifested as originary interest.

The case for interest being caused by the valuational spread between means and ends is found

[238] See Howden (2008) for a look at the distinction between Newtonian and Bergsonian time in the ERE.

lacking. Time preference can only result through the temporal passage, which, prior to now has not been explained in a satisfactory manner. As *real* time is the central issue that confronts us, it will prove instructive now to embark on an new exploration of time preference theory, from this perspective.

A Bergsonian Defense of Time Preference

Humans value because they understand that time is limited regarding their actions. An individual who had no concept of the scarcity of time would be unable to assign a value to any action. However, we see that the source of action is the *expectation* that the fruit of such action will be valued greater *when realized* than the costs of action necessary for the attainment for that end are valued in the present. An individual unaware of the temporal passage is unable to partake in action; they are merely reacting to external stimuli.[239]

However, if we only mentioned that action occurs with the realization of a temporal passage, we would be simplifying the situation in a detrimental way. Animals too may (and indeed, in some cases, do) realize that time is passing them by as they partake in action. However, there is an additional realization that becomes instrumental and takes action to a more certain level – the knowledge that the temporal passing is finite; time is limited.

Of course, Mises (1949: 485n4) forwarded that animals *do* exhibit time preference, as evidenced by the storing of food for the winter. This observation led him to the conclusion that time preference was an "inherent feature of the behavior of all living beings." However, in his next breath, he correctly notes that time preference tradeoffs are conscious acts by humans, and not instinctive, as is the case with animals. An animal, or a young human infant, lacks the consciousness of time's true scarcity to act accordingly. Animals *may* however realize that a change in the environment will make future conditions unsuitable for the end they seek. For example, chipmunks may store food for preparation of a winter during which food will be scarce, or unavailable. This is not a display of time preference however. What they *are* doing is substituting present goods for future goods, *without regard for the need to economize upon time*.[240]

As Mises (1949: 12) illuminates the topic, action is not merely preference, it always involves "taking and renunciation." Both of these concepts – taking and renunciating – are meaningless in the

239 On this point see Mises (1949: 11-29).

240 In fact, this display may best be described as "time allocating" rather than "time preference." Mises' view on time preference has been criticized on this ground by Moss (1978: 161), with the distinction between these two time actions – allocation and preference – drawn.

absence of a temporally limited vantage point. If one had no conception of a time that would elapse at a defined (if ill-defined) point in the future, there would be no cost associated with giving up an object – the foregone alternatives would be eliminated. There would always be another time, and another opportunity, to assume the alternative. One can only renunciate an alternative at a point in time, and for an amount of time – time must be limited for this possibility to occur.[241]

Time preference is the product of actors desiring want satisfaction sooner rather than later. An endless time would preclude the concepts of sooner and later. *Was*, *now*, and *will be* would be effectively eliminated.

A clarification should be offered before progressing. The want that we prefer temporally sooner is satisfaction, never a concrete physical good causing such satisfaction. For example, the want we wish satisfied may be the satiation of our hunger. This does not mean we prefer one apple now to one apple tomorrow. Rather, it means we value the satisfying of our hunger now to its satisfaction at a temporally more distant time. The crux of time preference is, therefore, seen as defining what it is exactly we are preferring through a temporal passage.[242]

All objects that fall into the sphere of our actions are necessarily limited. A good available in abundance would not be subject to action, we would assign no value to it (Wieser 1889: 2-3; 21). Time exists as an endless flow, a timeless flux. It is, however, the distinctly human knowledge that time, as it pertains to action, *is* limited.[243] We gain knowledge, as Shackle (1961: 21) fatefully explained it, aware that we are "prisoners of time." This occurs in two ways. First is that we gain knowledge that our lives are not infinite, they must terminate at some unknown point in time that, *with certainty*, we know to exist. Second is the knowledge that any action we partake in will be necessarily limited. Our lives are built with singular actions, therefore, the gaining of awareness of our first point necessarily leads to the

[241] It may be raised that this assumes time to exist solely one-dimensionally – sooner and later – and that it ignores the different dimensions of the present – *nows* in which to undertake action. Further, for an actor to not be unable to discern between the present and the future, not only must the future be endless, but the present opportunities for action must always be endless. However, we see that the second case arises from the first. In fact, our present opportunities are only limited owing to the limited total scope of time available – the limitation stems from the *mere* fact that a now exists as time is scarce. However, if time were endless, *now* would no longer have a meaning; there would be no difference between the present and the future without end. As a result, the present opportunities for action would also be endless.

[242] Considerable confusion has arisen in the past due to this point (see above fn 12). Böhm-Bawerk's original ambiguity in defining between physical and value productivity has been carried forward throughout time, and reappears in many treatments of originary interest. Fisher (1930: 62), for example, following in the time preference tradition, defines the temporal choice between "one more unit of *present* goods over the *present* marginal wants for one more unit of *future* goods. He does, however, later try to rephrase the problem in terms of *real* income (64-65).

[243] As Bergson (1896: 275) viewed it, "to perceive consists in condensing enormous periods of an infinity diluted into a few more differentiated moments of an intense life, and thus summing up a very long history." It is this very perception of limited time that defines us as humans. Hoppe (2001: 4-5) seems to recognize this recognition of limited time as a factor influencing an actor's time preference.

knowledge of the second.

However, there is one element of action that is unlimited, yet still has value to us: wants.[244] Human wants are without end, they are the continual motive for action, and spur us to continue to live. A life without a want could not exist. And yet we know that objects that exist in unlimited form are not subject to the acting human; they are not economized. The reason we are able to act upon our wants is that we realize the inextricable link between time, wants and action. As Rothbard (1962: 5) tells us: "Time is *scarce* for man only because whichever ends he chooses to satisfy, there are others that must remain unsatisfied." Humans rank their wants in an ordinal manner, from most desired to least. The most desired action is the one undertaken at any given time. If we lacked a concept of a limited time, there would be no impetus to rank wants. Any want would be as desired *at any given time* than any other want, as there would be no reason to distinguish between the *now* we wish to satisfy a want during, and the *future* when a less pressing want will be satisfied.

Some have tried to ground time preference in biological reasoning. Fetter (1915: 240), for instance, tries to outline time preference on a biological basis whereby humans have an "impulse" to want things now. Reisman (1990: 56) likewise thinks that:

> The nature of human life implies time preference, because life cannot be interrupted. To be alive two years from now, one must be alive one year from now. To be alive tomorrow, one must be alive today. Whatever value or importance one attaches to being alive in the future, one must attach to being alive in the present, because being alive in the present is the indispensable precondition to being alive in the future.

However, life is not a distinctly human quality. Animals, insects, and plants all enjoy life. Instead, time preference is forever a value problem rooted in the realization of time's true nature – its limited nature. It is this factor that makes us distinctly human, and hence time preference a purely human phenomenon.

Human action is always directed towards the satisfaction of an end at some point in the future. This future must, always, contain an element of uncertainty in it. However, it is not uncertainty that is the driving force of time preference, although it may strengthen our want for present satisfaction (Fetter 1915: 241). We could envision a world of pure certainty and see that time preference would still exist.

[244] As Menger (1871: 83) noted, human wants are infinite, but become quantifiable, and hence, subject to economization, once we realize the limited nature of means and time periods available to satisfy them.

Whether we know for certain everything the future will hold, or have limited knowledge of the future state of affairs, we will still have to economize our wants temporally by satisfying them sooner rather than later. In fact, the only certainty that would eliminate time preference completely would be the certainty that time would last forever. We know this cannot happen. However, uncertainty as to the quantity of time we have remaining, and the expectation thereof, will temper our degree of time preference. If we are uncertain as to the amount of time outstanding for our actions, and the expectation that results from this is that we have a more limited amount of time remaining than would be the case if the situation was more certain, we will begin to increase our time preference as we value our time more highly. We try to accomplish more in the present based on our expectations of the shortened future duration our present will have.

As Mises (1949: 477-478) tells us, action is always concerned with a "definite and limited fraction of the future," never the future in general. When we commence an action, we have a period of time we expect it to continue for. Likewise, we see the marginal insight come into play as we never value an unlimited amount of time, but only that portion that is applicable, or expected to be applicable, to suit our sought after end.[245] It is with this reasoning that time stops being a valueless object due to its super-abundance (limitless nature), but an object with a duration, subject to be valued by the acting individual.

Böhm-Bawerk reckoned that time preference existed due to psychological factors.[246] Mises found this unsatisfactory, and tried to explain it in terms of the human desire for life and the necessity of consumption. Both explanations are lacking for a science based upon deductive logic seeking explanations for universal human actions. We prefer want satisfaction in the present over the future because we realize the temporal restrictions placed upon us by the very life we continually create through action in the present.

When we realize the connect that lies between wants and time, we see where the true source of time preference stems from. Time is limited for humans, *although we must become aware of this fact before it affects our actions*. Wants, on the other hand, forever remain in a state of endlessness – they

[245] Indeed, Fisher (1907: 92) stressed that the rate of time preference between any two periods depends not only on the value spread between two *points* of time, but potentially between the present point in time and a future *stream* of satisfaction. The utilities being valued each have their own respective expected temporal durations that they will deliver value over.

[246] Likewise, Fisher grounded his later views of interest determination in the concept of "impatience," arguing that: "Impatience is a fundamental attribute of human nature. As long as people like to have thing today rather than tomorrow, there will be a rate of interest. *Interest is, as it were, impatience crystallized into a market rate*" (1911: 387, as quoted in Pellenghar 1996: 82; see also Fisher 1912: 371). Mises would also deny this cause (impatience) as a psychological approach, unsuitable for the field of economics. (See Mises 1949: 486).

can never end until our life expires. Within this limit of time humans move to satisfy as many wants as is possible (or wanted), before our time runs out. As our wants are always ranked from most preferred to least, it follows that we continually prefer our most highly desired want in the present to a later date, due to the limiting notion of time. Our preference for *sooner* over *later* is how the acting human economizes on time; it is how we treat time as a scarce resource.[247]

Entrepreneurial Profit and Time

Time is of central importance to the economist. As Garrison (1984: 200) correctly points out, time is the medium of all action. We have seen that the true source of time preference lies in two distinctly human qualities. Our ability to value means to achieve ends creates a hierarchy of satisfactions in need of fulfillment. Our ability to see the true limited nature of time allows us the ability to economize upon our endless wants within this restrictive temporal framework. As we gain the realization that our means, including time, are limited, but our ultimate ends desired are unlimited, we maximize upon this by preferring the fulfillment of ends sooner to later. The temporal spread then creates the source of originary interest, a concept with a plethora of consequences for the economist to explore.

Originary interest was a concept explaining many facets of human action, in need of a theory. When time is assessed as it exists for acting humans, we see its limited and dynamic nature become apparent. As wants are always ordinally arranged, from most to least valued, every action undertaken is necessarily the most valued at any given time. These wants must always be fulfilled through an action occurring during a specific temporal duration. The distinctly human realization of the occurrence of limited time – the future end of time as it concerns individual action – brings rise to the desire to satisfy wants earlier in time to later. The human preference for time to fulfill wants sooner rather than later is how the actor is able to economize on this limited resource to attain their most valued ends.

It may also be seen that due to the unlimited nature of wants, there is no way that this valuational spread – between want satisfaction now compared to later – may be arbitraged away. There will forever remain a non-negative value attached to want satisfaction occurring in the present rather than at a later date. Entrepreneurial profit, in contrast, has a *tendency* to be arbitraged away through the competitive forces of diverse entrepreneurs. Profits in one area of the economy provide a signal to

[247] It must be stressed that time preference is never preferring the satisfaction of a defined want to another in the present. Instead, it is the satisfaction of *a* want in the present over the future. As our desires for satisfaction continually change, we can never predefine what exactly it is that would be preferred in the present to the future. Instead, we may only say, without hesitation, that some want is preferred to another in the present.

other entrepreneurs that there are apparent disequilibria discovered which require exploiting. As entrepreneurs move into this previously unexploited area of the economy, the entrepreneurial profit rate declines through these competitive forces.

Entrepreneurs have one alternative method to earn a return which is separate of any of the three ways identified in the previous chapter. By exploiting the valuational spread between present and future wants, a return can be earned which will always remain non-negative. Hence, entrepreneurs can loan their resources to other entrepreneurs, and earn a profit in this way. However, the rate of originary interest, that stemming from the true renunciation of want satisfaction in the present over the future, cannot be arbitraged, implying that although competitive forces can place downward pressure on it, it may never be entirely eliminated. Hence, entrepreneurs need not search for a monetary return solely through pure profit seeking activities outlined earlier. Instead, the *mere* act of renunciating present want satisfaction until the future will earn return.

As this return is produced through a method drastically different than the entrepreneurial profit, it falls into a wholly separate category. At the end of the previous chapter we noted that entrepreneurial profit could only be considered as that which was produced *above* that which would have been produced regardless of any of these entrepreneurial activities. It should be clear that the rate of any level of entrepreneurial profit must, then, exceed that of the rate of originary interest to be considered as true profit – earned as a result of the entrepreneur's ability to foresee future uncertainty or mitigate risk. Therefore, any true measure of entrepreneurial profit will be that return which they earn on owned resources (inclusive of all resources used to satisfy a consumer's want) *minus* that which they would have earned if they had done nothing with their resources, that is to say, if they renunciated all use of these resources by loaning them to another entrepreneur on the time market.

VII. EQUILIBRIUM

The last chapter should have made clear that time is an integral part of the market process. In fact, by definition, the market *process* must pay consideration to the temporal passage which it is confined to. At the same time we saw that time on its own creates a valuational spread between present and future want satisfaction. However, by viewing market outcomes as equilibrium situations valuable insights may be gleaned. We should keep in mind that it is generally not the equilibrium itself that we are concerned with but rather the process which leads to this point which becomes the focus of our attention.

Mises' equilibrium conception – the ERE – suffers a fatal flaw in that is assumes away the source of one of the things we wish to see. By removing *real* time, originary interest is also removed from the construct. Interesting conclusions may still be gleaned, but these are helped substantially by modifying the ERE to better account for originary interest.

By creating an equilibrium incorporating pure entrepreneurs and real time, the true source of growth can be deduced – future uncertainty. As entrepreneurs move towards bearing future uncertainties, the opportunities for sustainable growth become apparent. In fact, this entails a feedback loop as the profits provided by the class of entrepreneurs that mitigates risk is used to provide the resources which the uncertainty-bearing entrepreneurs require in order to create the future that they see.

Uncertainty should not be shunned, but rather embraced, as it is the ultimate source of continual profits, and hence, economic growth.

1. Mises' Evenly Rotating Economy

Mises (1949: 244-251) created an artificial construct, the evenly rotating economy (ERE) in order to explain the entrepreneurial function better by removing two elements: time and changes in market data. In the course of removing time from this conception, Mises does not mean that time ceases to exist. The removal of time completely would imply the non-existence of everything it is that we wish to see through the use of the ERE. Instead, the removal of time implies the removal of *Bergsonian* time.[248] Newtonian time still passes, but it progresses as a static wave, existing as the true antithesis of real time. Market actors must move through it, but they do not feel it, or alter their behavior accordingly. The essential point that makes time come alive to us as acting humans is removed – it incites no new knowledge in us, and alters nothing of us. We see that as a result of this removal of the essence of time, the economy of the ERE still functions, and progresses, but it does so without change. In Mises' (247) words:

> The system is in perpetual flux, but it remains always at the same spot. It revolves evenly around a fixed center, it rotates evenly. Therefore prices – commonly called static or equilibrium prices – remain constant too.

Hence, as change is removed from the ERE, entrepreneurially acting humans are likewise removed. There are no longer disco-ordinations that require attending to, nor is there an equilibrium in the distant future to be reached; it already exists in the present. This state of equilibrium infers that all factors of production are allocated to the area where their discounted marginal value product is highest, as determined by consumers (Rothbard 1962: 514), and this allocation will remain set indefinitely.

Mises utilized this construct, well aware of its limitations for one purpose: as a purely abstract model to demonstrate other principles, in particular, entrepreneurial profit and loss.[249] It was only in constructing a system that removed the entrepreneur wholly, that changes concerning this role could be introduced and analyzed. For example, with the elimination of change (and the entrepreneur by association), we find that profits will also disappear. A yield will still be realized between the stages of production however, equivalent to the time preference yield dictated through the passage of *Newtonian*

[248] See O'Driscoll and Rizzo (1985: esp. chap. 4) for a comparison of Newtonian and Bergsonian time, as it applies to action.

[249] Mises was well aware of the fallacy of using static constructs: "The problem of economic calculation is of economic dynamics: it is of no problem of economic statics" (1936: 139).

time.

The Plain State of Rest

The use of an artificial construct may seem odd for an economist of the Austrian school, one that prides itself on realism. Mises' endeavor was not the first attempt at static modeling however. As Salerno (2006: 45) points out, Mises' teacher, Böhm-Bawerk, and in turn, his teacher, Menger, had utilized similar constructions. How do these alternative viewpoints dovetail with the Austrian conception of the market as a dynamic process?

Menger (1871: 188) notes that:

> [T]he foundations for economic exchanges are constantly changing, and we therefore observe the phenomenon of a perpetual succession of exchange transactions. But even in this chain of transactions we can ... find points of rest at particular times... At these points of rest, no exchange of goods takes place because an economic limit to exchange has already been reached.

Likewise, Menger's pupil Böhm-Bawerk speaks of the exchange process as taking place temporally. Much like monetary valuations allow for actors to reduce the bid-ask spread, or "zone" of prices that exchange will occur within, the process of valuation will also move towards reducing the temporal "zone" between changes in prices. As this temporal zone is reduced to a point in time, the market achieves a "momentary equilibrium" (1889: 231).

Hayek (1937a: 1948) defined equilibrium as the compatibility of plans. As these plans are dispersed, and potentially tacitly given, Hayek views the coordination of these plans to an equilibrium setting as highly suspect. He would, however, make subtle use of momentary equilibria – a "temporary state of rest" – to explain the coordination process (see Hayek (1941: 19) for one such example).

Hence, there was a long precedence for equilibrium constructs within the Austrian school before Mises' ERE. However, the ERE took the static nature to a new level, introducing an element of unrealism not evident in these earlier approaches. Mises bridged this gap by introducing a similar concept to Böhm-Bawerk's momentary equilibrium – "the plain state of rest" (1949: 245).

In his view, this was not a wholly artificial construct, but instead represented an "adequate description of what happens again and again on every market." If time ended at any given moment, this would be true. This state is achieved in the sense that the state of affairs that has occurred, and has taken all of history to reach this point in time, ends at every moment. Hence, the plain state of rest can describe accurately the world that exists at any given point in time.[250] Hayek (1937b: 22) shows that the international monetary flow mechanism can be analyzed through comparisons of plain state of rest balances. Likewise, Salerno (1994a) demonstrates the process of monetary adjustment from disequilibrium to the plain state of rest. The plain state of rest is not of mere theoretical importance, but of practical importance as well.

In contrast to the passing occurrence of the plain state of rest, the ERE can never exist as a realizable state, *no matter how fleeting*. Mises takes great pain in stressing the plain state of rest is not imaginary, but an everyday occurrence on the market.[251] In contrast, the ERE is an *imaginable* construct, but only within its own known boundaries.

A Consistently Rotating Economy?

The use of the ERE has seen increased use over time, but has also come under fire. Cowen and Fink (1985) argue that the ERE fails in its role as an artificial construct as it: (1) fails to be totally unrealistic, and (2) is internally inconsistent. However, through their criticism, they demonstrate a poor understanding of the ERE's construction, and its proper use. It was never created to be wholly non-descriptive of reality, this would serve against everything Mises forwarded as being methodologically necessary for economics. Instead, it was to be sufficiently unrealistic so as to demonstrate the problem intended – entrepreneurial profit and loss. Second, it is only viewed as internally inconsistent if it is misused, as they portray it to be. Mises made note of the fact that *exchanged* money ceases to exist in the ERE, as uncertainty is a natural precondition for the existence of this factor. This only precludes the possibility of explaining monetary changes through the ERE, something that it was never originally

[250] Of interesting note is that Mises viewed the possibility of a stable purchasing power of money as being an impossible realization. However, it was only in this final state of rest (Mises' later adopted term for equilibrium) that money could have a stable purchasing power – if only for a fleeting moment. See Mises (1944: 47).

[251] As Garrison (1984) reminds us, there is no significant issue *theoretically* by viewing single actions in this static manner. All action does, after all, transpire within its own single moment. It is only the masking of the processes that connect singular actions together, or the market process, that creates a significant issue when viewing the world to exist in this unrealistically static construct. We see the meaning of this statement in light of constructs such as the plain state of rest. Hayek (1941: 22) views the usage of stationary states as useless as it eliminates the crux of the problem we wish to study: "[T]he construction of a stationary state is particularly useless because the main problem ... arises just because people intend to do in the future something different from what they are doing in the present."

proposed to do.

Furthermore, Cowen and Fink (1985: 867) fault the ERE as not sufficient in explaining how it would be reached. In their eyes, the removal of change and time does not adequately explain why an equilibrium, such as is described in an ERE, would come to exist. However, as Garrison (1991: 95) has countered, "[i]t is not necessary for the initial conditions to preclude all kinds of disequilibria but only to preclude systematic intertemporal disequilibrium – the kind of disequilibrium for which the theory itself accounts." The removal of Bergsonian time eliminates the root of intertemporal disequilibrium.[252] Again, we see that when used *within its own limitations* the ERE provides a consistent analytical tool.

Lastly, as Gunning (1989: 126) points out, the definition that Cowen and Fink use regarding equilibrium is mistaken:

> The problem with [Cowen and Fink's] criticism is that it is based on a *mathematical* definition of equilibrium and not a *logical* definition consistent with Mises' pure logic of action. In the logical definition, the concept of disequilibrium is meaningless. To say that there is a tendency toward disequilibrium is like saying that individuals do not make choices.

Hence, we see that Mises' ERE is an invaluable tool, *as it has been heretofore developed*, understood within its known limitations, and used appropriately. Its construction as a partly unrealistic representation does not fault its results, but instead gives added meaning and clarity to them.[253]

---

[252] O'Driscoll and Rizzo (1985, 82-83) remind us that "the state of *ex ante* coordination is not enough for equilibrium; there must also be no *logical* impossibility standing in the way of the actual consummation of intentions." The removal of dynamic Bergsonian time provides the circumstances necessary for an equilibrium to be achieved.

[253] Cowen and Fink make two additional errors in their conception of the ERE that deserve quick mention. First is their (1985, 866) contention that Mises implies the ERE to contain money prices, and their note that the lack of a futures market implies a deficiency in its construction (869). When Mises ([1949] 1998, 416) mentions money prices, he implies money as a *numèraire*, not as an exchange medium. The lack of a futures market stems naturally from the elimination of an uncertain future, a point which deserves no further discussion.

2. Dynamic Equilibria

Mises' ERE viewed equilibrium from only one of several possible manners that could be used. By removing *real* time, he was able to achieve a dynamic equilibrium in which time passes by, but no change in market conditions results. We may remember that there are several methods with which entrepreneurs serve to increase consumer want satisfaction, or, to borrow Hayek's terminology, increase plan co-ordination. Since the use of equilibrium serves to demonstrate the nature and source of entrepreneurial profit, it may prove pertinent to look at this construct in light of other equilibrium constructs – both static and dynamic – and see what results can be yielded. Table 4 from chapter V has been reprinted below to show the three sources of entrepreneurial profit and the significance that time serves for each.

Time Dynamics and Event Types		
	Risk	**Uncertainty**
Static (one-period)	Non-seriable/non-divisible events	
Dynamic (multi-period)	Class Probability	Case Probability

Table 4

By looking at the three temporal sources of entrepreneurial profit, we see that Mises' ERE typically works by changing the dynamic temporal passage from real time to linear time. This implies that time exists, but as no change is enacted by it, the entrepreneur is eliminated. Hence, we find that entrepreneurial profits represented by the bottom two quadrants become eliminated by design. Also as a result, there can be no static profits as they are removed with change – they become arbitraged away. As Mises concluded, in this type of situation pure entrepreneurial profits tend to zero, but the passage of time still represents a value spread between different temporal stages.

However, returning to the previous chapter, we saw that time preference does not result from the *mere* passage of time, but rather through the economization of our wants which are unlimited owing this fact, coupled with the reality that time is of a limited nature. In the ERE while it is true that time is still limited, as it lacks a dynamic component, human wants are no longer *un*limited but embody a finite character. This stems from the fact that the source of our infinite wants comes through the passage of time which introduces us to new temporal viewpoints, and hence, opens up possibilities for want satisfaction which were not apparent before. For our wants to increase, it must be possible for us to imagine a future that is different than the present – an occurrence ruled out unless one's perception of time is necessarily *real* and not *linear*.[254]

In Mises' ERE, we find instead that time *preference* does not prevail, only time *economization*. Individuals in such an equilibrium prefer something now to later as it becomes more optimal in the production and consumption cycle to do so. Hence, goods between different temporal stages of production may exhibit a positive value spread between them, however, this will not stem from a preference for satisfaction sooner to later (as per conventional time preference) but rather it will result from a greater value placed on a good at a certain time in production compared to another (i.e., in more conventional terms, a marginal rate of substitution greater than 1 will prevail).

This problem may seem moot, but is easily surmountable to obtain much the same conclusion as Mises. Instead, we may create a general dynamic equilibrium where time is real (that is, it insights change), *and* the entrepreneurs in the economy are pure. That is to say, all future uncertainties are planned for accordingly, *and* all risks are mitigated. Only in this case can we attribute a value spread between temporal stages to time preference. With the addition of real time, it is now a real possibility that wants become unlimited, and hence, the requirement of time preference is created. Now a value spread will obtain between stages given by the prevalence for consumers to withhold consumption (that is, to renunciate want satisfaction) in the present, in exchange for that at a later date. Entrepreneurs exist to foresee this later date and use the resources provided by the renounced present wants to build this desired future. As all entrepreneurs are pure, in the sense that they do not err, profit differentials disappear (as they tend to 0) and the remaining residual is the originary interest rate.

We may determine that what is required of an equilibrium construct to demonstrate the difference between entrepreneurial profit and originary interest is not the removal of real time, as per Mises, but instead the addition of pure entrepreneurs. Mises' ERE was aptly named as it represented an economy with action, one that is rotating in the same spot, and never changing nor growing. However, if originary interest were apparent, as Mises reckoned, then it is true that the economy would be rotating consistently, but it would also be growing at a rate equivalent to the originary interest rate less the cost of depreciation and replacement of old capital goods. Mises makes no reference to this

[254] See Rizzo (2000: 175): “It is the *process* of acting, rather than the mental picture of completed acts, that discloses time as real duration or the continuous flow of novelty. This is time-as-lived rather than as thought. As we work through an action, we must conceive of a future different from the present and thus be conscious of a heterogeneous time-flow.”

depreciation component, and how it affects growth. For Mises' result to obtain, we find therefore that two additional caveats need be employed. The first is that entrepreneurs are pure, in that they can perfectly foresee the uncertainty future (which is uncertain due to the passage of real time). Second is that the rate of savings prevailing (and hence, that causing originary interest), is equivalent to the rate of depreciation on the capital goods. This second caveat will give rise to an economy that is not growing, nor shrinking, but rather retaining its size while simultaneously "rotating."

The Process to Equilibrium

We have seen previously that although forever trending toward an equilibrium, this point is never reached. Much like Rothbard (1962: 322) maintained, the drive towards equilibrium is like dogs chasing a mechanical rabbit at the race-track. They follow it around the track, but never catch up with it. Equilibrium is never a state that can be reached, however, it does shed important light on which end-state the process of action is trending towards. It is this process that becomes the focal-point of our attention as the problems of economics become apparent, not in equilibrium, but on the path to this unrealizable state.

In Mises' original conception of the ERE, the removal of real time also removed the entrepreneurs that exist to predict the uncertain future. However, as we have seen, in the dynamic sense there are two roles for the entrepreneur, only one of which concerns shouldering future uncertainty. The entrepreneur whose existence is owed to risk mitigation would still exist in the ERE. In fact, absent any uncertain future, it would be this exact group of entrepreneurs who move the market toward equilibrium – they *are* the process which embodies the equilibrating tendency. Once equilibrium is reached, there will no longer be a need for this type of entrepreneur as all risk has been mitigated, and the lack of continued uncertainty leads to no additional risks developing. Hence, although in equilibrium both types of entrepreneurs will be absent, during the process shifting to that point, risk-mitigating entrepreneurs will serve a significant equilibrating role in the economy.

We may also note that in this conception of the ERE, as the temporal passage is not real – it is static time – that there cannot be a source of time preference. Instead, time economization prevails as wants are finite and more optimally met at one location in the temporal sequence than another. Competition in the entrepreneurial realm always tends towards eliminating all profits above and beyond those that could be attained by refraining from partaking in entrepreneurial activity. This, as was seen in chapter V, implies that entrepreneurial profits only entailed those above the prevailing originary interest rate, and that these entrepreneurial profits will trend towards zero. With no originary interest rate prevailing in the ERE, we see that entrepreneurial profits will decline to the point where the *total* return is zero. With no profits or positive returns, the economy will reach a finite point of maximum growth. No excess returns can be used to increase growth further as they disappear by the entrepreneurial process. At the same time, it may become evident that the entrepreneur has also brought forth their own demise.

This conclusion makes much intuitive sense as we consider that an economy not discovering any new opportunities to exploit, will have a finite level of growth until all pre-existing disequilibria are discovered and repaired. Although economic growth will be experienced during this process owing to the increased efficiencies of the risk-mitigating entrepreneurs, this will be a short-lived outcome as no new risky disequilibria will be created by the uncertainty bearing entrepreneurs. In the absence of fresh, risky situations, the economy will become efficient to a specific point, and then the evenly rotating process which Mises described will begin. Much like Huerta de Soto (2004) makes clear, economies need new uncertain discoveries in order to grow (i.e., dynamic efficiency) instead of the more conventional risk-mitigating growth techniques focused on optimizing a given end-state (i.e., static efficiency).

We may turn to our modified conception of the ERE, in which there is the presence of real time, and simultaneously this is fully mitigated by pure entrepreneurs fully foreseeing the future uncertainty. There will continuously be new growth, as real time occurs with uncertainty-bearing entrepreneurs continually foreseeing these future opportunities. At the same time, there will now be an originary rate of interest separating different temporal stages. This is owing to the fact that real time gives rise to the limitless nature of wants, which is a precondition for time preference (itself the cause of originary interest).

If entrepreneurs are fully foreseeing future uncertainties their profit rate will be arbitraged to zero. Originary interest will still yield a return for inter-temporal investments, but this return will be fully exclusive of any uncertainty-bearing entrepreneurial profit. However, as real time passes and entrepreneurs continual meet the *future* uncertainty, there exists in the present *risky* disequilibria that are constantly evolving and needing mitigation. The class of entrepreneurs who work towards eliminating risky situations are faced with a continual stream of changing circumstances in need of adjustment. This continual stream of fresh risky situations, however, will not be able to be fully arbitraged away. The entrepreneurial class focusing on risky mitigation will earn a positive profit rate above the rate of originary interest that will continually prevail – no tendency will exist for this entrepreneurial profit rate to tend to zero. At the same time, competitive forces will have a tendency to drive this entrepreneurial profit rate towards a long-term equilibrium rate of return on risk-bearing, but this long-term rate will never prevail owing to the continual source of fresh risk-factors.

The economy will be growing in a dynamically equilibrated manner. Uncertainty bearing entrepreneurs fully foreseeing the uncertain future ensure that future needs are met in the present. Risk-mitigating entrepreneurs faced with a continual stream of new risk-factors continually ensure that efficiency is maintained, and resources are not wasted. A question that may well be raised at this point is, “As risk-mitigating entrepreneurs are earning a positive profit for their efforts, what happens to this profit?” The answer provides the source of the sustained growth of the economy. Risk-mitigating entrepreneurs provide the uncertainty-bearing entrepreneurs with the resources they need to continually plan for the future through these profits. These profits, then, provide the resources that the economy requires to grow in this dynamically stable manner.

This type of economy is also in an equilibrium, similar to Mises' ERE. However, this one will be marked by dynamically stable growth. The uncertainty of the future gives rise to a profit source in the present. This profit source, owned by the risk-mitigators, is loaned to the uncertainty-bearers, and this resource provides the input to drive the economy forward. As the entrepreneurs are pure, the future is met perfectly, leaving no opportunity for inter-temporal disco-ordination. However, this also implies that a fresh stream of risk-factors will continually prevail, giving rise to the a continual profit source for the risk-mitigating entrepreneurs. This continual profit is only made possible through the efforts of the uncertainty-bearers, who now become apparent as the ultimate source of profit, and reason for the economy's advance. As risk-mitigators reinvest their profits with the uncertainty-bearers, resources are provided for growth, which is dynamically sustainable. Much like Mises' ERE, the economy is rotating, however, it is simultaneously expanding into fresh, previously unknown possibilities.

Growth and the Entrepreneur

It now becomes apparent that if the economy is to grow, the emphasis is placed on the uncertainty-bearing entrepreneur. This group is the ultimate source of all *sustainable* growth potential in the economy. An economy may maintain the appearance of growth owing to the profitability and efforts of the risk-mitigating entrepreneurial class increasing efficiencies. However, lacking any sort of uncertainty bearer to foresee the future and move towards creating it, any risk-mitigation will be short-lived. As entrepreneurs continually compete against each other, risky situations become exploited and a state of equilibrium will be reached. It is only with the appearance of new uncertainties that fresh risky situations may appear to be exploited. However, risk-mitigating entrepreneurs provide the resources for the uncertainty-bearers to use in order to achieve the future state that they envision. We find that both sets become instrumental for each others long-term sustainability and an equally essential part of the continued growth of an economy.

VIII. ACTION

Action thus proceeds a results of the continual uncertainty and risk that actors face. The dichotomy of entrepreneurial roles that has been previously outlined – risk mitigation and uncertainty bearing – provides the essential occurrences that evolve the state of consumer want satisfaction to higher and higher levels. Both have been shown to be essential from the limitations that our mental capacity has at discovering and storing information as knowledge. This limitation has repercussions important to the study at hand through the concept of bounded rationality which must be stressed when viewing any result of human action.

The view of action to this point has, however, been quite simplistic and limited implicitly to autarkic situations. In the real world we find that social institutions are as much a result of our actions as they are an influencing factor on them. Concepts such as consumption and production have, until now, been used in loosely defined ways. Developed society and markets cannot focus solely on isolated instances of these essentials of action – production and consumption – but instead must incorporate a theory of why they arise and how they manifest in the market. In Book II we shall delve into the creation of institutions that assist these functions, as well as determinants that shape the extent to which both are practiced.

REFERENCES

Arrow, Kenneth J. 1951. Alternative Approaches to the Theory of Choice in Risk-Taking Situations. *Econometrica* 19(4): 404-437.

Arrow, Kenneth. 1959. *Economic Welfare and the Allocation of Resources for Invention*. Santa Monica, CA: The RAND Corporation.

Arthur, Brian W. 1991. Designing Economic Agents that Act like Human Agents: A Behavioral Approach to Bounded Rationality. *The American Economic Review* 81(2), Papers and Proceedings of the Hundred and Third Annual Meeting of the American Economic Association: 353-359.

Arthur, Brian, W. 1994. Inductive Reasoning and Bounded Rationality. *The American Economic Review*, 84(2), Papers and Proceedings of the Hundred and Sixth Annual Meeting of the American Economic Association: 406-411.

Baudin, Louis. 1954. Irrationality in Economics. *Quarterly Journal of Economics* 68(4): 487-502.

Becker, J. W., and F. O. Brownson. 1964. What Price Ambiguity? Or the Role of Ambiguity in Decision Making. *Journal of Political Economics* 72(1): 62-73.

Bergson, Henri. [1889] 1910. *Time and Free will: An Essay on the Immediate Data of Consciousness*. (*trans*.) F. L. Pogson. London: George Allen & Unwin.

Bergson, Henri. [1896] 1911. *Matter and Memory*. (*trans*.) Nancy Margaret Paul and W. Scott Palmer. London: George Allen & Unwin.

Bewley, Truman. 2002. Knightian Decision Theory: Part 1. *Decisions in Economics and Finance 25*(*2*): 79-110.

Bikhchandani, Sushil, David Hirshleifer and Ivo Welch. 1992. A Theory of Fads, Fashion, Customs and Cultural Change as Informational Cascades. *Journal of Political Economy 100*(*5*): 992-1026.

Birnbaum, Michael H., and R. J. Gutierrez. 2007. Testing for Intransitivity of Preferences Predicted by the Lexicographic Semiorder. *Organizational Behavior and Human Decision Processes* 104: 97-112.

Birnbaum, Michael H., and Ulrich Schmidt. 2008. An Experimental Investigation of Violations of Transitivity in Choice Under Uncertainty. *Journal of Risk and Uncertainty 37*(*1*): 77-91.

Blavatskyy, Pavlo. 2006. Axiomatization of a Preference for Most Probable Winner. *Theory and Decision 60*(*1*): 17-33.

Block, Walter. 1978. The Negative Interest Rate: Toward a Taxonomic Critique. *Journal of Libertarian Studies* 2(2): 121-124.

Boehm, Stephan. 1994. Hayek and Knowledge: Some Question Marks. In (*eds.*) Colonna, M., H. Hageman, and O. F. Hamouda. *Capitalism, Socialism and Knowledge: The Economics of F. A. Hayek*, Vol. II. Aldershot, UK: Edward Elgar.

Böhm-Bawerk, Eugen von. [1884] 1959. *Capital and Interest. Volume I: History and Critique of Interest Theories*. South Holland, IL: Libertarian Press.

Böhm-Bawerk, Eugen von. [1889] 1959. *Capital and Interest. Volume II: Positive Theory of Capital*. South Holland, IL: Libertarian Press.

Boring, Edwin G. 1933. *The Physical Dimensions of Consciousness*. New York: The Century Co.

Čapek, Milič. 1961. *The Philosophical Impact of Contemporary Physics*. Princeton, N.J.: D. Van Nostrand.

Čapek, Milič. 1971. *Bergson and Modern Physics: A Reinterpretation and Re-Evaluation*. Dordrecht, Holland: D. Reidel.

Carter, C. F. 1958. Time in Economics: A Review. *The Economic Journal 68*(*271*): 543-545.

Chomsky, Noam. 1976. *Reflections on Language*. Glasgow: Fontana.

Cohen, John. 1964. *Behaviour in Uncertainty and its Social Implications*. New York: Basic Books Inc.

Collins, H. M. 1992. *Changing Order. Replication and Induction in Scientific Practice*. Chicago: University of Chicago Press.

Collins, H. M. 2001. Tacit Knowledge, Trust, and the Q Sapphire. *Social Studies of Science* 31(1): 71-85.

Conlisk, John. 1996. Why Bounded Rationality? *Journal of Economic Literature* 34(2): 669-700.

Cowen, Tyler, and Richard Fink. 1985. Inconsistent Equilibrium Constructs: The Evenly Rotating Economy of Mises and Rothbard. *The American Economic Review* 74(4): 866-869.

Čuhel, Franz. [1907] 1994. On the Theory of Needs, (*trans.*) William Kirby. In *Classics in Austrian Economics: A Sampling in the History of a Tradition. Volume I: The Founding Era*, (*ed.*) Israel M. Kirzner. London: William Pickering.

Dennett, Daniel C. 1995. *Darwin's Dangerous Idea. Evolution and the Meanings of Life*. London: Penguin Books Ltd.

Dequech, David. 2000. Asset Choice, Liquidity Preference, and Rationality Under Uncertainty. *Journal of Economic Issues* 34(1): 159-176.

Edelman, Gerald M. 1990. *The Remembered Present*. New York: Basic Books.

Edelman, Gerald M., and Giulio Tononi. 2000. *A Universe of Consciousness*. New York: Basic Books.

Egerton, R. A. D. 1960. *Investment Decisions Under Uncertainty*. Liverpool: Liverpool University Press.

Ellsburg, Daniel. 1961. Risk, Ambiguity and the Savage Axioms. *Quarterly Journal of Economics 75*(*4*): 643-669.

Fama, Eugene F. 1970. Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance* 25(2): 383-417.

Ferguson, Adam. [1767] 1966. *An Essay on the History of Civil Society*. Edinburgh: Edinburgh University Press.

Fetter, Frank A. 1902. The 'Roundabout Process' in the Interest Theory. *Quarterly Journal of Economics* 17(1): 163-180.

Fetter, Frank A. 1914a. Capitalization versus Productivity: Rejoinder. *The American Economic Review* 4(4): 856-859.

Fetter, Frank A. [1914b] 1977. Interest Theories, Old and New. Reprinted in *Capital, Interest, and Rent: Essays in the Theory of Distribution* (*ed*.) Murray N. Rothbard. Kansas City: Sheed Andrews and McMeel.

Fetter, Frank A. 1915. *Economic Principles*. New York: The Century Co.

Fisher, Irving. 1907. *The Rate of Interest*. London: Macmillan.

Fisher, Irving. 1911. The Impatience Theory of Interest. *Scientia, Rivista di Scienza, International Review of Scientific Synthesis* IX: 380-401.

Fisher, Irving. 1912. *Elementary Principles of Economics*. New York: Macmillan.

Fisher, Irving. 1913. The Impatience Theory of Interest. *American Economic Review* 3.3: 610-618.

Fisher, Irving. 1930. *The Theory of Interest*. New York: Macmillan.

Foldes, Lucien. 1958. Uncertainty, Probability and Potential Surprise. *Economica*, *New Series 25*(*99*): 246-254.

Garrison, Roger W. 1979. In Defense of the Misesian Theory of Interest. *Journal of Libertarian Studies* 3(2): 141-150.

Garrison, Roger W. 1984. Time and Money: The Universals of Macroeconomic Theorizing. *Journal of Macroeconomics* 6(2): 197-213.

Garrison, Roger W. 1991. New Classical and Old Austrian Economics: Equilibrium Business Cycle Theory in Perspective. *The Review of Austrian Economics* 5(1): 91-103.

Garrison, Roger W. 1997. The Lachmann Legacy: An Agenda for Macroeconomics. *South African Journal of Economics* 65(4): 459-481.

Garrison, Roger W. 2001. *Time and Money: The Macroeconomics of Capital Structure*. London: Routledge.

Garrison, Roger W. 2002. Capital, Interest, and Professor Kirzner. *Journal des Economistes et des Etudes Humains* 12(2-3): 343-356.

Gigerenzer, Gerd, Zeno Swijtink, Theodore Porter, Lorraine Daston, John Beatty, and Lorenz Krüger. 1989. *The Empire of Chance: How Probability Changed Science and Everyday Life*. Cambridge: Cambridge University Press.

Gilboa, Itzhak, and David Schmeidler. 1989. Maxmin Expected Utility with a Non-Unique Prior. *Journal of Mathematical Economics* 18(2): 141-153.

Gilboa, Itzhak, Andrew W. Postlewaite, and David Schmeidler. 2008. Probability and Uncertainty in Economic Modeling. *Journal of Economic Perspectives* 22(3): 173-188.

Gollier, Christian. 2001. *The Economics of Risk and Time*. Cambridge, MA: MIT Press.

Grampp, William D. 1948. Adam Smith and the Economic Man. *Journal of Political Economy* 56(4): 315-336.

Greaves, Percy L., Jr. 1974. Review of Competition and Entrepreneurship by Israel M. Kirzner. *Wertfrei: A Review of Praxeological Science* 1: 17–20.

Grossman, Sanford J. and Joseph E. Stiglitz. 1980. On the Impossibility of Informationally Efficient Markets. *American Economic Review* 70(3): 393-408.

Gunning, J. Patrick. 1989. Mises and the Evenly Rotating Economy. *The Review of Austrian Economics* 3(1): 123-135.

Gunning, J. Patrick. 2005. Interest: In Defense of Mises. *The Quarterly Journal of Austrian Economics* 8(3): 81-91.

Hayek, Frederick. [1928] 1994. The System of Intertemporal Price Equilibrium and Movements in the 'Value of Money'. (*trans*.) William Kirby. In *Classics in Austrian Economics: A Sampling in the History of a Tradition. Volume III: The Age of Mises and Hayek*, (*ed*.) Israel M. Kirzner. London: William Pickering.

Hayek, F. A. 1931. Reflections on the Pure Theory of Money of Mr. J. M. Keynes: Part I. *Economica* 33(2): 270-295.

Hayek, Friedrich A. [1937a] 1948. Economics and Knowledge. In *Individualism and Economic Order*. Chicago: University of Chicago Press.

Hayek, Friedrich A. [1937b] 1971. *Monetary Nationalism and International Stability*. New York: Augustus M. Kelley.

Hayek, F. A. 1941. *The Pure Theory of Capital*. London: Routledge & Kegan Paul.

Hayek, F. A. [1945] 1948. The Use of Knowledge in Society. Reprinted in *Individualism and Economic Order*, F. A. Hayek (*ed*.). Chicago: University of Chicago Press.

Hayek, F. A. [1946] 1948. The Meaning of Competition, Reprinted in *Individualism and Economic Order*, F. A. Hayek (*ed*.). Chicago: University of Chicago Press.

Hayek, F. A. [1952a] 1979. *The Counter-Revolution of Science*. Indianapolis, Ind: The Liberty Fund Inc.

Hayek, F. A. 1952b. *The Sensory Order: A Inquiry of Theoretical Psychology*. Chicago: University of Chicago Press.

Hayek, F. A. 1960. *The Constitution of Liberty*. London: Routledge & Kegan Paul Ltd.

Hayek F. A. 1962 [1967]. Rules, Perception and Intelligibility, in *Studies in Philosophy, Politics and Economics*, F. A. Hayek (*ed*.). Chicago: University of Chicago Press.

Hayek, F. A. [1968] 2002. Competition as a Discovery Procedure. (*trans*.)Marcellus S. Snow. *The Quarterly Journal of Austrian Economics* 5(3): 9-23.

Hayek, F. A. 1973. *Legislation and Liberty, Volume I: Rules and Order*. Chicago: University of Chicago Press.

Hayek, F. A. von.1976. *Law, Legislation and Liberty, Volume II: The Mirage of Social Justice*. Chicago: The University of Chicago Press.

Hayek, F. A. 1978. *New Studies in Philosophy, Politics, Economics and the History of Ideas*. Chicago: University of Chicago Press.

Hayek, F. A. 1988. *The Fatal Conceit: The Errors of Socialism*. London: Routledge.

Haynes, John. 1895. Risk as an Economic Factor. *Quarterly Journal of Economics* 9(4): 409-449.

Hazlitt, Henry. 1974. Review of Competition and Entrepreneurship by Israel M. Kirzner. *The Freeman* 24(12): 756–60.

Heidegger, Martin. [1953] 1996. *Being and Time*, (*trans*.) Joan Stambaugh. Albany, NY: State University of New York.

High, Jack. 1982. Alertness and Judgment: Comment on Kirzner, in *Method, Process and Austrian Economics: Essays in Honor of Ludwig von Mises*, Israel M. Kirzner (*ed*.). Lexington, MA:

Lexington Books.

Hirshleifer, J. 1973. Where Are We in the Theory of Information? *The American Economic Review 63*(*2*): 31-39.

Hirshleifer, J., and John G. Riley. 1995. *The Analytics of Uncertainty and Information*. Cambridge: Cambridge University Press.

Holcombe, Randall G. 1998. Entrepreneurship and Economic Growth. *The Quarterly Journal of Austrian Economics* 1(2): 45-62.

Holcombe, Randall G. 2003. Progress and Entrepreneurship. *The Quarterly Journal of Austrian Economics* 6(3): 3-26.

Hoppe, Hans Hermann. 1995. *Economic Science and the Austrian Method*. Auburn, AL.: Ludwig von Mises Institute.

Hoppe, Hans-Hermann. 1997. On Certainty and Uncertainty, or: How Rational Can Our Expectations Be? *Review of Austrian Economics* 10(1): 49-78.

Hoppe, Hans-Hermann. 1999. Time Preference, Government, and the Process of Decivilization: From Monarchy to Democracy. In *The Costs of War* (*ed*.) John V. Denson. London: Transaction Publishers.

Hoppe, Hans-Hermann. 2001. *Democracy - The God that Failed: The Economics & Politics of Monarchy, Democracy & Natural Order*. New Brunswick, NJ: Transaction.

Howden, David. 2008. *Fama's Efficient Market Hypothesis and Mises' Evenly Rotating Economy: Comparative Constructs*. Unpublished manuscript.

Huerta de Soto, Jesús. 2004. La Teoría de la Eficiencia Dinámica. *Procesos de Mercado: Revista Europea de Economía Política* 1(1): 11-71.

Huerta de Soto, Jesús. 2005. *Socialismo, Cálculo Económico, y Función Empresarial*. Madrid, Unión Editorial.

Hülsmann, Jörg Guido. 1997. Knowledge, Judgment and the Use of Property in Society. *Review of Austrian Economics* 10(1): 23-48.

Hülsmann, Jörg Guido. 1999. Entrepreneurship and Economic Growth: Comment on Holcombe. *The Quarterly Journal of Austrian Economics* 2(2): 63-65.

Hülsmann, Jörg Guido. 2002. A Theory of Interest. *The Quarterly Journal of Austrian Economics* 5(4): 77-110.

Hülsmann, Jörg Guido. 2003. Facts and Counterfactuals in Economic Law. *Journal of Libertarian Studies* 17(1): 57-102.

Khaneman, Daniel, and Amos Tversky. 1972. Subjective Probability: A Judgment of Representativeness. *Cognitive Psychology* 3(3): 430-454.

Kahneman, Daniel, and Amos Tversky. 1973. On the Psychology of Prediction. *Psychological Review* 80(4): 237-251.

Kahneman, Daniel, and Amos Tversky. 1979a. A Prospect Theory: An Analysis of Decisions Under Risk. *Econometrica* 47(2): 263-291.

Keynes, John Maynard. [1921] 1962. *A Treatise on Probability*. New York: Harper Torchbooks.

Kirzner, Israel M. 1973. *Competition and Entrepreneurship*. Chicago:University of Chicago Press.

Kirzner, Israel M. 1979. *Perception, Opportunity, and Profit: Studies in the Theory of Entrepreneurship*. Chicago: The University of Chicago Press.

Kirzner, Israel M. 1982. Entrepreneurship, Uncertainty, and Austrian Economics, in *Austrian Economics: Tensions and New Directions*, Bruce J. Caldwell and Stephan Boehm, (*eds*.). Boston: Kluwer Academic Publishers.

Kirzner, Israel M. 1982b. Uncertainty, Discovery, and Human Action: A study of the Entrepreneurial Profile in the Misesian System. in *Method, Process and Austrian Economics: Essays in Honor of Ludwig von Mises*, Israel M. Kirzner (*ed*.). Lexington, MA: Lexington Books.

Kirzner, Israel M. 1985. *Discovery and the Capitalist Process*. Chicago: University of Chicago Press.

Kirzner, Israel M. 1988. Advertising in an Open-Ended Universe, forward to *Advertising and the Market Process*, Robert B. Ekelund Jr. and David S. Saurman, (*eds*.). San Francisco: Pacific Research Institute for Public Policy.

Kirzner, Israel M. 1989. *Discovery, Capitalism, and Distributive Justice*. New York: Basil Blackwell Inc.

Kirzner, Israel M.1992. *The Meaning of the Market Process*. London: Routledge.

Kirzner, Israel M. 1993. The Pure Time-Preference Theory of Interest: An Attempt At Clarification. In *The Meaning of Ludwig von Mises*, (*ed*.) Jeffrey M. Herbener. Norwell, MA: Kluwer Academic Publishers.

Kirzner, Israel M. 1996. *Essays on Capital and Interest*. Brookfield, VT: Edward Elgar.

Kirzner, Israel M. 1997. Entrepreneurial Discovery and the Competitive Market Process: An Austrian Approach. *Journal of Economic Literature* 35(1): 60-85.

Kirzner, Israel M. 1998. Creativity an/or Alertness: A Reconsideration of the Schumpeterian Entrepreneur. *The Review of Austrian Economics* 11(1/2): 5-17.

Kirzner, Israel M. 2005. Information-Knowledge and Action-Knowledge. *Econ Journal Watch* 2(1): 75-81.

Knight, Frank H. 1921. *Risk, Uncertainty and Profit*. Boston, MA: Houghton Mifflin.

Lachmann, Ludwig M. [1945] 1994. A Note on the Elasticity of Expectations. Reprinted in Don Lavoie (*ed*.) *Expectations and the Meanings of Institutions: Essays in Economics by Ludwig Lachmann*. London: Routledge.

Lachmann, Ludwig M. [1956] 1978. *Capital and Its Structure*. Kansas City: Sheed Andrews and McMeel, Inc.

Lachmann, Ludwig M. 1959. Professor Shackle on the Economic Significance of Time. *Metroeconomica* 11(1-2): 64-73.

Lachmann, Ludwig M. 1976a. From Mises to Shackle: An Essay on Austrian Economics and the Kaleidic Society. *Journal of Economic Literature* 14(1): 54-62.

Lachmann, Ludwig M. 1976b. On the Central Concept of Austrian Economics: Market Process. In *The Foundations of Modern Austrian Economics*, Edwin G. Dolan (*ed*.). Kansas City: Sheed & Ward, Inc.

Lachmann, Ludwig.1977. Methodological Individualism and the Market Economy. Reprinted in *Capital, Expectations and the Market Process* (*ed*.) Walter Grinder. Kansas City: Sheed Andrews and McMeel.

Lange, Oscar. 1944. *Price Flexibility and Employment*. Cowles Commission for Research in Economics, Monograph 8. Bloomington, IN: The Principia Press, Inc.

Langlois, Richard N., and Metin M. Cosgel. 1993. Frank Knight on Risk, Uncertainty, and the Firm: A New Interpretation. *Economic Inquiry* 31(3): 456-465.

Langlois, Richard N., and Paul L. Robertson. 1995. *Firms, Markets and Economic Change*. London: Routledge.

Lavoie, Don. 1985. *National Economic Planning: What is Left?* Cambridge, MA: Ballinger.

Lewin, Peter. 1997. Rothbard and Mises on Interest: An Exercise in Theoretical Purity. *Journal of the History of Economic Thought* 19(1): 141-159.

Lewin, Peter. 1999. *Capital in Disequilibrium: The Role of Capital in a Changing World*. London: Routledge.

Lipman, Barton L. 1999. Decision Theory Without Logical Omniscience: Toward an Axiomatic Framework for Bounded Rationality. *The Review of Economic Studies* 66(2): 339-361.

Loasby, Brian J. 1982. The Entrepreneur in Economic Theory. *The Scottish Journal of Political Economy* 29(3): 235-245.

Loasby, Brian J. 1989. *The Mind and Method of the Economist, A Critical Appraisal of Major Economists in the Twentieth Century*. Aldershot, UK: Edward Elgar.

Loomes, Graham, Chris Starmer., and Robert Sugden. 1989. Preference Reversal: Information-processing Effect or Rational Non-transitive Choice? *The Economic Journal* 99 (395): 140-151.

Loomes, Graham, Chris Starmer., and Robert Sugden. 1991. Observing Violations of Transitivity by Experimental Methods. *Econometrica* 59(2): 425-440.

Machlup, Fritz. 1972. *The Production and Distribution of Knowledge in the United States*. Princeton, NJ: Princeton University Press.

Maritain, Jacques. 1946. *Formal Logic*. New York: Sheed and Ward.

Mars, J. 1951. A Study in Expectations. Reflections on Shackle's 'Expectation in Economics'. Part II. *Bulletin of Economic Research* 3(1): 1-35.

Maturana, H. R. and F. J. Varela. 1980. *Autopoiesis and Cognition. The Realization of the Living*. Dordrecht, D. Reidel Publishing Company.

Marshall, Alfred. [1890] 1961. *Principles of Economics*. London: Macmillan.

Menger, Carl. [1871] 2007. *Principles of Economics*. (t*rans*.)James Dingwall and Bert F. Hoselitz. Auburn, AL: Ludwig von Mises Institute.

Meredith, G. Patrick. [1954] 1957. Methodological Considerations in the Study of Human Anticipation. In C. F. Carter, G. P. Meredith, and G. L. S. Shackle (e*ds*.) *Uncertainty and Business Decisions*. Liverpool: Liverpool University Press.

Mises, Ludwig von. [1912] 1980. *The Theory of Money and Credit*. (*trans*.) H. E. Batson. Indianapolis, IN: Liberty Classics.

Mises, Ludwig von. [1920] 1975. Economic Planning in the Socialist Commonwealth. Reprinted in *Collectivist Economic Planning*, F. A. Hayek (*ed*.). Clifton, NJ: Augustus M. Kelly.

Mises, Ludwig von. [1933] 1976. *Epistemological Problems of Economics*, George Reisman (*trans*.). New York: New York University Press.

Mises, Ludwig von. 1936. *Socialism: An Economic and Sociological Analysis,* J. Kahane (*trans*.). London: Jonathan Cape.

Mises, Ludwig von. [1944] 1994. Monetary Stabilization and Cyclical Policy. In *Classics in Austrian Economics: A Sampling in the History of a Tradition. Volume III: The Age of Mises and Hayek*, Israel M. Kirzner (*ed*.). London: William Pickering.

Mises, Ludwig von. [1949] 1998. *Human Action: A Treatise on Economics*. Auburn, AL: Ludwig von Mises Institute.

Mises, Ludwig von. [1951] 2008. *Profit and Loss*. Auburn, AL: Ludwig von Mises Institute.

Mises, Ludwig von. [1957] 1985. *Theory and History*. Auburn, AL: Ludwig von Mises Institute.

Mises, Ludwig von. 1960. Epistemological Relativism in the Sciences of Human Action. reprinted in *Money, Method and the Market Process*, Richard M. Ebeling (*ed*.). Auburn, AL: Ludwig von Mises Institute.

Mises, Ludwig von. [1961] 1994. The Treatment of 'Irrationality' in the Social Sciences. In *Classics in Austrian Economics: A Sampling in the History of a Tradition. Volume III: The Age of Mises and Hayek*, (*ed*.) Israel M. Kirzner. London: William Pickering.

Mises, Ludwig von. [1962] 1978. *The Ultimate Foundation of Economic Science*. Kansas City: Sheed Andrews and McMeel, Inc.

Mises, Richard von. [1928] 1957. *Probability, Statistics and Truth*. London: George Allen and Unwin Ltd.

Moss, Laurence S. 1978. The Emergence of Interest in a Pure Exchange Economy: Notes on a Theorem Attributed to Ludwig von Mises. In *New Directions in Austrian Economics*, (*ed*.) Louis M. Spadaro. Kansas City: Sheed Andrews and McMeel, Inc..

Murphy, Robert P. 2003. *Unanticipated Intertemporal Change in Theories of Interest*. PhD dissertation, New York University.

Nisbett, R. E., D. H. Krantz, C. Jepson, and Z. Kunda. 1983. The Use of Statistical Heuristics in Everyday Inductive Reasoning. *Psychology Review* 90(4): 339-363.

O'Connor, D. J. [1954] 1957. Uncertainty as a Philosophical Problem: 2. In C. F. Carter, G. P. Meredith, and G. L. S. Shackle (*eds*.) *Uncertainty and Business Decisions*. Liverpool: Liverpool University Press.

O'Driscoll, Gerald O. Jr., and Mario J. Rizzo. 1985. *The Economics of Time and Ignorance*. Mew York: Basil Blackwell.

Pasour, E. C. Jr.. 1989. The Efficient-Markets Hypothesis and Entrepreneurship. *The Review of Austrian Economics* 3(1): 95-107.

Pellengahr, Ingo. 1996. *The Austrian Subjectivist Theory of Interest: An Investigation into the History of Thought*. Frankfurt am Main, Germany: Peter Lang.

Pigou, Arthur C. 1932. *The Economics of Welfare*. London: Macmillan.

Polanyi, Michael. 1958a. *Personal Knowledge: Towards a Post-Critical Appraisal*. London: Routledge & Kegan Paul.

Polanyi, Michael. 1958b. *The Study of Man*. London: Routledge & Kegan Paul.

Polanyi, Michael. 1967. *The Tacit Dimension*. Garden City, NY: Anchor Books.

Polanyi, Michael. 1969a. "Knowing and Being" in *Knowing and Being*, M. Greene (*ed*.). London: Routledge.

Polanyi, Michael. 1969b. "The Logic of Tacit Inference" in *Knowing and Being*, M. Greene (*ed*.). London: Routledge.

Polanyi, Michael, and Harry Prosch. 1975. *Meaning*. Chicago:University of Chicago Press.

Pozzali, Andrea. 2008. Tacit Knowledge, Implicit Learning and Scientific Reasoning. *Mind & Society* 7(2): 227-237.

Reisman, George. [1990] 1996. *Capitalism: A Treatise on Economics*. Ottawa, IL: Jameson Books.

Rizzo, Mario J. 1979. *Time, Uncertainty, and Disequilibrium*. Lexington, MA: Lexington Books.

Rizzo, Mario J. 1994. Time in Economics. In *The Elgar Companion to Austrian Economics*, (*ed*.) Peter Boettke. Aldershot, UK: Edward Elgar.

Rizzo, Mario J.1995. Time and Ignorance: After 10 Years. Economic Research Reports, New York, New York University C.V. Starr Center for Applied Economics (August).

Rizzo, Mario J. 2000. Real Time and Relative Indeterminacy in Economic Theory, in *Time in Contemporary Intellectual Thought*, Patrick Baert (*ed*.). Amsterdam: Elsevier Science B.V.

Robinson, Joan. 1980. Time in Economic Theory. *Kyklos* 33(3): 219-229.

Rothbard, Murray N. [1962] 2004. *Man, Economy and State with Power and Market*. Auburn, AL: Ludwig von Mises Institute.

Rothbard, Murray N. 1974. Review of Competition and Entrepreneurship by Israel M. Kirzner. *Journal of Economic Literature* 12(3): 902–04.

Rothbard, Murray N. [1985]1997. Professor Kirzner on Entrepreneurship, in *The Logic of Action Two: Applications and Criticism from the Austrian* School. Lyme, N.H.: Edward Elgar Publishing.

Rothbard, Murray N. 1987. Breaking Out of the Walrasian Box: The Cases of Schumpeter and Hansen. *Review of Austrian Economics* 1(1): 97-108.

Rothbard, Murray N. 2006. *Classical Economics: An Austrian Perspective on the History of Economic*

Thought, Volume II. Cheltenham, UK: Edward Elgar.

Salerno, Joseph T. 1994a. Ludwig von Mises's Monetary Theory in Light of Modern Monetary Thought. *The Review of Austrian Economics* 8 (1): 71-115.

Salerno, Joseph T. 1994b. Reply to Leland B. Yeager on Mises and Hayek on Calculation and Knowledge. *Review of Austrian Economics* 7(2): 111-125.

Salerno, Joseph T. 2006. A Simple Model of the Theory of Money Prices. *Quarterly Journal of Austrian Economics* 9(4): 39-55.

Salerno, Joseph T. 2008. The Entrepreneur: Real and Imagined. *Quarterly Journal of Austrian Economics* 11(3-4): 188-207.

Sautet, Frédéric. 2000. *An Entrepreneurial Theory of the Firm*. London: Routledge.

Savage, Leonard J. 1954. *The Foundations of Statistics*. New York: John Wiley and Sons.

Schmeidler, David. 1989. Subjective Probability and Expected Utility Without Additivity. *Econometrica* 57(3): 571-587.

Schumpeter, Joseph A. [1911] 2002. *Theorie der wirtschaflichen Entwicklung*, chapters 2 and 7, (*trans*.) Markus C. Becker and Thorbjørn Knudsen. *American Journal of Economics and Sociology* 61(2): 405-437.

Schumpeter, Joseph A. 1942. *Capitalism, Socialism and Democracy*. New York: Harper.

Schumpeter, Joseph A. 1949. *The Theory of Economic Development*. Cambridge, MA: Harvard University Press.

Schumpeter, Joseph A. 1989. The Creative Response in Economic History, reprinted in *Essays on Entrepreneurs, Innovations, Business Cycles and the Evolution of Capital*, Richard V. Clemence (*ed*.). New Brunswick: Transaction Publishers.

Schumpeter, Joseph A. 1991. Comments on a Plan for the Study of Entrepreneurship, in Richard Swedberg (*ed*.), *The Economics and Sociology of Capitalism*. Princeton, NJ: Princeton University Press.

Shackle, G. L. S. 1938. *Expectations, Investment, and Income*. Oxford: Oxford University Press.

Shackle, G. L. S. [1949a] 1955. A Non-Additive Measure of Uncertainty. Reprinted in *Uncertainty in Economics*. Cambridge: The University Press.

Shackle, G. L. S. [1949b] 1955. Expectations in Economics': Some Critics Answered. Reprinted in *Uncertainty in Economics*. Cambridge: The University Press.

Shackle, G. L. S. [1949c] 1955. Probability and Uncertainty. Reprinted in *Uncertainty in Economics*.

Cambridge: The University Press.

Shackle, G. L. S. 1952a. *Expectation in Economics*. Cambridge: The University Press.

Shackle, G. L. S. [1952b] 1955. On the Meaning and Measure of Uncertainty: I. Reprinted in *Uncertainty in Economics*. Cambridge: The University Press.

Shackle, G. L. S. [1953] 1955. On the Meaning and Measure of Uncertainty: II. Reprinted in *Uncertainty in Economics*. Cambridge: The University Press.

Shackle, G. L. S. [1954] 1957. Expectation in Economics. In C. F. Carter, G. P. Meredith, and G. L. S. Shackle (*eds*.) *Uncertainty and Business Decisions*. Liverpool: Liverpool University Press.

Shackle, G. L. S. 1958. *Time in Economics*. Amsterdam, NL: North-Holland Publishing Company.

Shackle, G. L. S. 1961. *Decision Order and Time in Human Affairs*. Cambridge: The University Press.

Shackle, G. L. S. 1965. *A Scheme of Economic Theory*. Cambridge: Cambridge University Press.

Shackle, G. L. S. 1972. *Epistemics and Economics*. Cambridge: Cambridge University Press.

Shackle, G. L. S. 1973. *An Economic Querist*. Cambridge: Cambridge University Press.

Shackle, G. L. S. 1994. The Imagined Deemed Possible: The Germ of the Process of History, in *Uncertainty in Economics*, David G. Dickinson, Michael J. Driscoll, and Somanth Sen, (*eds*.). Aldershot, UK: Edward Elgar Publishing Ltd.

Shafer, Glenn. 1986. Savage Revisited. *Statistical Science 1*(*4*): 463-486.

Shostak, Frank. 1999. Economic Growth and Its Causes: Comment on Holcombe. *The Quarterly Journal of Austrian Economics* 2(2): 67-71.

Simon, Herbert A. 1957. *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*. New York: John Wiley and Sons.

Simon, Herbert A. 1978. Rationality as a Process and as Product of Thought. *The American Economic Review* 68(2), Papers and Proceedings of the Ninetieth Annual Meeting of the American Economic Association: 1-16.

Simon, Herbert A. 1986. Rationality in Psychology and Economics. *The Journal of* Finance 59(4), *Part 2: The Behavioral Foundations of Economic Theory*: S209-S224.

Slovic, Paul, Baruch Fischoff, and Sarah Lichenstein. [1980] 1981. Facts Versus Fears: Understanding Perceived Risk. Reprinted in Daniel Kahneman, Paul Slovic and Amos Tversky, (*eds*.) *Judgment Under Uncertainty: Heuristics and* Biases. New York: Cambridge University Press.

Steele, G. R. 2002. Hayek's Sensory Order. *Theory & Psychology* 12(3): 125-147.

Strigl, Richard von.. 2000. *Capital and Production*. Auburn, Al: The Ludwig von Mises Institute.

Taleb, Nassim Nicholas. 2004. *Fooled by Randomness*. New York: Random House.

Taleb, Nassim Nicholas. 2007. *The Black Swan: The Impact of the Highly Improbable*. New York: Random House.

Theirry, Aimar. 2008. Self-ignorance: Towards an Extension of the Austrian Paradigm. *Review of Austrian Economics* 21(1): 23-43.

Thomsen, Esteban F. 1992. *Prices and Knowledge: A Market-Process Perspective*. London: Routledge.

Tulving, Endel. 1972. Episodic and Semantic Memory. In *Organization of* Memory, E. Tulving and W. Donaldson, (*eds*.). New York: Academic Press.

Tversky, Amos, and Daniel Kahneman. [1971] 1982. Belief in the Law of Small Numbers. Reprinted in Daniel Kahneman, Paul Sovic and Amos Tversky (*eds*.) *Judgment Under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.

Tversky, Amos, and Daniel Khaneman. 1974. Judgment Under Uncertainty: Heuristics and Biases. *Science* 185(4157): 1124-1131.

Tversky, Amos, and Daniel Kahneman. 1980. Schemes in Judgments Under Uncertainty. In M. Fishbein (*ed*.) *Progress in Social Psychology.* Hillsdale, NJ: Erlbaum.

van den Hauwe, Ludwig. 2008. John Maynard Keynes and Ludwig von Mises on Probability. *Procesos de Mercado: Revista Europea de Economía Política* 5(1): 11-50.

Vaughn, Karen I. 1994. *Austrian Economics in America: the Migration of a Tradition*. Cambridge: Cambridge University Press.

Vercelli, Alessandro. 1999. The Recent Advances in Decision Theory Under Uncertainty. In Luigi Luini *(ed.) Uncertain Decision: Bridging Theory and Experiments*. London: Kluwer Academic Publishers.

von Neumann, John, and Oskar Morgenstern. 1944. *Games and Economic Behavior.* New York: Wiley.

Watikins, J. W. N. 1955. Decisions and Uncertainty: A Review. *British Journal of the Philosophy of Science* 6 (21): 66-78.

Wicksell, Knut. [1898] 1965. *Interest and Prices*, R. F. Kahn (*trans*.). New York: Augustus M. Kelley.

Wieser, Friedrich von. [1889] 1956. *Natural Value*. (*trans*.) Christian A. Malloch. New York: Kelley & Millman.

Williamson, Oliver E. 1975. *Markets and Hierarchies.* New York: Free Press.

Williamson, Oliver E. 1981. The Economies of Organization: The Transaction Cost Approach. *American Journal of Sociology* 87(3): 548-577.

Yates, Steven. 2005. What Austrian Scholars Should Know About Logic (and Why). *Quarterly Journal of Austrian Economics* 8(3): 39-57.

BOOK II

ACTION AND ITS STRUCTURES

I. INTRODUCTION

Part I of this work explained that action has many components that affect the way decisions are made. From the fundamental basis of logic, our empirical senses continue learning additional information which augments our knowledge base. Risk and uncertainty forever shroud our decisions, and leave us under a dark cloud of ignorance concerning the future. We are never certain that our expected results will obtain from our actions, however, we undertake them under the pretense of a plan that we think will obtain, given our expectations of results and their accompanying level of projected uncertainty. The process was shown to be reduced to a comparison between two expected results – the focus-gain and focus-loss – which entice us to act depending on our personal propensity for uncertainty.[255]

In this part, we build considerably from Eugen von Böhm-Bawerk's concept of the structure of production. In fact, what Böhm-Bawerk has illustrated is that action is not a timeless occurrence, but actually proceeds temporally along a structure. As actions are combined with other actions, larger events are allowed to manifest.

All action has a unique goal in mind – the satisfaction of a want. As all wants are not created equally, we find that there is a dichotomy of ways that these desires may be satiated. The first are simple uni-stage desires. These are those wants we wish satisfied which require no significant temporal passage, and require no complex process to obtain. For example, an itch on our arm may provide an opportunity for a want to be satisfied. This can easily be seen to be accomplished without the need for complex time-consuming instruments; we may satisfy the want instantaneously or almost so. However, this range of wants is very limited compared to the greater total that we wish satisfied. For there are those wants which require very complex structures to complete. For example, we may wish to take a drive in our car on a Sunday afternoon to see some of the country-side. The auto that is required for this cannot be produced in any simple manner. Instead, we see that the collaboration of many individuals using intermediary goods will be necessary to create this good.

Böhm-Bawerk's structure of production describes those processes which are complex and necessary to produce the goods required to satisfy that set of wants that are too complex to be possible for an individual without either waiting a temporal period prior to satisfaction, or enlisting the help of others to make the satisfaction possible.

[255] Although, we should note that the trade-off need not always be between only two expected outcomes. For reasons of simplicity, at this point, the analysis has been limited to only one unique set of focus-pairs. The possibility remains, however, that the comparison may be more than binary.

Much work has previously been focused on elaborating this concept – the structure of production. However, the influence of Hayek's (1935) triangle as an heuristic tool for exploring the possibilities that production processes yield has had both advantages and disadvantages. The three chief disadvantages that are explored and rectified in this present work are: 1) the confusion over production as a temporal process *or* a capital intensive process, 2) the neglect of a corresponding structure of consumption, and 3) the exclusion of fixed capital with the focus paid almost exclusively to circulating capital in production.

Production has been previously a somewhat misunderstood concept. What is it that we mean when we say that we are producing something? Followers using Hayekian triangles have wanted to explain production as a series of stages, where goods of higher order are continually transformed into goods of lower, and eventually, the lowest order – consumers' goods. However, this has created the deficiency of thinking of production solely in terms of pre-defined stages. Garrison (2001), for example, uses five stages of production to illustrate his structure of production. The reason as to why five stages is chosen is that it gives just the right amount of complexity to see that production is a timely process, but without adding too much complexity so as to make the triangle's use as an heuristic device useless, or unnecessarily complex.[256]

As a result of this line of thought, the structure of consumption has been neglected, while production has received much attention. The reason this is so is that while it is clear that production does, in fact, have stages (i.e., manufacturing, distribution, retail, etc.), there are no such clearly delineated stages to consumption. Goods are obtained at t_0 and used throughout their useful life, whether long or short. The influence of viewing production as stage-specific is somewhat of a useful heuristic device, but suffers grave theoretical deficiencies.

For what is production at its core other than a series of value adding actions? And, the corollary, what is consumption at its core other than a series of value-subtracting actions? As we consume a good, we personally gain value (the direct end of consumption), and at the same time, the good in question loses value to us. Hence, consumption is the exact corollary to production – both actions affect the value in a good,and do so not necessarily in specific stages, but through a series of infinitesimally small individual actions. If a structure of production exists, based upon this series of intertwined value-adding actions that create consumers' goods, then a similar, and off-setting, structure of consumption exists

[256] Anyone doubting the use of the "stages" concept in Austrian production theory should take a random sampling of the literature where the term and concept are thoroughly ingrained in discussions. Foldvary (2006: 800), for example, states: "Austrian School economists have recognized that capital goods have a structure based on two elements: stages and time."

that continually removes value from these same goods as they are used.

The emphasis laid on Hayek's triangles lay almost exclusively on circulating capital, at the expense of its fixed counterpart. As circulating capital continually progresses through the structure, it is transformed into a good of lower and lower order until eventually it becomes a consumers' good. However, this viewpoint overlooks the necessity of fixed capital, with detrimental consequences.

First, we see that production is split between circulating and fixed capital. The split is made as entrepreneurs forecast the demand for consumers' goods. This in turn stems from the time preference that individuals display, as they chose to forgo present consumption and save for the future. The relationship between circulating and fixed capital production, then, is fundamentally *derived* from the consumers' trade-off between consumption and saving. As consumers demand fewer consumption goods, a relative shift occurs to produce more fixed capital. This shift has a secondary effect owing to the need to replace depreciated fixed capital. An increase in production of this capital will also require an additional increase to satisfy its own need of an allowance for depreciation.

Hence, the structure of production is comprised of three sub-structures. The first is the structure of circulating capital, the most developed action structure until now. Second is the structure of fixed capital. These produced means of production are necessary to create the circulating capital and hence, the structure thereof. Last, owing to the finite life of fixed capital, a sub-structure need be created to allow for the depreciation that occurs with fixed capital. Hence, a portion of fixed capital production will be directed towards the allowance to maintain the existing structure, and therefore, keep the output level steady.

Changes in the structure of consumption reverberate throughout the structure of production to determine and direct what types of capital are to be produced.

Lastly, much confusion has arisen over what exactly it means to lengthen the structure of production. One of Böhm-Bawerk's greatest contributions is in developing his theory whereby lengthening the productive structure will result in increased productivity. There are two ways that this has been interpreted to this point. The first is that the structure of production actually undergoes a temporal expansion. Hence, productivity is increased as a result of longer temporal processes being used. This is one interpretation, but it is erroneous and leads to poor conclusions. Hayekian triangles have bred an emphasis on this interpretation, as the x-axis is typically defined as being a temporal axis. However, as we will see, even if we treat it as a temporal axis, it is difficult to conclude that production takes place at a point in time more distant from the consumer or not.

Instead, as we shall see, a lengthening of the structure of production involves a shift in the ratio between circulating and fixed capital. If the ratio is high, than much circulating capital is being produced at the expense of fixed capital. Hence, this is equivalent to saying that more consumers' goods are in production, than capital with which to create them. The build-up in capital necessary to increase consumption capacity comes from two factors.

The first is that the ratio between circulating to fixed capital decreases. In this way, there is a relative shift along the structure of production, with greater amounts of production being directed toward producing that capital which will produce consumers' goods (fixed capital). This need not imply a temporal lengthening of the structure of production. Indeed, that conclusion would be somewhat paradoxical as the goal of production is to satisfy our wants in the least amount of time, or with the least amount of goods renunciated to do so. Hence, to lengthen the structure of production to increase productivity may be an unfortunate terminology, as what we really wish to imply is that we increase the amount of fixed capital relative to circulating goods produced along this structure.

There is one additional way that the structure may be lengthened. The use of fixed capital of greater durability will mean that less production need be directed towards an allowance for its depreciation. Hence, production may be maintained at its current output more easily, or through the production of a relatively greater amount of circulating capital. The intensity of the degree of capitalistic processes employed in the structure must be viewed relatively, then, depending on the degree of durability the goods in question have.

We find that with these small short-comings corrected, the concept of production and consumption take on increased significance. As we set these actions within their proper contexts of their respective structures, a complete and intertwined system of action results. We are able to see concretely through the use of these structures how individuals may create more complex want satisfactions through the use of consumers' goods and hence, increase their own satisfaction over time. Action's *structures* become an integral part of the analysis, which yield their most enlightening results when viewed as a unified whole.

II. ACTION AND ITS FORMS

In Part I, we laid out a foundation of action. The process through which the limitations of our mind involving the perception, storage, and use of knowledge necessarily provided a boundary to which we could use the absolute logic system endowed in us as humans through life. The concepts of risk and uncertainty become inextricably linked to this action process, as we move to an uncertain future which implies an ever-changing present.

The indeterminacy of the future was proven to be not a hindrance to the development of society, but a boon, as it gave rise to ever more possibilities for expansion and progression. The entrepreneur has been identified as the individual who conducts this role in moving the economy forward through three different methods. First is the foresight to see the future state of affairs that will be desired, and acting in the present to move towards this condition. Second is the bearing of uncertainty in the present, as tasks which are non-repeatable involve an uncertain element which can be mitigated in no way other than judgment and foresight from the entrepreneur. Lastly, and somewhat controversially, we saw that *mere* risk-mitigators in the present also move the economy forward. As the entrepreneurial process is defined as that which moves consumers toward ever greater degrees of ultimate want satisfaction, risk-mitigating entrepreneurs can function in the present to move to this more satisfying future state by reducing the costs associated with exchange. However, it was also pointed out, and should be remembered, that this secondary class of entrepreneurs – risk-mitigators – may only function as long as their exists an active class of uncertainty bearing entrepreneurs. The discovery of fresh future states is what continually shifts the present state of risk, and hence, gives rise to the continual need for an entrepreneurial class to mitigate this risk.

Two issues exist in our theory of action as has been previously developed that now bear attention. Action has been heretofore viewed in a very homogeneous manner – no distinction has been made concerning the categories of ends that an action may serve to complete. In actuality, our theory of action has been mostly implicitly concerned with that category of action which only *produces* want satisfaction through entrepreneurial action. It has also made little note of the actual resources through which we proceed with this action. What must be developed further, then, is a theory of action which looks not only at the different categories that may be manifested, but also a theory of the resources that are used to realize this future state of affairs that action drives toward.

Production – exchange – consumption. This trichotomy of actions forms the praxeologic categories through which our future state of affairs may be moved towards. This trichotomy has been recognized, implicitly at least, for much time already. However, recent years have seen attacks on these categories. Barnett and Block (2007) argue that action consists of a binary distinction between consumption and production. As we will outline, the omission of exchange as a pure category of action is done with great neglect, and leads to grave consequences. Some of these grave consequences have already begin to appear, notably in Barnett and Block's (2005) work demonstrating that, as there may only be a dichotomy of action, there may only be a dichotomy of goods in existence to both aid and result from these actions.[257] The demonstration of goods' categories will have to wait for a further chapter, but what must be shown in the following pages – the trichotomy of action's categories – will provide the core of the theoretical foundation that will be used throughout this present part.

[257] See also, Rothbard (1962: 33) for an earlier view of goods' nature being binary.

1. Consumption

Earlier, we have identified the entrepreneurial process as that which moves individuals to ever increasing levels of want-satisfaction. Want-satisfaction as a term is unobjectionable, and conveys all that we wish to express, however, it may prove helpful to use somewhat different terminology to better express what it is exactly what we mean by "want-satisfaction." Consumption may be a better word choice, at it delineates with a more modern term, that which *provides* greater want-satisfaction.

In this way, we may return to the value scale of preferred satisfactions that an individual uses when acting. Consumption may be better defined as the removal of a ranked preference from their scale. Also, when we define consumption in this manner, it becomes evident that our desires for consumption are limitless. As long as life continues, we find that our desires will be without end. By definition, in fact, we find that the continuation of life requires the continuation of unfulfilled desires. Breathing, for example, provides a continual desire that we have provided our life is to continue. It becomes clear, then, that the very nature of life implies a limitless end to unfulfilled desires or wants in need of satisfaction. Consumption is that process by which we ultimately satisfy existing wants and allows us to move forward to satisfaction of ever greater amounts of desires.

Time Preference and Consumption

As was already seen in Part I, chapter VI, there is a built-in mechanism, a distinctly human quality trait, which prohibits us from acting solely to enjoy continued want satisfaction in the present. Our preference for time (or time preference) implies that we are able to look to the future, and defer potential fulfillment of our wants until some future period of time. It was shown that this will always be a positive factor. That is to say, our time preference will always prefer to have a want satisfied in the present, to one in the future. This was demonstrated to be so as it allows us to maximize the total wants that we may eventually satisfy, owing to the limited nature (real and perceived) that our available time allows us to satisfy.

What our time preference does allow for, is not the the *absolute* deferral of consumption. Rather, it may slow the pace of satisfaction through consumption. It can be seen that the complete renunciation of all want satisfaction may arise only through a situation of death. Indeed, the want for survival through breathing may be seen as the last want which we may satisfy prior to leaving this world. The absolute denial of all satisfactions would occur in a state only available in death – a state defined as the negation of action. Such a state is the only conceivable *time* where where we could be said to exhibit no (or zero) time preference.

However, while the ultimate and absolute negation of time preference implies an impossibility for acting humans, we may definitely slow this rate of preference, and defer many satisfactions to a later date. In this way, we see that what time preference ultimately implies is not the negation of consumption, but rather its gradual deferral. Individuals characterized by high degrees of time preference will prefer a high rate of present want satisfaction over that in the future. Figure 1 demonstrates the trade-off that exists between time and our want satisfaction through time preference.

Time and Want Satisfaction

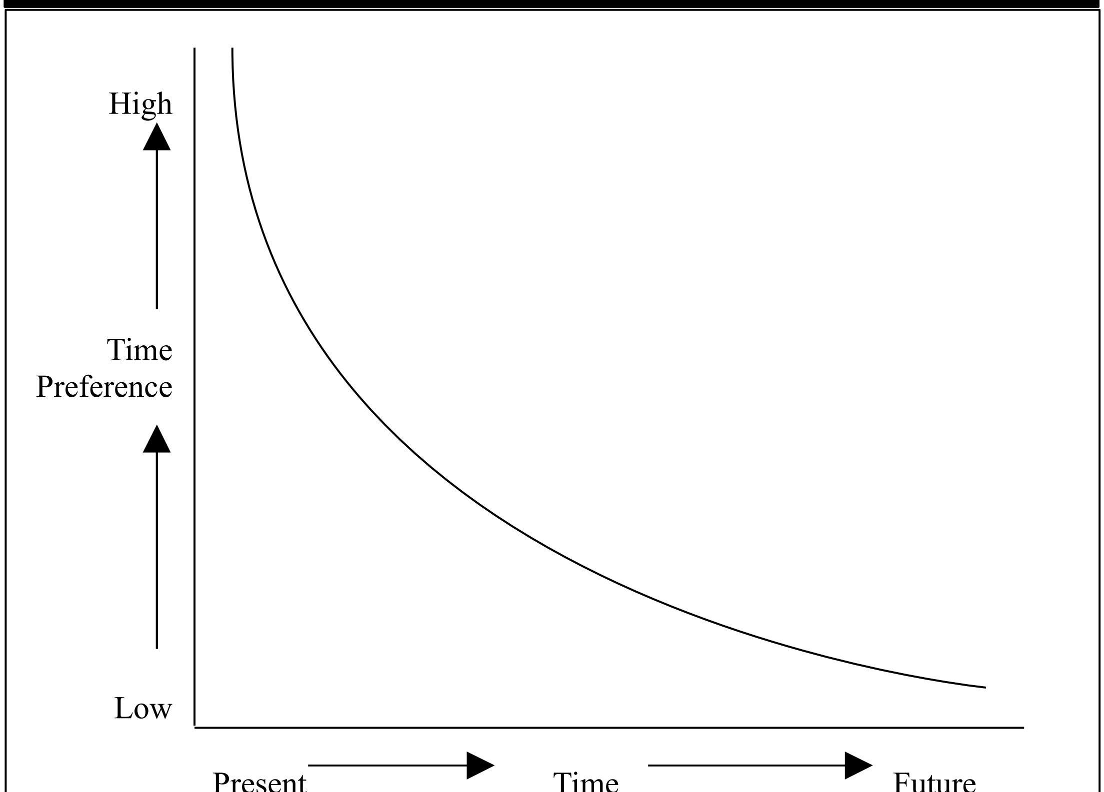

Figure 1

Two important conclusions can be drawn concerning our rate of consumption deferral through time preference. The first is that an individual may never be defined as having an infinite amount of time preference. High rates are observable, however, some degree of wants will forever have to be foregone in the present in exchange for their expected future fulfillment. This arises necessarily as our wants are limitless, but our time with which to achieve them is necessarily of limited nature. In fact, the important factors that define *real* time (as discussed in part I, chapter VI) bring light to this. Time is not a strictly uni-dimensional element. Instead, there is a definite *now* with which individuals chose to partake in actions directed at an expected future based upon knowledge learned in a known past. This gives rise to a multi-dimensional viewpoint of time, where we have an expected length of time with which to undertake actions and fulfill wants (as shown in figure 1), but it also gives rise to an extendable but limited present which places a restriction on the amount of wants that can be satisfied at any *one* moment. As this is a limited element, our wants can never *all* be fully fulfilled in the present, and hence, some degree must forever be deferred for the future. This gives rise to the impossibility of an infinitely high degree of time preference – although we may make note that some individuals exhibit *very* high degrees of time preference.

Second, we may note that time preference is never a negative factor, as shown in figure 1. We have noted already the necessity for continued consumption in the present, as a necessary condition of life. Now we may go one step further and comment that time preference not only may never be negative, it also may never be zero. Instead, as it approaches this lower bound of zero it asymptotically converges nearer and nearer this absolute boundary, without actually ever reaching this zero-point. The reason for this becomes evident when we think of what a zero degree of time preference would imply for action. Again, we see this as just the graphical expression for what we have previously explained as the reason for positive time preference. For not only can we never have negative time preference owing the continuation of life, we cannot even envision a zero degree of time preference owing to this necessary continuity and net-positive degree of consumption through want satisfaction.[258]

The concepts of time preference and consumption become inextricably linked. One cannot envision the act of consumption – that of want satisfaction – without thinking of a temporal moment when this action occurs. When we think about the general case for consumption, we can envision different rates at which this will manifest. Through the differential between high and low degrees of time preference we see the general tendency for the *rate* of want satisfaction to occur.

[258] Hülsmann (2002) makes the point that there can be observable categories of action that seem to neglect the concept of positive time preference – martyrs or warriors who seem to pay little heed or attention to conserving and prolonging their own lives. However, this misses the point that there is still exhibited positive time preference. As it is a definite preference that these actors partake in this type of action in the present, over deferring it for the future demonstrates their preference for present want satisfaction. That their expected future is so short only makes the example that much more pertinent, as it necessarily implies a very high *relative* degree of time preference.

It might become clear that if an individual prefers to not consume in the present, the ubiquity of action will require that some type of action occurs. It may also become clear that consumption can be increased or decreased, both in the present and the future, through the degree of alternative actions that we wish to undertake. In the next section further attention must be heeded to these action-alternatives.

2. Production

Production might be easily viewed as the corollary to consumption. In fact, more commonly, we may define production as that act which is expected to add value later and lead to greater amounts of consumption at a future date. It may become evident that the end of all production can only be a future state of increased want satisfaction. If we take this definition as adequate, we find that all production must occur in the present, forever aimed at some expected future state of consumption.

Temporal Viewpoints of Production and Consumption

Production and consumption may seem to be two sides of the same coin *prima facie*, but their seeming uniqueness has deeper rooted differences that manifest in many hidden places. One such place that deserves a closer look is the temporal element that each occurs in and is focused towards.

Consumption necessarily always occurs in the present. It makes no reference to the past state of consumption, but is forever tempered by an expected state of affairs in the future which may come into existence *if* an alternative degree of consumption is undertaken in the present. For, as the negation of present consumption allows for greater amounts of present production, the expected future state of affairs will be one which allows for greater levels of ultimate want satisfaction to occur. The concerned temporal viewpoints of consumption are those of the present state, which are known with near certainty, and the future, which are shrouded in great amounts of potential uncertainty.

Production, in contrast, is that which results from foregoing consumption in the present – it is a residual of sorts. Production too must occur in the present. However, it is also deeply concerned with that which has occurred in the past. The consumption of a car makes no reliance on the decision to consume a car (or more pertinently, the want-satisfaction of driving a car, transportation to work perhaps), yesterday. However, the decision to produce a car is deeply reliant on the decision to produce a car yesterday. The processes involved in production are more complex and lead to a necessary link between that production which has been established in the past and may be furthered in the present, and that which will have to be started anew. Not only does production necessarily have to focus on an inter-temporal viewpoint, it is also undertaken focused at an *inter-satisfaction* viewpoint. Production is not undertaken to increase production in the future. Rather, it is undertaken in the present with the expectation that future consumption will be allowed to be increased as a result.

Hence, we see the complexities arising in the production realm compared to that of consumption. Consumption necessarily is only concerned with the trade-off that exists between present and future levels of want satisfactions. Production, in contrast, must be concerned with past levels of production which may be utilized in part or kind, and expected future levels of consumption which will obtain should production be undertaken in the present.

Problems with the Theory of Production

Much confusion arises over what exactly it means to say "I have produced something." Take this quote as indicative of the confusion, as Barnett and Block (2007: 131 [footnotes omitted]) state:

> Production is action that is expected to add value; i.e., it is expected, in the ex ante sense, to in the future increase human want satisfaction, directly or indirectly. Alternatively, it may be thought of as action that transforms goods from a higher (earlier) to a lower (later) order. Goods can be categorized as consumers' goods or producers' goods (capital goods). Consumers' goods are used to directly satisfy wants; capital goods are used in the production of yet other goods, either consumers' or capital. Production, then, consists of producing consumers' services, tangible consumers' goods, capital services, or tangible capital goods.

They provide three definitions of production, which seem to not be fully coherent. As follows, we may sums them as:

[1] Production as a value adding process
[2] Production as a physical transformation process of creating more consumers' goods.
[3] Production as producing capital and consumers' goods.

A confusion arises in two regards. First is the more egregious error of confusing physical productivity with value productivity.[259] What production ultimately aims at is never a greater amount of physical goods in the future, but rather a greater amount of want satisfaction. That these two

[259] Böhm-Bawerk (1884) dedicated one volume to refuting the myth that interest stems from this very same error. That increased physical productivity may not always lead to greater amounts of future value should be clear not only concerning interest rate theory, but also in production theory in general.

occurrences may commonly be associated is only a *mere* historical fact, not an a priori deduction. For example, there is no questioning that an accumulation in capital goods has led to vehicles (cars, trucks, etc) that are more want-satisfying than existed 100 years ago. However, it is a mistake to reverse the causal process and deduce that the build-up of capital goods resulted in the increased want-satisfaction provided by the vehicles. Instead, the causal relationship is that an expected future level of want-satisfaction was foreseen (expected) and that the value placed on this enabled the entrepreneurs to accumulate the amount of capital necessary to make this expectation a reality. Further, the accumulation of capital could only occur under the lowering of time preference which necessitated the foregone consumption in the present, and hence, allowed for the accumulation of capital to occur.

The second error is the confusion which arises between ex ante expected value increases, and ex post realized value increases. Taking Barnett and Block's first definition, they seem to be implying that any activity that aims at increasing want satisfaction in the future may be considered as production. However, in their latter two definitions, the focus is on a much more ex post element, whereby the actual *creation* of capital and/or consumers' goods is considered to be the productive element. However, we may very well raise the question as to what would occur if the creation of capital goods was undertaken in the present aimed at some future state of increased consumer want satisfaction based upon their existence one year hence. However, could we really consider the mere production of these capital goods in the intervening period as value adding if upon completion consumers' wants have changed and they now prove to be useless for this want increasing purpose? It will be true that an increase in physical capital goods will have occurred, but without the alignment of consumers' wants to these newly produced goods, no such increase in value can be considered to stem from them. Furthermore, their creation will have resulted from some degree of consumption foregone in the present, which will have been enacted in the expectation that future want satisfaction will be increased. However, it can now be readily seen that no such increase in future want satisfaction will materialize.

Production cannot be defined as the production of capital or consumers' goods. It instead faces a limitation that these goods must be ultimately desired by individuals once produced. Production then can only be that which results in an ultimate increase in consumers' want satisfaction.

Temporal Considerations with Production

One problem which may now become evident is that, while the level of consumption may be measured at any given moment, the level of production is a much more difficult undertaking.[260] This arises as we know the level of consumption in existence at any temporal moment. However, production must necessarily await for its ultimate consumption before it may correctly be viewed as value-producing activity. It becomes clear, then, that at any given moment we may approximately know the value of consumption but the total value of *true* production must await for a future date. Total values of goods related to these actions may be determined at any given time, however, the question as to the value-added by the actions may only be determined ex post.

[260] For the moment we will ignore consumption such as that provided by leisure, which ultimately proves a much more difficult element to measure. However, we may at any given moment calculate the monetary value of consumption goods in existence, which provides an approximate measurement of the level of consumption. Intangible consumption, such as provided through leisure, will be assessed later.

3. Exchange

In opposition to the previous view of action as a dichotomy of undertakings – consumption and production – we see that this conclusion may only rest on a very shaky assumption. The denial of error in decision-making will support the dichotomy fully, and indeed, seems to be implicit assumption for supporters of this binary choice in acting.[261] However, as has been demonstrated in Part I, chapter IV, due to the uncertainty of the future, error forever remains an instrumental part of action. As such, we may not jump to the conclusion that all action that is marked by the foregoing of consumption may be considered as production. For such a conclusion rests on the idea that all actions that are not consumption oriented will necessarily lead to greater levels of consumption in the future.

What then may we coin those actions that belong to the class which we traditionally refer to as a type of production, but is really not of that type at all. One easy answer may be that we may retain the nomenclature and refer to these actions as *unsuccessful actions*, in opposition to their more successful counterparts which actually do result in production. But this solution would not solve our problem, as activities which we now wish to refer to as production may only be revealed as such ex post, after they have demonstrated to result in an increase in ultimate consumption or want satisfaction.

Instead, it seems more fitting that these ex ante productive activities instead be referred to as *mere* exchanges until the point where they can be established to be productive in the sense we wish to ascribe to the term. For in using the term exchange, we not only retain our distinction between production and consumption, but also are allowed to differentiate between successful and unsuccessful actions. Of course, Mises (1949: 97) has alluded to this already, but failed to bring the trichotomy to its full conclusion:

> Action is an attempt to substitute a more satisfactory state of affairs for a less satisfactory one. We call such a willfully induced alteration an exchange. A less desirable condition is bartered for a more desirable. What gratifies less is abandoned in order to attain something that pleases more.

What Mises calls exchange implies that the acts of consumption and production are derivatives thereof.

[261] In fact, as Hülsmann (2000) demonstrates, only in equilibrium may the possibility of error be eliminated. Hence, provided we operate within a disequilibrium setting, we need to acknowledge that error may occur, resulting in the trichotomy of actions as outlined herein.

In distinction, as has been shown above, this type of exchange may only be classified as a special instance of production – that which is undertaken in the present but unknown to be successful or not.

The Temporal Cases for Exchange

What may now become clear is that exchange has two temporal cases for occurrence. The ex ante versus ex post possibilities not only give rise to the necessity of this third category of action, but also give rise to two distinctly different types of exchanges which may occur, and may only be determined depending on their temporal occurrence.

In the first case – ex ante – we see that all which we wish to classify as *not*-consumption is exchange. Since we cannot at this time yet know if success will result or not (success as defined as greater levels of future want satisfaction), we cannot determine this in advance; although we may definitely have expectations as to whether these exchanges will be successful or not. Ex ante, then, it would seem that the case is made for only a binary distinction between consumption and exchange. Production must necessarily be excluded.

However, in the case that production may be realized ex post, and hence, from an historical perspective we may always identify that which was value adding production, that that which was want satisfying consumption, and that which was neither – only an exchange of goods from one location (temporal or physical) to another.

The overlapping nature of temporal actions means that the ex ante case for binary action – exchange and consumption – will never materialize. At any given point in *real* time we may see the fruits of previous exchanges materialize as consumptions. This, by definition, will imply that what we previously had to define as exchange was actually productive activity. Hence, at any given *point* in time we will be able to see not only those previous activities which are now revealed to be production, but also those being undertaken in the present which are still *uncertain* to be successful (those that may be defined as true production).

Entrepreneurship and Production

In part I, chapter V, we had looked at entrepreneurship as a duality of roles. In one instance, those maximizing individuals who face no uncertainty, but are quite apt at mitigating risk, can be seen as moving individuals towards a state of greater want satisfaction. This is so as there are clearly defined cases where no (or more correctly, little) uncertainty abounds concerning a particular action (i.e., demand is established, a product developed to suit this demand, etc) however, by lowering the cost necessary to be renunciated to attain a satisfaction, these risk-mitigators may increase consumer satisfaction.

Now it may become more clear that the case made previously between ex ante production as *always* being exchange due to the uncertainty surrounding its success is not as clear-cut as one may like. As one group of entrepreneurs may operate in a way with which little uncertainty abounds concerning the success of their current undertakings, we find that those activities which they undertake may be almost assuredly be defined as production- they will with little to no uncertainty result in greater levels of consumer satisfaction once completed.

This level of certainty regarding the risk-mitigating entrepreneur as producer increases as the temporal element their productive process requires is lengthened. For instance, the only true source of uncertainty that a risk-mitigating entrepreneur may face in the future are shifts in consumers' preferences for different products. It may become clear that time is an important source of uncertainty. A production process which takes one day may be said, then, to operate under very little day-to-day felt uncertainty as the limited temporal element leaves little room for uncertainty to manifest. However, if we were instead to look at a production process defined typically as solely undertaken by risk-mitigating entrepreneurs that takes one year, we see a greater degree of felt uncertainty operative as the greater intervening period between when the production process is undertaken and when it will be completed leaves more room for ultimate uncertainty concerning shifting consumer demands and preferences.

Uncertainty bearing entrepreneurs, by their very definition, are undertaking action whose result is uncertain. This evidently enough cannot be considered production in an ex ante sense. Work undertaken by this class of entrepreneur will forever be defined as exchange until that point when it may be ascertained that it is productive in nature.

Hence, we see that when we view production as an entrepreneurial element, a distinction arises as some entrepreneurial activities will actually be identifiable in the present as productive, while others will forever need to wait until the fruits of their labor result in that which we may consider productive activity. Not all action which is considered as not-consumption must be labeled as exchange prior to its known success. Instead, we see that there are elements of action, those undertaken solely under conditions of risk, which may be identified ex ante as productive activity.

The Temporal Viewpoint of Action

All action is undertaken aimed towards a future state. It bears comment in light of this that although we will only know if an undertaken action proves to be successful as production ex post, each entrepreneur performing this action will believe that they will be correct, and hence, that their action represents true production. We may also be in a position to state that not everyone can be correct in their projections concerning the future. Some will err as their plans are disrupted by others. Educated guesses may be made as to what will constitute production and what will forever remain in the realm of exchange, but these will by necessity be only guesses, for only the certainty of ex post realization will reveal which plans are correct and which are not.

All entrepreneurs will act thinking that they are partaking in true production, however, only a portion of these actions will actually result in productive activity. To the degree that entrepreneurs are correct in their expectations, we may expect that ex ante exchange will result in ex post demonstrated production. Entrepreneurial error will result in reduced production, and an offsetting increase in exchange activity which results in no net increase in ultimate consumer want satisfaction.

III. GOODS AND THEIR FORMS

The trichotomy of action – exchange, production, and consumption – leads naturally to a definite trichotomy of goods following the same classification. Menger (1976: 52) set out to distinguish between those objectively useful things which we may not object to admitting existence of, and those which are actually useful and within our realm of use that we may define as goods. In fact, as Menger clearly delineates, there are four prerequisites that a thing must fulfill in order to be considered a good:

[1] A human need must exist.
[2] The "thing" in question must possess sufficient causal connection to bring about the satisfaction of this need.
[3] The individual must recognize, that is to say, have knowledge of, the causal connection that exists in number [2].
[4] The individual must have the "thing" in their own command to exploit the causal connection.

The initial absence of any of these four prerequisites denies a "thing" from attaining a proper existence as a good. Likewise, a good may exist, but if one of these four points should no longer apply, its status as a good proper will be lost.

Both Menger (1871) and Böhm-Bawerk (1889) had set out to refute the previous concepts of "classes of wants" in favor of "classes of goods." The impetus for doing so was to eliminate and refute the erroneous concepts that existed concerning questions such as why water and diamonds are valued differently. As Mises (1949: 123-124) shows us, concerning action it matters not what class of wants is being satisfied, indeed:

> Classes are not in the world. It is our mind that classifies the phenomena in order to organize our knowledge. The question of whether a certain mode of classifying phenomena is conducive to this end or not is different from the question of whether it is logical permissible or not.

As we saw in chapter I, there is in fact only one type of "want" - that which concerns the wants that a consumer wishes satisfied in order to reach a state of greater satisfaction. To speak of classifying wants would be akin to merely stating what the preference rank of wants an individual had was at any given moment. We see the futility that results from delineating such a distinction as these will exist only in a relative state within an individual's mind, and never form any sort of absolute categories from which to build a theory of action from. We may, however, categorize *goods* according to the role that they serve in the action process. As our action process necessarily has three definable categories, our classification of goods may also take on this same form.

1. Consumers' Goods

Those goods which directly serve in the satisfaction of consumers' wants may be considered as consumers' goods. This definition may prove to be unassuming, but it requires three elaborations to provide the true meaning which we wish to convey with the term. The temporal element of the serviceableness of want satisfaction through a good gives rise to a necessary consideration as to the expected duration of serviceableness embodied in the good. Second, these goods need not always be considered consumers' goods. In fact, in many instances we may find that those which we consider consumers' goods hold a dual role as capital goods as well. Last, we will see that not all things we wish to refer to as consumers' goods are actually goods. In fact, a large portion of these will be services lacking a physical component – services. Each of these points will be discussed in further detail.

Duration of Satisfaction

Consumers' goods need not yield their want satisfying services at only one instant. We find that many consumers' goods will have a life-span of sorts, with which they will continue yielding services to satisfy our ultimate wants. Two categories of consumers' goods become apparent – those which provide satisfaction only for an instant, and those which are renewable to continue for an extended period of time.

Those consumers' goods which only provide satisfaction in passing may be referred to as non-durable consumers' goods. For example, food provides the satisfaction of hunger removal but only for an instant. Once a food is eaten, it will have lost all power at being used again at some future date. The duration of serviceableness of this good is, for all intents and purposes, instant – there is no lagging period of services provided.

The class of consumers' goods which yields satisfaction over a temporal dimension we may refer to as durable consumers' goods. Houses, for example, if used for purposes of shelter provide this service over an extended period of time. For while a house may be purchased in the present, it will essentially yield sheltering services for an extended period of time. Durable consumers' goods may require an allowance for depreciation in order that they continue providing want satisfying services. Houses, for example, will require continued upkeep – new roofs, repairs on their property, etc. Taxes, in this sense, also provide a component which must be factored as a cost for continued service from a durable consumers' good. For what good would a house be as a provider of shelter if the owner's failure to pay taxes resulted in their eviction by the city? As Garrison (2001: 48) points out, consumers' goods are durable for two reasons. First, as these goods wear out, it is evident that new goods provide better service than old ones. Second, as in production goods, a time discount is applied to consumers' goods as well. Satisfaction in the future is valued less than in the present, hence, these durable consumers' goods that will have increasingly less emphasis placed on their future satisfying abilities than those in the present.

Earlier, we looked at time preference as resulting from consumption in the present deferred to a future date. Time preference was shown to be *the* cause of originary interest (see part I, chapter VI). Time preference also finds itself manifested through the degree of durability of consumers' goods. We may notice that there is a general, although not constant, trade-off between the degree of serviceableness of consumers' goods and the cost of acquiring those goods. For example, houses of short serviceableness may be made at lower cost compared to those with a higher duration of serviceableness. In particular, we may notice that a more costly steel roof may be used that will last (or provide want satisfaction) for 50 years, as opposed to a less costly asphalt roof which will provide services for only 15 years before requiring replacement or repair. It becomes evident that if an individual has a lower time preference, they will be inclined to invest in the longer yielding consumer good. By sacrificing more present satisfaction by using a more costly option in the present, they will enjoy the benefits of the good over a longer time horizon, or duration.[262]

Second, not all consumers' good need be considered as solely consumers' goods at all times. A truck, for example, may provide pleasure as a driving vehicle on the weekends, but be used for work purposes during the laboring days. A certain amount of consumer-capital switching will become evident. A house may provide services as a durable consumers' good when used as shelter. However, when it is used as a rental unit, its role now becomes two fold. First, it continues to be a consumers' good for the renters as they receive shelter from it. However, for the true owner of the house, it now becomes a capital good which serves to increase their future want satisfaction through the stream of income it generates pertaining to the future.

Last, a brief note should be made on services, which provide the present-oriented consumers' good *par excellence*. These are only able to be enjoyed in the present, although their effects may continue for an extended duration. For example, haircuts are typically cited as examples of services

[262] This, of course, in general ignores the entrepreneurial motive of searching for ways to provide consumers with satisfaction while having to renunciate less in return.

provided. While these are services in the sense that they require few capital goods to achieve, the effect of a haircut has an extended durable life which requires continual reinvestment in order to compensate for deterioration in serviceableness over time (i.e., use). Hence, while my haircut generally makes me look presentable for a few months, eventually my hair will reach a point where a new trim is required. The satisfaction of having a presentable haircut requires a continual reinvestment in order to maintain this look. In contrast, that class of services may exist which requires no such continual reinvestment. Advice, regarding taxes for example, provides a satisfaction for required knowledge which will not need to be reinvested in in order to maintain its benefits in the future.

Not only do consumers' goods have a three-fold nature – non-durable, durable, and services – but they also have the property that they are not always considered the same good. Some degree of switching may be possible between consumers' and capital goods. However, the nature of this switching does not imply that a good can have a dual-classification at the same time. Take an example of a house which is owned to be used as a rental (a capital good) and is rented by a person who views it as a durable consumers' good. When we return to Menger's original four requirements for a "thing" to take on the character of a good, we see that the fourth – that the individual must have the "thing" in their own command to exploit the causal connection – may pertain only to the owner of the good. In effect, the question of ownership becomes central to the question of what role, or category, a good takes on. Hence, in our example of a house having a dual role as a capital and durable consumers' good, we find that only that role which it serves for its owner – as a capital good – may be counted towards its true goods' character.

Leisure and Time

Leisure provides an interesting case of a service which is provided by the same person who uses it. These self-made and administered leisure activities have the same inherent problem as the switching example cited above where a consumers' good may take on multiple roles at once. Leisure is generally seen as the "amount of time not spent in labor" (Rothbard 1962: 46n30). Leisure as a good becomes problematic when we try to assign a value to it. Using this Rothbardian definition of leisure as time spent not working, there are several interesting points to be gleaned. Earlier in chapter II we saw the trichotomy that existed in actions. The problem which existed when viewing actions is that when it comes time to assign a value to them, difficulties arise. Particularly, we see problems with consumption as action, as value becomes only implicitly given through counter-factual activities.

First is that, in reality, the distinction is not between working and not working, or leisure, to keep terminology consistent. Instead, the distinction is between producing, exchanging and consuming. It should become clear that what is commonly referred to as leisure is a type of consumption. In fact, leisure represents that time which we have at our disposal for activities which are typically defined as consumption activities. As such, we must first make a distinction between time spent on the three types of activities. Value through our time spent in production is easily valued through the explicit wages which are provided to us for this activity. Leisure constitutes not a good as such, however, it does represent a value through our time available for consumption. We may value our time for leisure as the foregone alternative that we could attain for this time. Leisure provided through our weekends spent enjoying leisure may be valued by the foregone alternative such as the wages we could earn by working a weekend job.

This value through leisure time becomes apparent when separated from the value of consumption goods. For instance, although a haircut may represent a consumption good, it is erroneous to consider the time spent having a haircut as a value adding leisure time. Instead, this temporal period represents a time spent in exchange – in this case, exchanging time plus money for a haircut.This temporal aspect represents a cost which must be considered in looking at the total value renunciated for the haircut. However, the cost is implicitly paid, as it is merely removed from the total amount of time we have available for our consumption.

This residual consumption time should not be considered a good as Rothbard (1962: 1323) does. When we look back to Menger's four criteria for a "thing" to take on a goods' character, we see that the fourth – the command over that good – becomes a suspect criteria concerning time. Time, after all, represents an element which we can never absolutely control. Instead, we can only make a choice between using something sooner or later. Our time spent enjoying leisure does, however, become a good when we view the trade-off we make between fixed in the present and foregoing this choice to instead partake in leisure activity.

We now see that leisure time does in fact represent a good which we must group in with the other consumers' goods previously assessed – durable, non-durable and services. While the value of these three consumers' goods may be ascertained through their respective exchange values, the value of leisure must be ascertained in a more round-about method – the counter-factual of the next-valued activity that could be partaken during this time period.

2. Capital Goods

Until this point, we have typically referred to producers' goods as those which contrast with consumers' goods. Although logically pleasing to use this terminology, more modern convention may make it more relevant to begin using the more common term – capital goods. This should not detract from the point made earlier that capital goods are used for the specific purpose of increasing consumer want satisfaction. In fact, the use of the term capital good allows us greater flexibility in usage. The distinction between ex ante expected success in the production of consumers' goods and the possibility of ex post failure to produce an increased amount of valued consumers' goods results may now be expressed using the unique term – capital good.

Capital goods may be viewed as those which are used to create the final consumption goods which directly provide satisfaction. They are the goods thought (or perceived) to be necessary in order to obtain consumers' goods in the future (Huerta de Soto 2006: 272). However, we see that this perception involves a mental context, whereby the individual's aims give rise to a subjective perceptive of *what*, in fact, aids them in pursuing this end. As Kirzner (1966: 38) notes:

> The principal point to be emphasized is that capital goods, thus defined, are distinguished in that they fall neatly into place in a *teleological* framework. They are the interim goals aimed at in earlier plans; they are the means toward the attainment of still further ends envisaged by the earlier plans. It is here maintained that the perception of this aspect of tangible things now available provides the key to unravelling the problems generally attempted to be elucidated by capital theory.

These capital goods are the accumulation of three factors – natural resources, labor, and time. In fact, although the first two are commonly acknowledged as being necessary for the creation of capital, time is an oft neglected factor (Yeager 1979). We may see the importance of the temporal element when we look at two important aspects of capital.

First is its formation. As Böhm-Bawerk (1889: 102-118) demonstrated, an act of saving must arise *prior* to the creation of a capital good. Specifically, using his Robinson Crusoe example, it was seen that only by foregoing his consumption in the present, was time made available to work on the creation of capital goods – in Crusoe's case, a stick to knock berries from a bush. We saw this previously as time necessarily implies an action, and action has three logical categories – consumption, exchange and production. Consumption is the natural state of action – we try to reach for an existence with ever more dedication to those we refer to as consumption activities. Due to our time preference, we are able to forgo present consumption, and dedicate time to the attainment of capital goods (i.e., through production).

Second is its maintenance over time. Hayek (1941: 54-58) defined capital as that stock of resources which are not permanent. Not only must some amount of consumption be forgone prior to the formation of capital, but due to its non-permanent nature, some amount of saving must be continual in order to offset the depreciation of used capital (see Strigl 2000 on this point). Hayek's stress on the non-permanent means of production may seem troublesome. It becomes evident that there are many resources which we would typically define as permanent (i.e., land) which directly serve and are necessary in the production process to create consumers' goods. Aiming to justify this definition, he (1941: 59 [footnote omitted]) states:

> … the non-permanent resources provide an income stream for a limited period; and that in consequence we are in a position to postpone the return from some of the current services of the permanent resources without reducing our consumption below the level at which it can be permanently kept. We are thus able to take advantage of the celebrated productivity of the 'round-about methods' of production which have been the cause of so much understanding... [T]he existence of non-permanent resources makes it at the same time possible and necessary to 'invest' some of the current productive services, that is to use them in such a way that they will not yield consumable services until a later date than they might otherwise have done, but will then yield a larger amount of such services than they would have done at the earlier date. It is only because of this that the provision of an additional amount of services for a limited period in the future puts us in a position to raise for all time the return which we may hope to obtain from the meagre [*sic*] supply of really permanent resources.

Hayek's mistake in so viewing a difference between permanent and non-permanent resources was in his confusion between physical and value productivity. For while it may be true that some resources retain their physical productivity over extended periods of time, value productivity is a completely different story.[263,264] Menger's imputation theory of value was so important as it refuted the existing objective theories of value. Value is never derived from some innate, absolute and time-invariant quality – such as would have to be supposed by assigning a class of resources the label 'permanent.' Instead, value is a shifting flux that varies as consumers' needs change. Hence, even land, that resource par excellence that we wish assign to the category of permanent capital cannot be so viewed. For what permanence would a diamond mine have if tomorrow the world's demand for diamonds was eliminated?

If we were to then view capital goods as those goods which combine to create the consumption goods demanded for want satisfaction – whether permanent or non-permanent – we may do well to revisit Strigl's (2000: 27) trichotomy of capital goods:

[1] Free Capital – the subsistence fund (or supply of consumers' goods) which allows for more roundabout methods of production.
[2] Intermediate products – raw materials at each temporal point of production prior to a consumer good being produced.
[3] Fixed capital – those that are "relatively durable" and are the produced factors of production.

Strigl makes note that consumers' goods are not capital, except when they are used in this way. Hence, consumers' goods that are used as savings so that more roundabout production processes may be undertaken serve as an important type of capital (see also Mises 1949: 260). Consumption goods must be accumulated prior to the creation of capital, and this in turn implies that some consumption goods assist in the process of creating further consumption goods in the future.

The end use is what will be considered when we assign a goods' character. To the extent that a good exists to fulfill a certain want, it may be considered a good of that character. Consumption goods

263 Viewing those resources as permanent those which last essentially forever may be conceptually easy, but is too extreme a definition to have meaning. For what is of importance to the individual is if something will remain permanent over *their* useful lifetime. Hence, we can see that it is not necessary to view land as a permanent resource for thousands of years, but rather as permanent throughout the lifetime of the individual wishing to use it, and their subjective and personally perceived time duration that they wish to use this resource over. Remember that this is not merely limited by the physical life of an individual, but may also be viewed to include dependents, heirs, or the like. In fact, we can see by right of a thriving insurance industry that individuals are not concerned merely with their own lifespans, but that of others' as well. Likewise, as Lachmann (1956:11) indicates, for the theory of capital," the question which matters is not which resources are man-made but which are man-used. Historical origin is no concern of ours."

264 Hayek's insistence on capital as the non-permanent resources available to be used in production may have been an unfortunate result of his debates with Knight over capital. With Knight's (1934: 264) view that capital was a permanent fund whose "replacement has to be taken for granted as a technological detail," may have caused Hayek to define capital as those resources which, by definition, would require replacement at some point in time. See, for instance, Hayek (1936) for an early elaboration on his own view of capital as non-permanent resource.

used for the increase in consumer want satisfaction come under the rubric of consumers' goods. However, there are also consumers' goods which may not be used to satisfy our wants, but also to further the production process to increase our future wants. We have already noted that goods' may switch between consumers' and producers' types, but have failed to appreciate the complexity of the situation concerning the use of consumers' goods for consumption ends in the production process.

An example may prove helpful. Assume that an individual has but one want – to drive a car for pleasure – but at the same time has 2 different cars. That only one may be used at once should become clear. Are we to consider both of these cars as consumers' goods? The short answer is: no. For only one may be considered a consumers' good, that is, a good satisfying a consumption want, *at any given time*. The extra car we may consider as a Striglian subsistence fund – a type of capital. The accumulation of consumers' goods in excess of that which is required for the fulfillment of our consumption wants may be used to offset future want satisfaction. This allows a lesser amount of consumers' goods to be produced in the present, in exchange for capital goods which are expected to yield greater want satisfaction at a future date. This allowance, then, allows the production process to focus on capital goods – in effect, it permits the creation of capital.

Heterogeneity

One of the hallmarks of Austrian capital theory is its emphasis on the heterogeneity of capital goods. This is obvious when we look at physical capital goods, but becomes manifest when we turn to the implications thereof. This heterogeneity of capital *goods* implies that no type of aggregation may be possible concerning these goods. Just as economists are well-aware of the difficulties of adding apples and oranges, so to do insurmountable difficulties abound when aggregating paint and steel to produce an auto body.

It is, in fact, an outcome of this heterogeneity of capital that we must not refer to a stock of capital goods. Heterogeneity of capital goods stems directly from the heterogeneity of our consumer wants and desires. Just as there are a multitude of desires that we wish fulfilled, there are many different ways with which consumers' goods may be produced to satisfy these wants. This process regresses back to the production goods as we see there are many (perhaps almost infinite) ways that a good may be produced to satisfy a consumers' want. Skousen (1990) gives the example of the complex production processes, and hence, capital goods, that are necessary for the production of a car today. If we look back over the past 100 years, we may see that there were perhaps millions of different capital goods that were used to create this good to satisfy the desire of transportation – the automobile. However, if we think of the various methods of transportation that are not even imaginable today, we may well note that capital goods of the future have a nearly unlimited existence – a heterogeneity will prevail in excess of that which we may even see with our historical perspective of today.

Heterogeneity of capital goods implies two important factors which must be commented on concerning the role these goods serve in the production process – substitutability and complementarity.

Substitutability

Lachmann (1947: 199; 1956: 56) would use his example of a delivery company to demonstrate the substitution effect of capital goods. A a delivery company may have many delivery vans at any given time, as long as planned use of each van is followed, no problems will arise. However, should one of the vans break-down, the delivery pattern will be shifted to account for this. This alteration will be possible to the extent that the vans are substitutes for one another, so that they may be shifted around within the production plan to achieve the same (or nearly so) result.

As Lewin (1999: 123) brings to light, substitution of capital goods will only prevail to the extent that a "certain set of contingency events can be visualized." Goods may only be substituted within the known uses that they may have. In the example of delivery vans, it is easy to see that these vans may substitute for one another; although not being identical in nature, they may be considered to have a high degree of homogeneity. However, what of goods that have considerably less likeness in form, and function? A restaurant may run out of steaks to serve for dinner, and have only different types of meat to be used as its main course at dinner service. Knowledge that chicken will be an acceptable substitute for beef will be an instrumental factor in determining if these goods are, in actuality, to be considered substitutes. We see that ultimate consumer demand will dictate the degree to which this is possible or not.

Complementarity

The complementariness of capital goods implies that some goods will be required in order to realize plan completion. Lachmann (1956: 54) defines two types of capital complementarity. First is "plan complementarity" which implies how goods assist each other within the greater co-ordination of a planned action. This is effected directly by entrepreneurial actions, through the making and revising of plans. The second type is "structural complementarity" which views capital goods and their inter-relations within the greater economic system. Hence, this type is brought about only indirectly, through the market process.

Capital

We may assign capital a distinctly different meaning than that which we have assigned capital goods. As Lachmann (1956: xv) delineates the concept of capital:

> The generic concept of capital without which economists cannot do their work has no measurable counterpart among material objects; it reflects the entrepreneurial appraisal of such objects. Beer barrels and blast furnaces, harbour installations and hotel-room furniture are capital not by virtue of their physical properties but by virtue of their economic functions. Something is capital because the market, the consensus of entrepreneurial minds, regards it as capable of yielding an income.

Capital may be best defined, then, as the market value of the capital goods in existence. The difference between this valuational concept, and the physical concept of capital becomes apparent as Huerta de Soto (2006: 282) points out that "'capital' [*is*] *the market value of capital goods*, a value estimated by individual actors who buy and sell capital in a free market." The necessity of independent capital markets to establish these capital values implies that there are many instances where this distinction becomes quite necessary. For instance, in the Soviet economy of most of the 20th century, lacking capital goods' markets to determine value, we can say without a doubt that many capital goods were in existence, but there was a total lack of *capital* (see Huerta de Soto 1992; 2008: 50). In fact, we see that by using market values, we are able to answer the Lachmannian (1956: 53) criticism of viewing capital in aggregate terms as there is no common denominator with which to measure it. Market values, as established through the use of money as a medium of exchange (and more pertinently, unit of account), allow just such a common denominator to obtain.

The concept of capital as *valued* capital goods implies that we must know that the goods in question will be valuable to the end they seek. As this end is the ultimate satisfaction of a consumption want or desire through the creation of a consumers' good, we may seek a definition of capital that is more concrete and which allows us to differentiate between capital which has value (i.e., it has been attained through exchange), which ex ante is not known to serve the consumers' end that it has been employed for. Reisman (1990: 132) provides one such definition:

> The aggregate of capital goods in the possession of an individual can be described as his *capital*. And capital can be defined as *wealth reproductively employed* – that is, as wealth employed in the production of wealth. (In the contest of a division-of-labor economy, capital is wealth employed in the earning of money.)

There are some problems with Reisman's definition, which deserve brief mention, before we note the truth of the matter as he has stated it. First, we have seen the subjective problem which arises from aggregating capital goods into capital. The solution provided by Huerta de Soto of using market-values overcomes the obstacle of capital heterogeneity by providing the common denominator of money for value establishment. It is wealth employed in the production of wealth that we wish to stress from this definition, as it seems to succinctly get to the crux of the matter. We may never know if a capital good will actually serve its role in producing greater levels of consumer want satisfaction until ex post, but we can see the expectations that individuals have regarding this occurrence through the yield on these goods at any intermediate stage.

The value that capital has at these intervening stages of production, prior to the completion of a consumers' good which is known to have value, may be ascertained through the expectations that individuals have in the present. Hence, the profit yield on capital goods may give us an ex ante expectation of what the final value will be. Capital goods not expected to add value to a final increase in consumer want satisfaction will not have a positive yield in the present, although this will forever be subject to change until the point where a capital good is used for the direct purposes of consumer want satisfaction.

Conclusion

This heterogeneity of capital results in the existence of a structure of capital goods. As these goods are linked together through their uses – their substitutability and complementarity – they provide an inter-locking system of production. As entrepreneurs work towards having a coherent plan through complementary assets at their personal plan level, the market as a whole is fixed towards a situation of greater amounts of plan complementarity (Lewin 1999: 124). The plans in question are always those expected to prevail in the future by consumers' demands.

The substitutability of capital becomes an important component when error occurs in plan co-ordination. For if assets are highly substitutable, there will be little cost associated with switching them from one plan to another. As Huerta de Soto (2006: 282) notes, although capital goods can be very difficult to convert to different forms, they can be given added mobility through our created institutions designed to serve this purpose – contract law for example allows ease of mobility with capital, even if this capital may be difficult and require a timely process to convert in form.

We may also see that capital goods become more difficult to convert, that is to become substituted for, the closer they are to the point of final consumption. Entrepreneurial errors become more pronounced then at these lower stages of production as goods created with one type of consumer want satisfaction in mind may become useless and unable to be converted to a different use easily. For example, raw iron ore in a higher stage of production is more easily substituted if its projected use has changed, than a completed automobile will be if the same event arises.

We find that these two concepts – substitution and complementarity – also create a complex linkage between plans both horizontally, between plans of disparate individuals of the economy, and vertically, between plans of one individual.[265] These dimensions of capital goods will be expanded on in further chapters.

[265] Wicksell (1951, I: 164) was perhaps the first to use this terminology, referring to a vertical and horizontal dimension in the structure of capital goods.

3. Exchange Goods

The final category of goods which we wish to look at, and will exhaust our praxeologically derived trichotomy of goods' types, is that of exchange goods. Previously, we have seen that this category arises for two reasons.

First is that category of goods which is forever used only as an intermediary step in the process which is necessary to bring goods one step closer to satisfying consumers' wants. We typically define these goods as media of exchange – money. The *result* of this exchange good is the unit of account by which we are able to establish the *nominal* prices which we wish to measure our capital and consumers' goods' values in. As Horwitz (2000: 52) points out, money serves not only at a disadvantage *vis-a-vis* capital goods, but that this also justifies its existence as a separate category. For while capital goods can be employed directly for production processes, money must always go through one step of exchange before production may occur, that is to say, it can never be directly employed in the production process. Money's general usefulness is one of its main advantages in exchange, but becomes a hindrance in light of its usefulness in production. Additionally, as Mises (1932: 61) demonstrates, capital goods only serve their purpose while in use, but money always serves a purpose – whether it be sitting in a till, hoarded under your mattress, or used in active exchange.[266]

The second case we may look at are those decisions which were undertaken to combine capital goods in the expectation that a valued consumer good will be produced in the future. The possibility of error makes it clear that not all investment decisions in the present will yield valued consumers' goods in the future. Indeed, as Myrdal (1939: 23) first reminded us:

> Similarly, one cannot assume that capital (investment) demand and capital (saving) supply are identically equal; for they, too, originate with non-identical groups of individuals. To treat supply and demand in these cases as being *identically*, rather than *conditionally* equal, would involve a highly unreal and abstract concept of equilibrium.[267]

[266] We may also heed Niehans (1978: 14) on the matter: “The problem was to explain precisely why money stocks are useful. It is clear that, except perhaps for irrational misers, cash balances are not one of the genuine consumer goods appearing in consumer theory... It is also clear that money is not one of the genuine producer goods appearing in an ordinary production function... Rather than from direct utility and production, the services of money arise from exchange, being derived from the utility of money spent.”

[267] Although typically credited with first delivering the emphasis on ex ante expectations versus ex post results, Myrdal was presaged by Rosenstein-Rodan by almost three years with this concept. See (1936: 274).

Previously, we looked at the case that arises whereby all productive activities that face *uncertainty* (but not necessarily risk) would have to be viewed as merely exchanges until that point where it could be demonstrated through consumption that these goods do result in an increase in consumer want satisfaction. The possibility of error precludes us from assigning the term of capital good to all those goods which we wish to use in the production process. However, generally we can see that capital goods with a positive present value will have this as the expectations of the individuals in the present is such that these goods are deemed to positively affect consumer want satisfaction in the future. This conclusion which we ended the previous section – that present positive yields on capital are an indicator of future success in want satisfaction – is an independent, necessarily subjective interpretation. For it could very well be, and often times does turn out, that these capital goods ultimately fail in their present form to satisfy consumers' wants.

These two types of exchanged goods – money, and erroneous capital investments – will require further attention to complete our look at goods' natures.

The Origin of Media of Exchange

One of Menger's greatest contributions to the science of economics is found in his writings on the origin of money. He, concurrent to Jevons' (1876) analysis, viewed money arising from the "double coincidence of wants" problem.[268] Two individuals wanting to engage in trade each hold goods that the other may not want. This single issue gives rise to generally accepted commodities that came to be used not only for their direct use-value, but their exchange-value in trade as well. Hence, for Menger, the emergence of money stems from spontaneously finding a commonly accepted medium of exchange. The end goal would be to find a medium that would reduce the bid-ask spread on goods to the maximum allowable amount. For instance, it may have been that any good was able to be traded to some extent. Approaching a butcher to buy a steak with only a hammer at your disposal for exchange may still result in a trade; the question would be at what price. The discount offered for your hammer may be so great so as to dissuade the trade at all, or place you at a considerable bargaining disadvantage.

[268] Say (1803) would also allude to the double coincidence of wants problem. In fact, Say was one of the primary influences on Menger's thoughts on money. See Rothbard (1995: 37), for a further elaboration on this prehistory. Also, see Wu (1939: 126) for the view that Mises was heavily influenced by both Ricardo and Nassau Senior, as well as Menger, while formulating his own theory of money. For a more modern elaboration of Mengerian monetary economics, see Howden (forthcoming).

While recognizing that money emerges as a spontaneous order (1883: 131), Menger would also note that the precious metals have taken on the role of money, but this is only an historical fact.[269] Hence, diametrically opposed to conceptions of money as a created institution, for Menger (1892) we find that, "[m]oney has not been generated by law. In its origin it is a social, not a state institution."[270] Precious metals were adopted due to their particular qualities; however, these have always been particular historical facts, never time-invariant axioms we could use to deduce where future money is to be produced from.

Mises (1953) would add considerably to Menger's insight with his "regression theorem." The adoption of money requires two specific individual components: use-value and exchange-value. For general commodities, it makes no difference if they contain exchange-value. However, as Mises (*ibid*: 97) pointed out, for money to have a use-value, it must have exchange-value. To develop this concept one step further, there is no subjective use-value in money, unless it contains an objective exchange-value (*ibid*: 98).

Mises' regression theorem would explain why some monies that seem to have no use-value have come into existence. This is explained as their demand would be sourced from their demand the prior day, and onward, in an infinite regression to the past. At some definite point in the past this regress must end, and at this point, money's exchange-value would be derived from the use-value available at that given moment. The result is that a given money might have no direct use-value in the present, however, for it to have demand in the present, it must have derived its exchange-value from some concrete use-value in the past. The direct implication of this is that no money can originate from nothing, or more appropriately, no money can emerge from something which has no use-value. The consequences for theories built on a concept of money with a created existence at its origin of money are evident.[271]

[269] See also Rothbard (1962: 192) for the actual physical medium of exchange being a historical realm, and outside the realm of economic theory as such. We can compare Menger's spontaneous emergence view of money in distinction to his contemporary Walras where he stated, "[l]a monnaie est une affaire d'Etat" (1898: 169). For a closer look at the differences between Menger and Walras' conceptions of money and monetary evolution see Arena and Gloria-Palermo (2008: 333).

[270] Interestingly, a later Menger (1909) viewed the state as being able to "perfect" money. Once money had previously emerged as a market institution, the state, primarily through legal tender laws, could improve upon its acceptance and increase its demand as a medium of exchange. This, however, relied on a pre-existing, market established money.

[271] Hayek (1976) would advocate a system of competing currencies operating alongside the current government fiat money. In his view, the competing currencies could also be fiat in nature. For a critique of this view point, from a Misesian perspective, see Herbener (2002: 6), who argues that all a state can achieve in this is to ratify an existing medium of exchange. This would refute earlier criticisms of the Misesian theory of money by Gilbert (1953: 149) and Patinkin (1956: 71) who argued that fiat money introduced after a monetary collapse could not be explained by Mises' regression theorem. However, we can see that fiat money can never be introduced *ex novo*, but must always be offered for exchange with an existing currency (Rothbard 1976). See also Tullock (1957) for a discussion of the historical failure of

Later, Mises (1949: 405) clarified his position on exchange-value and money:

> The purchasing power which we explain by referring to the extent of specific demand is not the same purchasing power the height of which determines this specific demand. The problem is to conceive the determination of the purchasing power of the immediate future, of the impending moment. For the solution of this problem we refer to the purchasing power of the immediate past, of the moment just passed. *There are two distinct magnitudes*.

The relationship between demand and purchasing power is thus a complex one. The purchasing power of the immediate past that Mises writes about is conditioned by two things. First is the relationship that exists between goods available at that point in time for purchase and the amount of money outstanding with which to purchase them. Second is where his regression theorem plays a pivotal role. The demand for money at that particular moment existed due to the finite regression of demands the money had previously experienced. Hence, we can say that demand for a given money was determined in the past by its previous demand and purchasing power thereof.

However, the demand and purchasing power of the "impending moment" is of a slightly different nature. Demand for the future is conditioned by the demand that exists due to the past purchasing power of the money. The purchasing power of the future will be tempered somewhat by the expected purchasing power it will contain at the moment it is expected to be used. Mises' distinction between the two "magnitudes" takes on great significance when viewed in light of this.

Hence, we can see the process through which money not only emerges, but also how it maintains value *even in the absence of a valued exchangeable commodity backing it*. While the theoretical questions as to what money *is* have been answered adequately in the general sense, there is still the question as to what we should consider money in the more concrete sense, so as to ascertain where value changes in the stock of exchange goods lies.

Elements of Monetary Exchange Goods

Rothbard (1963: 87-91) set out to concretely define the money supply, in light of the monetary

a spontaneously introduced paper money in ancient Persian civilization. He concludes the failure was due to the fact that the Persians had no prior use-value in fiat money, hence, they held no reason to believe it would have an exchange-value.

substitutes that abound in the present era. One of the problems with defining the money supply is the confusion between what actually *is* money, and what people *perceive* as money. Traditionally, we would view cash on hand (i.e., in circulation) and that readily redeemable in demand deposits as comprising the monetary stock. Rothbard includes five categories of goods that comprise the monetary base, as they are *perceived* to be money by the general public:

[1] Cash in circulation
[2] Demand deposits
[3] Time deposits
[4] Savings and Loan Capital
[5] Life Insurance Cash Surrender Values

However, these last four classes retain value as money only to the extent that they are perceived as being exchangeable at par for true money (i.e., cash). As Rothbard (1978: 145) reminds us, even for the relatively safe category of demand deposits:

> It is important to recognize that demand deposits are not automatically part of the money supply by virtue of their very existence; they continue as equivalent to money only so long as the subjective estimates of the sellers of goods on the market think that they are so equivalent and accept them as such in exchange.

As Salerno (1987: 1) points out, "money serves as the *final means of payment in all transactions*." All five of the above listed components fit this description, as all are available (or perceived to be so) to be exchanged for what we more commonly refer to as money (or cash in circulation in the above scheme), at par at any time. As Shostak (2000:73) notes, the crux in differentiating what should and should not be included in the money supply rests on the distinction between claim and credit transactions. The essential element, then, is that what we define as money remains a present good, always available on demand. Indeed, whether money fulfills this requirement will be subject to change as individuals' perceptions are altered as to a goods' suitability in fulfilling this criteria. As Shackle (1967: 6) states:

> It is uncertainty which gives to money every character and capability which distinguishes it from a mere numéraire. Money is the refuge from specialized commitment, the postponer of

the need to take far-reaching decisions. Money is liquidity. Money is not mechanical nor hydraulic, but psychological.

Base money may always be established as that which we know will forever be available instantly, at par, and on demand. This roughly translates into what is commonly referred to as MZM (money of zero maturity) today, or cash in basic terms. However, the subjective interpretation of what may actually function as money makes analysis of this factor less determinant, and more subjective than is commonly assumed.

Erroneous Capital Investments

In chapter I, we looked at the reasons why production actions can not be considered as productive if they contain an uncertainty as to whether they will succeed in increasing consumer want satisfaction. One general conclusion that we drew was that these actions were not to be considered consumption, or production, but instead represented an element of exchange only. While this applies for the actual actions that are undertaken, what of the intermediary (or capital) goods that are used or created through these actions?

These capital goods which have been created through the process which was thought to be production may turn out to be erroneously produced. If the goods prove futile in increasing consumer want satisfaction, they will have to be assigned as a role that is no more than merely exchanging goods from one place to another. However, we have also seen that this will only be known in an ex post sense.

A complication arises owing to the substitutable aspect of capital goods. As capital goods that are produced for one purpose turn out to be not fully suitable for this role, this aspect allows them some amount of leeway. Hence, although a truck for personal use may have been produced thinking their was consumer demand for its services, it may be that it will also serve as a truck for a company as a capital good. Hence, its value may be retained, or rather, its value will now become its next best use.

As Huerta de Soto has demonstrated the difference between capital goods and capital, we can see the same difference between exchange goods and exchange. For, as the exchange goods are only a special type of capital, we run into the same difficulties that have previously been looked at in aggregating them. However, by using monetary prices and freely established exchange values, we may come up with a homogeneous measure of the capital in existence at any given time. Capital goods will forever be valued according to the role they serve in the production of consumers' goods. The best forecast available to entrepreneurs as to the true value this capital will serve in ultimate consumption satiation will be assigned to it any every given moment.

As an historical exercise, once it is known if the capital goods actually did or did not provide value in increasing consumer satisfaction, we may then state ex post whether they were capital or exchange goods. Ex ante, we can only take the expectations that are priced in through the market process as the best guess as to what value will manifest in the ultimate consumption goods that the capital will create at an unknown time in the future.

4. Conclusion

Goods can thus take on three forms. Consumption goods directly aid in want satisfaction that ultimately provides the cause of all action. Production goods contribute to creating these consumption goods. One special case for goods exists only as exchange. These goods are only held to be exchanged for other goods – either capital or consumers' – which directly or indirectly affect our ultimate want satisfaction. Ex ante, the only type of exchange good we find is that of money – the generally accepted medium of exchange. Ex post, however, we see the possibility of entrepreneurial error implies that what amounted ex ante to capital goods, may not serve any indirect role, or may serve only a diminished role, in the satisfying of consumer wants.

Although historically we may look at failed capital investments as mere exchanges, from an ex ante perspective we will have to base our calculation of the capital in existence on the value appraised on the market by entrepreneurs. As this will be a subjective evaluation made on the expectation that the capital goods will yield consumers' goods that provide want satisfaction at some uncertain future date.

Substitutability is an advantageous aspect to capital goods viewed in this respect. As capital goods may not be fully able to be used for the originally planed consumers' good they were intended for, this factor will allow for some value-retaining leeway. As their use can change, value will be saved, and the capital good need not be fully abandoned, but rather can be salvaged and used for its next-best use.

IV. PROBLEMS WITH THE STRUCTURE OF PRODUCTION

The concept of a structure of production is now a mainstay within Austrian school macroeconomics. As early as Jevons (1871: 231) a corresponding structure of consumption has been utilized as well. One would think that this group of economists would readily embrace this offsetting structure of consumption for two reasons. First is the emphasis placed on consumers' wants and the value placed on them in the determination of the values imputed to the capital goods. Second is the acknowledgment that all production is undertaken to create a good which will eventually satisfy a human want – a consumption good. However, while this stress on the consumptive aspects of the economy are given *prima facie* attention, the reality is much less clear.

Three broad groups of Austrian economists seem to comprise the current opinion concerning the need, or desirability, for a structure of consumption.

The first are those who choose to ignore it, out of neglect, or by outright dismissal. In fact, this group seems to be in the majority, with mentions of this concept receiving little to no attention.

The second group are those who profess a need or desire for this concept, but fail to develop it further. Lewin (1999: 57-58) notes that resource utilization depends on the link between the "*structure of consumption and the structure of production*." Barnett and Block (2006:64) provide a lengthy critique of the structure of production, as illustrated through a Hayekian triangle, and note that a structure of consumption should become an integral part of Austrian business cycle theory (ABCT). One of the most telling calls for the structure of consumption comes from Skousen (1990: 372) who opines that an important area of investigation is:

> … how long the period of consumption might be. Much research has already been completed on the durability of machines, producers' goods and consumers' goods, especially to determine accurately the depreciation schedules for tax and accounting purposes for machinery, equipment, and transportation. A thoroughgoing study of the age structure of durable capital and consumer goods would be extremely helpful in accurately describing the standard of living in a country.

However, although all four authors cited stress the importance and advantages of conceptualizing a structure of consumption, none has as of yet made any headway into more concretely defining the aspects of it that would prove helpful to further analysis.[272]

The other group is embodied by Garrison (2001: 47-48) who states that although a structure of consumption is conceptually possible, it has nothing significant to offer other than increased complexity. Regarding the role of consumers' goods in a structural representation, Garrison (2001: 48) states:

> The notion of stages of consumption has much [*sic*] more limited interpretation than the corresponding stages of production... Although the allowance for consumption time as well as production time may constitute a move in the direction of realism, there is little to be gained analytically by replacing multistage Hayekian triangle with the Jevonsian investment figure Durable consumption goods and durable capital goods are obvious and, in some applications, important features of the market process. But to include these features explicitly would be to add complexity while clouding the fundamental relationships that are captured by the simpler construction. Instead, we avoid this graphical complication and rely on informal discussion to qualify our applications of the simple capital-based framework.

Unfortunately, it seems to be the case that Garrison, fixed within the confines of his three-diagram model, cannot reconcile consumption with production activities adequately to be shown to have a strict relationship able to be diagrammatically illustrated. He takes great pain to define the temporal units that define his *x*-axis of the structure of production. Hence, what is measured is not strict temporal units, but rather production-time – as noted as "stages of production." However, there is no such corresponding temporal aspect to consumption, at least not in Garrison's eyes. The reason is that there are no obvious stages of consumption.

If we view stages of production as strictly defined value-adding stages, as Garrison does, than we can see there are no such corresponding stages to consumption. This arises naturally from Garrison's (2001: 46) simplifying assumption that stages of production should be set at five in his diagram, which he assumes creates just the proper amount of complexity to correctly see what is occurring within the productive structure:

[272] This could be due to the influence of Mises (1949: 315-316) who viewed consumers and producers as only two sides of the same coin: "[P]roducers and consumers are identical. Production and consumption are different stages in acting. Catallactics embodies these differences in speaking of producers and consumers. But in reality they are the same people." However, as will be shown, as the structure of production is divided into stages, so to will these consumptive stages gain relevance if included.

> The identification of the individual stages is strictly for illustrative purposes. The choice of five stages rather than six or sixty is strictly a matter of convenience of exposition. To choose two stages would be to collapse the triangle into the two-way distinction between consumption and investment – the distinction that gets emphasis in the PPF. To choose more than five stages would be to add complexity for the sake of complexity.

This limitation arises necessarily from the conception of "stages" of production. We may say that there are material stages whereby production goods undergo value-adding changes, however, this is an *almost* meaningless concept when speaking of production as a *process*. The reality of the situation is that there are not a defined number of productive stages comprising the structure of production, but rather a continuous stream of stages, nearly infinitely small so as to satisfy the fact that each action performed on the higher-order good is adding value and transforming it to a state nearer to the final point where its fundamental nature is altered – the change from a good being *merely* productive to where it is used directly for consumptive purposes.

In fact, Garrison realizes the limitations this places on his conception of a productive structure moving through stages, stating (2001: 47):

> The continuous-input/point-output process that is depicted by the Hayekian triangle takes time into account but only as it relates to production. Adopting the point-output configuration gives us a straightforward link to the consumption magnitude features in our PPF quadrant. But point output implies that consumption takes no time.

Hence, by treating production as a type of "point-output" process, the elimination of time as it pertains to consumption becomes apparent.[273] The more realistic nature of production, however, is not that it occurs throughout stages whereby its value is transmitted to the next lower stage of production, but rather that it occurs continuously with value being transmitted to a point *closer* to the point where the good is transformed to a consumers' good. This confusion and its ancillary problems result from viewing the productive process in objective terms which are strictly defined in terms of concrete stages.

[273] Without a structure of consumption, we fall half-way into Knight's (1934: 275) trap of 80 years ago whereby "*production and consumption are simultaneous*." Although Garrison thinks production takes time, the duration of consumption has no effect on his 3 diagram model. Consumption must occur fully and simultaneously with the completion of production.

1. Hayekian Triangles or a Structure of Production?

Hayek, fixed within the confines of his triangles, had a different view on the business cycle than his contemporary Mises. The latter, while continuing to use pure verbal logic to explicate the issues, was freed from the constraints that plagued much of Hayek's work on ABCT. The emphasis Mises placed on the two-fold problem with business cycles contrasts sharply with the uni-causal approach stressed by Hayek. Thus, as Garrison (2004) elucidates, Mises was well-aware of two opposite forces working to disrupt healthy business conditions – *malinvestment* and *overconsumption*. Only the former was able to gain notice in the writings of Hayek.

The concept of overconsumption is debatable from a purely theoretical perspective. For the very term – overconsumption – seems to imply that there is an optimal amount of consumption from which to compare these levels to. However, the comparison is never between absolute levels of consumption, but rather between that level which would normally obtain, and that which is enticed through artificial conditions. To use Garrison's loanable funds model, an increase in the supply of loanable funds, in excess over that level freely supplied by individuals, acts to increase the total amount of investable funds available. At the same time, the interest rate on these funds is reduced below what it would have been absent the credit manipulation, an event which provides a disincentive to save. The result is that not only have total funds available for investment been increased, but that portion attributable to *real* savings has decreased owing the disincentive to save at the new, lower interest rate. The corollary to the reduction in savings is the increase in the rate of consumption – an event which gives rise to overconsumption.

Part of the difference between Mises' stress on the concept, and Hayek's neglect, is the latter's definition of forced saving. Hence, "Hayek's *forced saving*, rather than being the antonym of *overconsumption*, is actually a synonym for *malinvestment*" (Garrison 2004: 328). Of course, this confusion concerning forced saving can hardly be attributed solely to Hayek, as Machlup (1943) was able to unearth some 34 different definitions of forced saving. One significant issue with Hayek's use of the term 'forced saving' is the use of the term saving to imply a pattern of investment (Garrison 2004: 329).[274] How Hayek preferred to express his concept of forced saving was as an artificially induced increase in capital accumulation. However problematic his definition of this concept was, it was to become a mainstay in the ABCT, with particular attention given to the concept by followers in this

[274] Of course, even Hayek himself would not disagree that the term was "a rather unfortunate expression" (1928: 220).

Hayekian tradition.[275]

In distinction, Mises views the forced saving process as a shift of available savings from workers (who tend to have low saving preferences) to entrepreneurs (who tend to have a higher preference for saving).[276] Hence, the shift causes the total amount of saving to be higher during a credit expansion than would have occurred absent the credit injection, and hence, the natural rate of interest undergoes a corresponding decline. For Mises, we find that this extra amount of savings is the forced portion, an amount much different than Hayek's. As Garrison (2004: 330) summarizes:

> [U]nlike Hayek's forced saving, the term in Mises's argument (as in Sraffa's) actually refers to a particular instance of saving rather than to a pattern of investment that is at odds with saving preferences. Mises differs from Sraffa, however, on the issue of the magnitude of such saving in comparison to the saving actually needed to see the policy-induced investments through to completion.[277]

The Misesian stress on overconsumption is an event which compounds the problems of forced savings (as defined and used by Mises). As Mises (1953: 362) writes:

> A time must necessarily come when the means of subsistence available for consumption are all used up although the capital goods employed in production have not yet been transformed into consumption goods. This time must come all the more quickly inasmuch as the fall in the rate of interest weakens the motive for saving and slows up the rate of capital accumulation.

Hence, malinvestment *begets* overconsumption.

Hayek fails to consider the overconsumption aspect of the credit expansion process. Hayek does (1939: 172) refer to the end of the boom as a period of "relative overconsumption."[278] However, the relative aspect in question is compared to the level of consumption consistent with the completion of the newly undertaken investment activities, not relative to the pre-expansionary rate of consumption.

275 At one point, Hayek (1939: 197) even found agreement with Keynes' early definition as investment in excess of savings.
276 See, especially, Mises (1949: 548-565) for this elaboration.
277 Mises makes no point this that forced portion of savings need be a positive amount, but rather that it will depend on the particular pattern of wealth redistribution affecting the relative saving preferences of the concerned credit recipients.
278 As Garrison (2004: 333fn6) insightfully muses: "We can imagine, however, that in light of the trouble Hayek had in defending the idea of forced saving (i.e., malinvestment) to his English audience in 1931..., he would not have been eager to add that overconsumption was occurring at the same time."

Early on in Hayek's trade cycle writings he considers overconsumption excluded by the occurrence of forced saving. Hence, “[t]his phenomenon, we are to understand, consists of an increase in capital creation at the cost of consumption, through the granting of additional credit, *without* voluntary action on the part of individuals who forgo consumption, and without their deriving any immediate benefit” (1928: 218-219). Later, Hayek denies the possibility of overconsumption completely (see, for instance, 1935: 88).

Garrison (2004: 334) explains the differences between the two prognosticators' viewpoints succinctly:

> In summary terms we can say that Hayek sees the boom-bust cycle as forced saving, which is eventually countered by intensified consumption demand; Mises sees the boom as malinvestment, which is immediately compounded by overconsumption. We now understand that Hayek’s forced saving and Mises’s malinvestment are the same thing.

Hayek's stress on the role of forced saving (malinvestment) at the neglect of overconsumption is evident in his use on Hayekian triangles to diagrammatically expose his view of the trade cycle's progression. Followers of this tradition have been latently influenced, with forced savings today being the core tenet of ABCT, at the expense of overconsumption.[279] This following should be replaced with a more fundamental understanding of the role overconsumption plays in the boom-bust cycle. The lack of attention explicitly paid to consumption in ABCT contributes to this misunderstanding. The addition of a *structure of consumption*, as we will see, will improve the understanding we have of this process.

[279] Garrison (1995), for example, makes no account for overconsumption in the boom-bust cycle. His amendment of this absence in Garrison (2001) adds this critical difference to increase logical application of the ABCT, and also a more encompassing rendition of the general (i.e., Misesian *and* Hayekian) theory of the trade cycle.

2. The Subjective Nature of Production

Hayek's original depiction of his famous triangles were structured as continuous-input/point-output processes. In describing his first use of a Hayekian triangle, he (1935: 38-40) comments on the relationship between production, consumption, and time as follows:

> In order to get a clear view of what is actually implied by these changes in the structure of production, it is useful to employ a schematic representation. [footnote omitted] For this purpose, I find it convenient to represent the successive applications of the original means of production which are needed to bring forth the output of consumers' goods accruing at *any moment of time*, by the hypotenuse of a right-angled triangle... The value of these original means of production is expressed by the horizontal projection of the hypotenuse, while the vertical dimension, *measured in arbitrary periods from the top to the bottom, expresses the progress of time*, so that the inclination of the line representing the amount of original means of production used means that these original means of production are expended *continuously* during the whole process of production. [emphases added]

In this short quotation, Hayek is very clear as to how he wishes to view the productive process. First, we may see that when he is describing the output of consumers' goods, this is a point-output occurrence. The value of total consumers' goods being created at "any moment in time" necessarily implies that these are not continually occurring, but rather we are only looking at a snap-shot of this longer process. What Hayek is really doing is viewing the final transformation of capital into consumption goods as an *equilibrium* event. Hence, this is not ongoing, at least not in the analysis he is undertaking. Rather, this is a momentary occurrence, with the transformation process really being removed and replaced with a unique transformation event – goods are changed from producers to consumers in a single moment.

Second, we may note that Hayek *originally* views production, in distinction, as a continual process. By viewing the progression not in terms of stages, but in terms of "arbitrary periods" of temporal progression, we see that a pre-determined length of productive stage is excluded. The use of the word "arbitrary" to describe this temporal length may be an unfortunate word choice, but it somewhat accurately purveys the conception of time that Hayek wishes to express. Time, as it relates to production, is not an objectively-defined quantity, one that would open itself to being described as "stages." Rather, it represents a continual flow which is "arbitrary" in the sense that it is subjectively defined by each value-adding action that is enacted on these intermediary goods.[280] Hayek found a fault in using the continual flow of time as a starting point in production, and quickly rescinded the view to be replaced with that of more discrete temporal units (1935: 43):

> A perfectly continuous process of this sort is somewhat unwieldy for theoretical purposes; more-over such an assumption is not perhaps sufficiently realistic. It would be open to us to deal with the difficulties by the aid of higher mathematics. But I, personally, prefer to make it amenable to a simpler method by dividing the continuous process into distinct periods, and by substituting for the concept of a continuous flow the assumption that goods move intermittently in equal intervals from one stage of production to the next... Probably the simplest method of transforming the picture of the continuous process into a picture of what happens in a given period is to make cross sections through our first figure at intervals corresponding to the periods chosen.[281]

The shift that occurred in Hayek's view on production went from one extreme to another. After viewing production as a continual process originally, he altered his view to account for strict objective stages to account for the temporal passage. However, in reality, the answer lies somewhere between these two extremes. It is evident that it is not the *mere* passage of time that changes goods from a higher order to a lower one. Instead, it is the value added through action at each intervening moment in time that moves these goods from higher to lower stage. If we wish to use the word "stage," it must be understood that, when referring to the production process, stage is really an intermediate value-adding action. With this in mind we must recognize that although production is discrete, these stages are almost infinitesimally small as to *appear* to be a continual process.

In fact, if we return to Hayek (1935: 40) we find that he does indeed realize that value is added

[280] Adding to this confusion is the use of temporal stages in the first Hayekian triangle drawn in *Prices and Production* (1935: 39). As we shall see shortly. The use of subjective, continuous time by Hayek was to be short-lived as he would replace the notion shortly after with a more objectively defined view-point.

[281] The simplifying assumption of "stages" of production receives almost identical justification by both Hayek and Garrison. Hence, Hayek (1935: 43) states: "In this way, in my view, the loss in precision is more than compensated by the gain of lucidity." Garrison (2001: 46) phrases it as: "The choice of five stages rather than six or sixty is strictly a matter of convenience of exposition... To choose more than five stages would be to add complexity for the sake of complexity."

“continuously” while production occurs discretely through stages. A bifurcation occurs latently as Hayek's original Hayekian triangle has a y-axis which measures time in stages, but a hypotenuse which measures value additions as continual.[282]

Hayek, and followers using his triangles (i.e., Garrison), have fallen into a trap where the confines placed on them by viewing production in a strict objective terms necessary for the use of the Hayekian triangle have lost the true meaning of the production process. Production when viewed as a structure of processes is not defined by objective stages that are compared with one another. Instead, these stages result from aggregating together many individual actions that are involved along the path towards altering higher order into lower order goods. It is more meaningful, and theoretically correct, then to define the structure of production as a *series of interlocking actions which each add value to an intermediary good in order to transform it to a consumption good*.

If we return then to the earlier criticisms leveled against the concept of a corresponding structure of consumption, we find that the rationales were misguided. It is evident that consumption has no strictly definable “stages” of use where value is eliminated from them. However, it is also evident (now at least) that production too suffers this deficiency. Stages – mining, refining, manufacturing, distributing, and retailing – exist only as the combination of many individual actions which add value to a good as it passes through these stages. The stages themselves do not exist to create the actions transforming the goods from higher-order to lower-order goods.

We then must ask the more pressing question – is the structure of production a theoretically sound concept? Earlier we have defined productive actions as those which result in a movement towards closer attainment of consumer want satisfaction. The structure of production becomes that series of small actions which combine to move a good from higher to lower order – or, from orders more distant from final want satisfaction, to those closer to consumer want satisfaction. The structure exists in objective terms – there are actions which will move us closer to out want satisfaction and those which will also fail to do so. However, these individual actions become encased in a subjective nature, as whether they move a good toward the final state of consumers' good or not may be fundamentally unknown, or at least indeterminable, to all but those who actually partake in the action.

The future is a fundamentally uncertain state of affairs – our actions in the present create it as we move forward in time. However, our actions are undertaken only with the expectation that we will create *a specific* future, never one that is already known in advance to obtain. However, the fact that

[282] Hayek's original rendition of his triangle had time placed on the y-axis. Garrison (1978) switched the axes to give the triangles the more modern appearance, which has placed time on the x-axis through convention.

production occurs demonstrates that there are processes which are successful completions of productive actions. Hence, we can see that the subjectivity of the *success* of a productive activity does not prohibit the use of the concept of the structure of production. Instead, it implies that we must recognize the need to phrase this conception in terms more conducive to truly describing how the process evolves. The removal of stages as the basis for this process, and the substitution of value-adding actions demonstrates the true nature of the production process.

3. Conclusion

The preference among current Austrians to describe ABCT using Hayekian triangles has had both beneficial and detrimental results. The triangle as an heuristic tool has been both praised (Sechrest 2001; Holcombe 2001) and lamented (see Hülsmann 2001; Barnett and Block 2006). In fact, as Hülsmann (2001, 36) comments concerning Garrison's use of Hayekian triangles as the core of his three-diagram model:

> The diagrammatical exposition of this model has much aesthetic appeal, and thus it might not be unreasonable to expect that Garrison-style macro-economics will make Austrian converts among mainstream economists and their students. Yet there is also the possibility that Garrison not so much succeeds in Austrianizing the mainstream as that he will mainstream the Austrians. For the fact remains that Garrisonian macroeconomics is essentially neoclassical macroeconomics, which he enriches with an Austrian model.

We have seen some of the limitations that the Hayekian triangle have placed on modern Austrian macroeconomics. In particular, they are not only of modern importance, but also historical, as we have seen how Hayek's own views towards forced saving and overconsumption were constrained as a result. One of the benefits Austrian economists have is there attention to the temporal element as it serves the action process. The neglect of this element concerning consumption has been a direct pitfall of Hayek's own neglect, especially given the stress laid on an heuristic tool – the Hayekian triangle – and the associated problems that illustrating consumption through it can entail.

A secondary problem has manifested more recently, with Garrison's use of the triangle, commencing with Garrison (1978). “Stages” of production have been placed on the *x*-axis, with the implication that time is no longer of an ill-defined or subjective duration, but rather takes on some sort of pre-defined and objective length – the length of distribution time, for example.

Both these issues have been rectified. Hayek's views on consumption, particularly overconsumption, have been contrasted with those of Mises, and the role this consumptive process plays in the economy has been stressed. We have also seen that production is not a lengthy process defined by objective stages, as per Garrison. Instead, the defining characteristic of production is that value is added to a production good as it is transformed step-by-step closer to satisfying consumers' wants directly. That is to say, production is always moving production goods to a state closer to becoming consumers' goods. Hence, production is not only a process progressing through aggregative stages – mining, refining, manufacturing, distribution, retail, etc – but one that is continual through the multitude of individual actions that move these production goods forward along the structure of production.

The structure of production has gone through many changes since Jevons first introduced the concept. It may only be seen as a little ironic that one of Jevons' main opponents in the Marginalist revolution – Menger – spawned a school of thought that would independently adopt his concept of the structure of investment, only to discard half of it as unnecessary. More than 125 years after the publication of Jevons' *Theory of Poliitical Economy*, economics reaches towards re-embracing this near-forgotten concept.

V. THE STRUCTURE OF CONSUMPTION

1. Consumption and Its Structure

Having theoretically justified the existence of a structure of production based upon many intermediary stages defined not in terms of their general position in the order of goods, but rather in their more specific quality as value-adding actions in this production process, we find that a similar concept may be applied to the consumption process as well. Consumption goods have a definite *point* in time when they switch their goods' nature – from producers' goods still requiring value to be added, to consumers' goods ready to satisfy wants directly. In fact, one of the greatest contributions of Hayek's *Prices and Production*, is the inclusion in his second Hayekian triangle to illustrate this point (see for instance, Hayek 1935: 44, figure 2). We have previously seen that no specific stages exist which demarcate different *stages* of production, but there is one definite *point* at which a producers' good is transformed to a consumers' good. This point also sparks the time whereby the act of consumption commences, and that of production ceases. Hence, these two related concepts – consumption and production *and* consumers' goods and producers' goods – are each inextricably linked, with no cross-usage possible (that is to say, no consumers' good can be used directly in the production process, lest it be a producers' good). In fact, this point is evident in what Strigl (2000: 27) termed the "subsistence fund." There can be a typically defined consumers' good which can really defined as a factor of production once we factor for the case that it is a necessary factor of production. The prior saving of consumers' good, such as food, to sustain us during the production process is really a form of capital which is necessary for the continuation, and completion, of the productive process.

It then becomes clear that only those goods which service consumption activities may be included in the structure of consumption. It is evident that consuming food is a necessary condition for our prolonged existence, and that eating food is neither a productive or an exchange activity – thus, it is consumption oriented. However, food that is eaten during the process of producing a good is necessarily a producers' good.

To take the classic example of Crusoe consuming all of his productive output of berries. By saving berries for later consumption, he finds that he may take a break from producing merely for consumption, but may dedicate himself towards production of a producers' good – a stick to knock berries off a tree in order to increase productivity. It is easy to see in this example that the saved berries are what makes the production activity possible – *they are in fact producers' goods*.

Problems with the Flow of Value in the Structure of Consumption

One of the assumptions we have used until now is that the structure of consumption must always be downward sloping. This arises for two reasons. First, satisfactions in the future are discounted relative to the present due to our time preference. Second, the serviceableness of consumers goods decreases over time – cars deteriorate, for example, and provide less want satisfaction when used than when new.

However, counterexamples to this second general tendency may be found if sought after. Some automobiles tend to increase in value over time, not decrease. We may note that a 1932 Ford roadster is worth more today if we were to try to purchase one than it was when brand-new leaving the factory in 1932. This example seems to neglect the use of a structure of consumption that slopes downward.

However, this example bifurcates between two opposite activities that have some similar features. It is obvious that the nature of the good – the car in our example – is of a consumptive nature rather than productive. Were it lying idle awaiting further value-adding actions it would clearly be a production good, and as such, a part of the structure of production. However, while the goods' nature is seemingly objectively that of a consumption good, subjectively the owner is using it for a much different purpose. For if the owner used the car as a consumption good (that is to say, as a car *qua* car) its value would decrease as it was used. However, the owner forgoes consuming the car's services, preferring to hold it as an exchange good. Through this manner, as the good in question is not used, its value is not destroyed (or *depreciated*, to give more general accounting terminology) and its value is preserved.

In this way, we may see that the car is held not to be used but to be exchanged at some future date. The sole reason then that an individual would hold onto a good for exchange at a future date, instead of exchanging today, is if the expected value in the future is greater than today. This comes full-circle to the reason why the value of such a good is expected to be increased, and why it is held as an exchange good.

Another example may show a different category of goods which shows this same confusing tendency for value to increase. Wine typically requires an aging period to realize its full value. We may see that an individual may purchase a bottle of wine, and hold it in a cellar for a period of time in order to consume it later when it has matured into a higher value product. Is this, then, not a counterexample of the tendency for consumers' goods to decrease in value over time?

Another bifurcation of goods' orders becomes apparent. Mises (1949: 479) demonstrates that the aging of wine is part of the production process. What has changed is not the good from producers' to consumers' good, but rather it is only the owner of such a good who has been changed. A consumers' good is that which is available to be consumed at the present time – all necessary value-adding activities of production are complete and it is ready for consumption. Normally, wine is purchased to be consumed immediately. The moment it is sold the wine has changed both owners and goods' character. However, if an individual buys wine with the expectation that they will hold it for a period of time to increase its value, they are not buying it as a consumer, but rather they are continuing to be a producer of the product. Although they may not pay an explicit price when they determine the wine has increased in value enough to consume, they are still holding the wine to further the production process, at the expense of present consumption of it; *the wine remains a producers' good*.

Consumers' goods serve their role as they are immediately available to be used to further want satisfaction. This can not be confused with producers' good which always require additional value-adding activity to realize this potential, or exchange goods which satisfy no want directly, instead being held only to exchange for another good – they satisfy wants only *in*directly.

In fact, we may state that consumers' goods satisfy wants directly, producers' goods satisfy wants eventually, and exchange goods may only satisfy wants *in*directly.

Consumption necessarily implies that value is being used from a good. Wine that is used as a consumers' good is drank, and hence, the quantity is diminished giving it a decreased *total* value as it is used.[283] Another defining characteristic of a consumers' good becomes clear in that its value *must* decline as it is used. Quantitative differences (i.e., the amount of wine remaining to be drank) as well as qualitative differences (i.e., the deteriorating condition of a car as it is driven) combine to decrease the total value of a consumers' good as it is used *qua* consumers' good.

Duration of Serviceableness

Previously we have seen that consumption goods have a definite life span. That is to say, their serviceableness towards the satisfaction of wants declines continually. This decline will continue until one of two eventualities are reached.

[283] Additionally, there is the marginal decrease in satisfaction with each additional unit of wine drank.

Serviceableness of a good is completely removed and it is discarded. In fact, many goods today can be seen to have a limited useful life, after which point they are discarded. For example, it may be seen that an individual would like a new television to satisfy their entertainment wants. Other things being equal, a new television is preferred to an older one. This stems from two main reasons. The first is the preference we have for newer, more technologically advanced goods which service our ends better than older goods, which have been replaced for this very reason. The second is that the quality (and also quantity) of want satisfaction is decreased as these goods age and lose their want servicing ability. The television in our example will see the quality of its picture deteriorate over time, so that it is valued less than a corresponding new television. The added value of the new television due to technological factors is apparent, as color picture, picture-in-picture, and a multitude of other options have become available that were lacking in previous television models. Hence, an individual will continue to use this television as a consumers' good until the point where its value has declined to a point where the individual chooses to discard the good in favor of a newer (or different) option.

The second option concerns when a good is no longer valued higher as a consumption good than as a production good. We may see that an individual may have a landscaping company where the same good is needed as a production good, and also as a consumers' good in the owner's personal life. For example, a sturdy truck is necessary to move equipment to the job-sites, visit customers, etc. The owner may prefer to drive a truck as their personal vehicle, thus satisfying their consumers' want for transportation and pleasure of driving. However, a point may be reached where the owner receives inadequate want satisfaction from their truck as a consumers' good, and instead introduces it to their fleet of work vehicles. The shift of the truck as a consumers' to a producers' good has occurred as the owner has determined the serviceableness of the truck as a consumers' good is insufficient, and that it would be in their interest to begin using it as a producers' good.

Note that these two aforementioned options are questions of value. For it could be that the truck in our second example retains its value as a consumers' good until that point where it is no longer serviceable even as a producers' good. In this case, the truck would be discarded, or, at the least exchanged to another individual who finds it of greater value.

Hence, we see that value as a consumers' good diminishes over time, with the result that the structure of consumption will be downward sloping. That is to say, consumers' goods have their greatest value *as consumers' goods* at that moment when they are obtained. As they are used to satisfy wants, their value generally declines due to their finite life-span.

A secondary factor serves a role in creating a downward sloping structure of consumption. The discount of wants in the future compared to those in the present, owing to our time preference, will cause us to place greater value on a want satisfaction in the present over the same at a future date. Hence, if we value the want satisfaction of transportation, and the *only* way this want may be satiated is through a car, then we place greater value on a car today than tomorrow. We find, then, that owing this reason, even if we purchase a consumers' good today, and then do not use it (hence, no depreciation occurs), its value will be greater at an earlier rather than later date. The *option* will be available to use the good, which adds to the value it has in servicing wants, *even if it is not readily apparent that no want is being serviced*.[284] We may add to this by noting that the demonstrated preference of purchasing a good at t_0 instead of t_1 signals that the good was valued more highly at that time than at a later time.

Future versus Present Value of Consumers' Goods

Suppose that a stock of consumers' goods is purchased today, and that over 10 years no further additions are made to that stock of goods. If the useful life of these consumers' goods is 10 years, we find that they will be completely depreciated at that time – they will no longer have value *as consumers' goods*. Provided that no replacement to these goods occurs, a general decline in the level of consumer want satisfaction will occur owing the less value these goods will have at a later date. Two main factors give rise to this diminishing value of satisfaction from a consumers' good over time.

The consumers' good in question will exhibit a diminishing value as serviceableness deteriorates over the life of a good. As the usefulness of a good deteriorates, its ability to satisfy a want is also affected. For example, it may be that a car bought today satisfies our want for comfortable transportation fully, but over time, we see that tires wear giving a rough ride, seats suffer tears creating less comfort, and other value affecting changes occur. The result is that, absent continual maintenance of these goods, a deterioration in their serviceableness will occur.

The second reason stems from the changing matrix of wants we wish fulfilled. If we have a want existing in the present, there are two options for its degree of desirability in the future.

First we may see that the want is desired less by the individual in the future. If a good has already been purchased to satisfy this want, it will now derive less value in the future as the desire it

[284] Consumers' goods may service wants even when not directly in use. For it is true that many people own a plethora of consumers' goods which lay idle, waiting for a future event in which to use them as consumers' goods directly for the servicing of a want. The option to use them at a later date has value which outweighs that of not having them readily available for want satisfaction.

serves to fulfill will have diminished importance. No matter how much utility the good in question will serve in achieving a want, if the want is of diminished value, the value of the good will also be.

Second, we may see that a want is desired in the same magnitude by the individual. In this case, the good will be valued less as it has deteriorated through use over the period. The value assigned by the maintained use-value of the good to satisfy the want is reduced by the depreciated value that the good actually has in satisfying this want.

Hence, we find that consumers' goods will always have a greater value in the present compared to the future. The diminishing of value over time coupled with a shift in the total want matrix an individual has will alter the value a good will have for satisfying a given want.

Changes in Serviceableness of Consumers' Goods

We also may note that newer consumer goods may change in two regards concerning their serviceableness over their useful lifetimes. The length of time that the good will produce services for will be taken into account along with the degree of serviceableness upon purchase of the good. Both of these forces will work together to provide a slope on the structure of consumption corresponding to that of the structure of production.

The first way is through increasing the want satisfaction available in the present. If we have wants that are in need of satisfaction, we can see that there are different levels of satisfaction available. The *degree* to which these wants become satisfied becomes important as it provides the value we place in the present on the goods delivering this end. Different degrees of luxury provide a good example for our case. A consumer may want transportation, but it is easy to see that there are a multitude of options available to be used. Walking may generally be seen as the least comfortable option, but also entails, in most cases, the least cost one. City-dwellers often find metro systems available to be used inexpensively, but this is also less comfortable, and often more time consuming, then taking a taxi. Likewise, intercity buses may be used, but these are normally seen as inferior to intercity trains, which in turn are seen as a less luxurious option then taking a car. When we make the progression to a car, we see that a Chevrolet may get you to the destination where you want to go, but a Bentley will get you there in a more comfortable manner. Likewise, a Citröen 2CV will *eventually* get you to a destination, but a BMW M5 will get you there faster and in greater comfort. In all these examples we notice that there is a general tendency for a trade-off between the *degree* of want satisfaction and the cost of obtaining it.

Of course, the subjective nature of the want is of the most pressing concern. A 1957 Mercedes-Benz is neither as sporty, nor comfortable as an entry level Ford of today. However, other factors enter the picture which make the Mercedes have a value much greater to the purchaser than the Ford. Other determinants enter the picture, as value is assessed on a particular good to service consumption wants. However, the common bond with all consumers' goods is that serviceableness will decrease over their life-span as usage causes wear on their abilities.

The second is through increasing the duration of serviceableness, or phrased differently, the life-span of the good. We may readily see that some goods provide services for extended periods, while others do so for only short periods of time. To keep our automobile example, we may see that there is a difference in durability between different makes of car. Mercedes-Benz diesels are famous the world over for their ability to continue working for what seems to be forever in some cases. Fiat autos from Italy, in opposition, are famous in some circles for having very short life spans.[285] The length of time that a good is expected to render consumers' services for conditions the value it will have in the present.

We may notice that this point will mark all goods. The *length* of serviceableness is generally one that humans can alter when creating consumers goods; more so than the *degree* of serviceableness. This arises from the fundamental activity that an entrepreneur serves, as discussed in Part I, chapter V. The two-fold role that the entrepreneur serves is to discover new opportunities for want satisfaction, and to improve upon old ones. The discovery of new avenues to achieve this is a fundamentally uncertain occurrence, whereby the entrepreneur creates the future based upon their expectations of what entrepreneurs will desire at this uncertain future date. The entrepreneurial role to improve upon goods that have an already established demand is also vital, however, it can also be seen as being of a fundamentally different nature than the aforementioned skill. Hence, the creation of an automobile of longer lifespan than previously existed may be seen as a more fundamentally certain occurrence than creating an automobile to satisfy a want – when no such good for satisfying the want currently exists. However, we may see that goods may be made more durable – their quality and longevity may be increased – as part of a perfecting process which requires no degree of uncertainty as to its desirability.

Hence, we see that the shape of the structure of consumption is dominated by two factors. The *y*-axis represents the degree of serviceableness that the good offers. It may be noticed that more is desired to less in cases of want satisfaction, however, this is constrained by the factor that wants may

[285] As one old joke goes: “What does FIAT stand for? 'Fix it again Tony.'”

be satisfied only through an act of renunciation. Hence, the question then becomes "What is the best trade-off I can obtain between my degree of want satisfaction and the amount of want renunciation that must be incurred to obtain this degree of satisfaction?" The *x*-axis is controlled by a much different factor. This is produced by the time preference that an individual has to consume in the present over the future, and will be assessed in more depth in the next part.

Time Preference and Desired Length of Serviceableness

The trade-off individuals make between satisfaction in the future and the present is known as time preference. Want satisfaction may only be achieved through the renunciation of something that we value less in exchange. This is evident in catallactic exchange, where we physically exchange something to gain a more satisfying good (or service) in return. Earlier, we looked at how even in autarkic situations there is the renunciation of something as a requirement for want satisfaction – even if this is only the renunciation of free-time as we undertake an act aimed at increasing our personal satisfaction.

As the act of want satisfaction involves the satisfaction obtained through some degree of sacrifice, it becomes apparent that we do not always aim to have the maximum satisfaction of our wants fulfilled at any given moment. Instead, the judgment as to if the satisfaction outweighs the costs of the sacrifice serves to determine if we will undertake such an action. From this reasoning comes two corollaries:

[1] The shortest production processes possible will be used to yield a satisfying result.
[2] If a longer production process is used, the process must be expected to yield a more satisfying result.

The same logic may be applied to the consumers' goods as we have just used for productive processes. If renunciation were not a relevant issue, goods would be purchased to satisfy our wants with an infinite life-span. However, we see that all goods must have a finite life, and also we find that the trade-off between want satisfaction in the present and duration of serviceableness interplay to determine the life-span of durableness that we desire. It has been demonstrated that goods of longer want satisfying durability can generally be seen as more costly. That is to say, more present renunciation is required to obtain a consumers' good which is expected to continue yielding want satisfaction for a longer future time period.

We can see, then, that the *x*-axis of the structure of consumption is defined as the duration of serviceableness of goods demanded by consumers. This time invariant concept will be better demonstrated through an example. It can be seen that a tendency to purchase a car will result in an expected useful life over which the car will provide services. In our previous example, we saw how we may compare the longevity of a Mercedes-Benz diesel with a Fiat. The Mercedes may have been purchased in the expectation that it would yield consumer want satisfaction for a period of 10 years. The Fiat, in distinction, may have been purchased at a lower cost, but only expected to have a useful serviceable life of 5 years. When this 5 year period expires, if the want for transportation still exists, the individual will have to find a new way of satisfying it – provided that the Fiat was correctly judged to have the limited life span of 5 years.

A more interesting example may be found if we look for the method of paying for our wants to be satisfied. Hence, we may find that cars can be either financed through purchase, or through lease (or financing which is essentially equivalent). The amount of time expected to be in ownership of these consumers' goods is seen to be drastically different. For the shift from car ownership to leasing is generally seen as one made depending on which is more profitable given the expected time period to hold the asset. Ownership *in general* makes sense over longer periods of expected duration of an asset's life. It is also in the trade-off that each individual makes between renunciating more in the present for a longer duration of want satisfaction, or renunciating less and receiving a less durable good in return.

Therefore, we may see that individuals with high degrees of time preference would prefer to consume in the present, and not pay heed to the future relative to those with low time preference. These individuals will be less concerned with the future, and instead will desire present want satisfaction. *Ceteris paribus* these people – those with high time preference – will desire greater want satisfaction in the present compared to the future. Hence, those consumers' goods which they desire will be those that yield satisfaction consistent with these desires.

We may sum up with the following dichotomy then:

[1] High time preference individuals will prefer goods yielding high want satisfaction in the present, with a comparatively shorter period of provision (or durability);

[2] Individuals with low time preference are able to forgo goods yielding high degrees of want

satisfaction in the present, in exchange for a long duration of want satisfying abilities in the future.

Depreciation of Consumers' Goods

A related but secondary question arises concerning the serviceableness of consumers' goods. Namely, what sort of allowance must be made for depreciating aspects of these goods. We have seen that all goods must have a finite life-span, and that value generally decreases as the temporal point moves within the life-span from the point the good was attained, to a more distant (i.e., future) point of time. However, value may be maintained by allowing for depreciation, and compensating accordingly. In this way, we may see that maintenance of goods applies not only to those of the higher orders (i.e., capital) but also those of the *lowest* order (i.e., consumers' goods).

An example may prove illustrative. A house when used by its owner as a house is categorized as a consumers' good. It is plain to see that its value decreases continually as it is used and depreciation occurs. However, we may also see that homeowners incur a significant expense at continually up-keeping their homes in good order. That is to say that periodic re-investments allow homeowners to maintain the value of their homes, despite the continual process of depreciation which naturally occurs with all goods' classes. New paint, windows, shingles, carpets, or a multitude of other value-adding tasks are performed in the hope that the value will be maintained in this consumers' good – the house.

However, we may also see that, owing the finite life all goods must have, there will be a point whereby additional allowances for depreciation will fail to maintain the value of the consumers' good, and it will fall to one of two fates. First, it may be traded (exchanged) for a another good (or goods) of greater value. This means that the stock of consumers' goods is maintained, with only a relative valuational and ownership change in the structure of consumption. A consumers' good of lesser value has been traded for one of greater value – the total quantity of consumers' goods has not changed, however, the values attached to these goods has been shifted. The second option is that the consumers' good is at the end of its life-span, and must be discarded. Hence, we see that a loss occurs in not only the quantitative stock of consumers' goods available, but also in the value outstanding of these goods.

Note that when consumers' goods are exchanged, the total value actually increases. As exchange represents a non-zero sum game, each trader will be made better off (that is, receive more valuable goods than they renunciate), and hence, total value will be increased. This differs from the case where consumers' goods are discarded, as the value is lost at this point. However, we may also note that while value will be lost, it will generally be negligible. This arises as the choice to discard a good will only be made if the expected costs of doing so are less than the expected profits that could be made from exchanging the good for another of higher value.

There are those consumers' "goods" which are not able to be maintained throughout time directly, instead, they must be obtained afresh periodically. Primarily, we wish to refer to services in this category. A haircut, for example, cannot generally be maintained, but instead, must be replaced several times a year with a new haircut. There is also that class of goods which is fully expended at its first use. Food which satisfies our nutritional wants is fully used at the moment we ingest it, and cannot be maintained to continually yield satisfaction over time. Instead, we find that this category of consumers' goods requires continual replacement – new food must be purchased on an on-going basis to provide nourishment.

Concerning consumers' goods and their depreciation, we have a trichotomy of categories that each may fall under – services, durable consumers' goods and non-durable consumers' goods. Services and non-durables are generally a class of goods unable to be extended over time. Their satisfaction must depreciate over a schedule that cannot be lengthened, except by changes made at their inception. For instance, it may be that if I get a shorter haircut, it will take more time before the length of my hair is sufficient to warrant me returning for another haircut. The length of satisfaction from the service has been extended, but only through a change that occurred at the service's delivery, it cannot be extended after. However, durable consumers' goods have that ability to have their services lengthened and maintained over time through periodic improvements. We have previously seen with the example of a house how this is achieved.

The length of time that we wish a good to provide services for is tempered through our time preference, as previously explored. However, the durability of the good in question is also tempered through this same process. More durable goods will, ceteris paribus, cost more than their less durable counterparts. This is a wholly different statement from saying that the cost of production of a good will increase if its life-span is also increased. Instead, this statement refers to the expectation that more replacement of depreciated portions of the good will be required in less costly goods, in opposition to their more costly, but otherwise identical, counterparts. For why would any individual incur more cost to purchase a good, if all else was equal, except with the expectation that it would require less maintenance of its expected life.

Time preference involves the preference for present want satisfaction over that same satisfaction at a later date. We may compare two consumers' goods. Good A which has a life span of five years and requires no allowance for depreciation over this time period – its full value is lost upon expiry of the serviceable life. Good B has an identical life span but requires a yearly allowance of $10 to maintain its value. Finally, after five years, despite the yearly maintenance, good B loses all value. We may assume that each good yields an identical amount of serviceableness each year of $20. If we ignore discounting due to time, we see that good A will be purchased at any price under $100, and good B at any price under $50. Although identical in every other respect, the individual time preference will dictate which option the individual will chose. Those with a higher degree of time preference (i.e., those prefering present to future want satisfaction) will choose good B. They will renunciate less today with two resultant occurrences. First, they will be able to enjoy more want satisfaction today as they will still have at their disposal more resources (money in this case) to exchange for more want satisfying goods. Second is that they will have less want satisfaction in the future, as they will be required to continually maintain the good over its expected lifetime. The opposite holds true for low time preference individuals. They will renunciate more in the present (i.e., purchase consumers' good A) but reap an increased amount of future want satisfaction as less allowance will be required in the future to maintain the serviceability of the good in question.

Hence, we may see that the trichotomy of consumers' good available – services, durables and non-durables – are tempered by their respective needs for replacement or allowance for depreciation. The value released by durable consumers' goods can be maintained over time as maintenance is allocated to these depreciating assets. Hence, we see that much like producers' goods, there is a special type of consumers' good which also requires an allowance for depreciation. Lastly, we may see that time preference proves to be an instrumental part in selecting the degree of durableness that a consumers' good may have. For if two otherwise identical goods have different amounts of required maintenance to upkeep their services, than time preference will dictate to what degree a consumer will be willing to forgo current want satisfaction, in order to sacrifice more in order to obtain the good requiring less *future* sacrifices (through maintenance) to up-keep the services.

Conclusion

We have seen the validity of using a structure of consumption to offset that of production. As all production is undertaken to serve this end – consumption – the means of production and how they are directed may be fruitfully developed further by focusing on what exactly are the consumptive ends that individuals desire.

Three types of consumers' goods are incorporated into the structure – services, non-durable and durable consumers' goods. The three act in an integrated way to affect the structure in two dimensions. First is the degree of want satisfaction that will be given in the present through these goods. Second is the length of time that these services will be provided for. An ancillary problem develops when we deal with durable consumers' goods as they will require an allowance for depreciation in order to be used continually.

The amount of want satisfaction in the present and the overall length of serviceableness represent a trade-off of options. Consumers' goods that provide an infinite amount of want satisfaction are impossibilities (they would, in fact, cause us to cease acting if all our wants were fulfilled), and hence all consumers' goods must have a finite life-span. Individuals always make a trade-off between the degree of want satisfaction they expect to obtain from a good, and the cost of renunciation they must incur to obtain such a good. Our personal time preference will dictate which trade-off we wish to incur – more satisfaction in the present, or a longer term of satisfaction into the future.

However, even if the above considerations are held constant among a cross-section of goods, that is to say, even if two goods were identical in the degree of satisfaction that they provided in the present, and had an identical expected life-span over which they would provide equivalent services, there is an important difference that still must be analyzed through our time preference. As goods must have a finite serviceable life, they will necessarily depreciate over this period. This value depreciation may be offset through an allowance for maintenance of the goods. As a comparative example, a Mercedes-Benz diesel automobile may be purchased today for $50,000 and require no maintenance until its life-span of 10 years is complete. Alternatively, a $25,000 Fiat automobile may be purchased today, and have the same expected duration of serviceableness, but at the same time be required to have $5,000 of repairs each year to maintain its service providing abilities. If we assume the same degree of want satisfaction is provided each year, the determining factor of which option to chose is the time preference trade-off between increased want satisfaction in the present (via less renunciation with the purchase of the Fiat) versus greater want satisfaction in the future (owing the additional goods that will be available in the future with the Mercedes due to its lack of maintenance requirement).

The analysis of the structure of consumption yields three important conclusions. First is that the amount of want satisfaction that we wish in the present as opposed to the future is reflected bidimensionally through the structure. An increase in time preference will cause an individual to demand more want satisfaction in the present, while causing a simultaneous shortening in the structure's useful life-span (ceteris paribus). The corollary is that decreases in time preference will cause a lesser degree of want satisfaction in the present, and extend the period of serviceableness of the structure further into the future (again, ceteris paribus). Second is that the degree of durability of these consumers' goods will be affected directly by the operative time preference as well. These goods will become more durable as time preference rates decrease. This results from the desire to have services rendered for a longer period of time, while renunciating more in the present to obtain these longer lasting goods. The final point we may notice is the link between the originary interest rate and the structure's slope. Previously we have seen the effect of time preference through the originary rate of interest (and also the more visible market interest rate as well). As time preference rates increase, the originary rate of interest sees a corresponding increase. In the structure of production, as time preferences increase the present value component increases, at the expense of the period of serviceableness of the goods in question. Hence, the slope of the structure increases alongside any increase in the rate of originary interest – they are positively related.

The way that our time preference and the desire to have our wants satisfied will affect the shape and size of the structure of consumption. To the degree that value is imputed backwards from these consumptive ends to the producers' goods that provide them, we see that this structure of consumption provides an important link in understanding how changes in the structure of production occur.

VI. THE STRUCTURE OF PRODUCTION

Previously we answered some criticisms of the structure of production. In particular, Hayek's (1935) rendition of production as taking a triangular form divided into “stages” of production has suffered several detrimental logical inconsistencies. One of the more unfortunate results of these inconsistencies has been that the focus has been on predefined stages of production, which may or may not exist in reality. Consequently, later economists writing in this Hayekian tradition have fallen into the trap of ignoring the “structure of consumption,” as we have just defined it, owing to consumption's lack of “stages” corresponding to production. With this fallacy exposed, we have rectified the situation to emphasize production not as “stages” which transform goods from higher to lower orders (a definition which necessarily excludes a similar analysis concerning consumption), but rather as a process of value adding actions. As actions modify an existing good to add value, it is “moved” closer to the final stage of production where a capital good may finally be transformed into a consumers' good.

Of course, with production viewed as a series of value-adding actions, we see that consumption is the opposite process – a series of value-subtracting actions. Production and consumption entail continuous processes of adding and subtracting value from the goods in question.

A different issue arises with the structure of production, when used in its Hayekian formulation, as it excludes fixed capital at the expense of circulating capital (Nurske 1935). As a result there are two problems which arise and require rectification to create a usable structure of production. The first is that both forms of capital – fixed and circulating – must be included to demonstrate the true origin of capital goods as the “produced means of production.” Second, is that the linear representation of circulating capital progressing ever closer to the final consumption stage does not show how fixed capital is used in this productive structure. To see how the true structure of production is altered, we must take into account both types of capital, something which has been heretofore neglected.

At the same time, the issue of depreciation on capital goods becomes problematic. We have previously seen how durable consumers' goods also face this value-subtracting occurrence. With capital goods we find that the structure of production produces fixed capital for two causes. One is to replace and replenish depreciated capital assets. The second is the production of new capital which will further develop and lengthen the structure. How exactly this lengthening process is achieved must be looked at to determine how relative changes within this structure affect the greater whole.

Lastly, as all production is geared towards the production of consumers' goods, we see that changes in the structure of consumption are what entice producers to alter their own production activities. As consumption goods' are demanded in differing durabilities, durations, or present value-providing abilities, the structure of production must be likewise altered to accommodate these changes. We see, then, that the ultimate sources of change within the structure of production are changes in that consumption structure which is imputed backwards to it.

1. Circulating Capital

During the debates between Böhm-Bawerk and Clark in the late 19^{th} and early 20^{th} centuries, the conception of capital as a homogeneous fund, aggregated in monetary terms was the center of intense dispute. While Clark would view capital as a monetary reckoning of a homogeneous fund, he would point to the realism of this assumption:

> Ask a manufacturer, 'What is your capital?' and he will probably express his answer in dollars. Ask him, "In what is your capital invested?" and he will specify the buildings, machines, land, materials, etc., in which his productive fund now chances to be embodied. These concrete things will figure in his thoughts as the containers of his capital; while the content itself will appear to him to be a value, an abstract quantum of wealth. He will think of it as a fund that is permanently his, though it may not retain for a single day its exact present form of embodiment .... Capital is, in this view, an abstract fund, the destiny of which is to migrate thru an endless series of outward forms. (1888: 9-10)

While Böhm-Bawerk would essentially agree with this assessment of *true* capital as a monetary assessment of the capital goods, it was Clark's emphasis on three-ancillary factors which caused so much debate (for example, see Böhm-Bawerk 1906: 5). Hence, capital's permanency, synchronous nature of production *viz* consumption, and full-mobility would cause considerable amount of disagreement between the two. As a result, Böhm-Bawerk stressed capital goods' heterogeneity above all other factors, in direct response to the emphasis of an alternative viewpoint by his opponents.

Hayek, while describing his first rendition of his productive triangles, would make note that he too was using capital in a homogeneous monetary sense: “It should be remembered that the relative magnitudes in the two figures are values expressed in money and not physical quantities, the amount of original means of production used has remained the same, and that the amount of money in circulation and its velocity of circulation are also supposed to remain unchanged” (1935: 51). We find that explicitly, a Hayekian triangle concerns itself solely with circulating capital (original means of production have remained the same), and that this capital is reckoned in homogeneous monetary terms. Although this type of capital is essential, it is after all the goods that will become consumers' goods eventually, but it is only part of the total which we wish to look at.

If we start with Hayek's assumptions, we find that given a stock of fixed capital, circulating capital will begin its life at a high stage of production and continue toward the consumption stage as it is transformed to a consumers' good. The structure, or Hayekian triangle in this case, grows wider as it nears the final consumption stage as continual value is added to these pieces of circulating capital.

Goods are taken at an earlier (or higher) stage of production, and over time are mixed with human labor and the assistance of further capital goods to create a final product of greater value than the original had.

A case in point may be taken with the production of a simple pencil. Originally, a piece of wood and a quantity of graphite are mixed together to yield a writing instrument. However, additional value adding features are then added to better satisfy consumers' desires. A rubber may be affixed to the end of the pencil to erase mistakes. Paint may be applied to the exterior to allow for a more pleasing appearance, or a more welcoming feel to the user's fingers. The basic writing instrument – the pencil – may be stopped at any time after the creation of the minimal want satisfying good, or it may continue along the structure as features are added to enhance the value for the final consumer.

Note that the creation of the pencil has been undertaken with the goal of satisfying a consumers' want – either existing or expected. However, there are two points to keep in mind when satisfying this want. The first is that the actual want desired is, in this case, a writing instrument. The pencil could have satisfied this basic want at the stage where the graphite was inserted into a wooden casing, thus yielding a consumers' good capable of satisfying the desire for a writing instrument. The second point is that differing *degrees* of want satisfaction are desired by consumers. Hence, it could be that the basic pencil may satisfy the *base* desire, but that what is really sought is something of greater complexity to add value to the writing experience. In this case we find that although the base pencil may have satisfied the want adequately, there was a higher want satisfaction demanded by the individual, *beyond that merely provided by the basic writing instrument*.

The use of "stages" confuses the production picture slightly. Garrison's (2001) use of predefined stages such as mining, or distribution, may be *prima facie* pleasing to read, however, they confound two similar topics. Value adding actions need not correspond with pre-defined sequential stages of production. With our previous pencil example, we can see that adding stages to the process, such as a more intensive distribution stage for example, would do nothing towards creating a more valuable pencil for the end consumer. Instead, we find that in this specific case, the sole culprit contributing to increased want satisfaction was an increase in the individual value-adding actions within the manufacturing stage. Hence, only a relative increase of activity within one stage as compared to another has added increased value.[286]

Circulating Capital in Monetary Terms

Hayek's use of circulating capital in monetary terms may seem to add an element of homogeneity to capital that we wish to avoid. Even though the use of separate stages of production can do little to rectify this apparent problem. As capital is reckoned in solely monetary terms (itself representative of consumer value) as it passes through separate stages, there is no differentiation between capital in one stage compared to that in another. They become identical entities separated only by a temporal element.

This seeming problem becomes a moot point when dealing with circulating capital, as Hayek must have realized when employing it. As all these goods fall into a general realm of "consumers' goods" there becomes little to differentiate them in the sense that their existence serves a homogeneous end – the servicing of consumers' want satisfaction. Just as there is little to define between one consumers' want and another, it becomes unimportant to define between one good that satisfies these wants over another. We may think back to Part I, chapter VI, where we spoke of originary interest as being caused from the realization that our wants are unlimited, but our time in which to achieve them is necessarily so. The continual desire to satisfy *a* want leaves the conclusion that although we may never say which want is valued more in the present than in the future, we may say that *a* want is valued in the present over the future. Likewise, when we look at the production of consumers' goods, we know that desires exist to be satisfied in the present (and future) but not necessarily *which* desires. The entrepreneurial incite is the ability to see these desires and meet them accordingly. Hence, the actual type of good being produced is not necessarily important, as long as it is the correct good to meet the expected need. Provided this caveat is met, there is no significant loss concerning the use of a monetary total in substitute for heterogeneous circulating capital goods.

[286] We find this is the exact point stressed by Hill (1933: 601): "A lengthening of the average production period does not necessarily mean any lengthening of the time between the first and last applications of resources (the absolute production period), but merely an increase in the proportionate expenditure of resources at the earlier states, or, in other words, an increase in capital used."

In equilibrium, circulating capital will progress along the structure from higher to lower stages at a uniform rate. Profits on these goods will be equalized all stages – that is to say, across all actions – and hence, all actions for a given time will result in an equivalent addition of value.

We may think back to the equilibrium construct formed in Part I, chapter VII. As profits have disappeared through competition to zero, there are not gains to be made by switching to one part of the structure of production over any other. Additionally, interest rates have equalized across all stages, giving a uniform increase in value. The result is that the value growth in circulating capital will be equivalent to the rate of originary interest. Actions that take the same amount of time to undertake will be rewarded equally through this originary rate.

Also note that in any equilibrium setting, there will be a lack of losses as well as profits. This implies that any planned action will obtain as per the ex ante expectation. While profits will not manifest due to the competition driving the rate of profit to zero, action will still be successful in light of these prior expectations. Losses will also not occur, however the reason will be somewhat different. Instead of being the corollary of successful actions earning a zero profit rate, there will be an absence of *unsuccessful* actions which will require a negative rate of profit. Not only will unsuccessful actions be eliminated to the point where their return will be zero, they will be eliminated to the point where the actions themselves are eliminated altogether.

We find that viewing capital value-growth in equilibrium there is no significant loss in viewing capital in monetary terms. Remember that Hayek viewed his triangle representation as an equilibrium construct,[287] and also that he used capital in a monetary sense.[288] Both of these simplifying assumptions pose no significant problem with circulating capital that is being produced in an equilibrium setting. However, as we shall see, once disequilibrium settings are assumed possible, the chance of entrepreneurial error creates problematic issues concerning capital growth.

---

287 See, for example, Hayek (1935: 40): "[W]e may conceive of this diagram not only as representing the successive stages of the production of the output of any given moment of time, but also as representing the processes of production going on simultaneously in a stationary society."

288 "It should be remembered that the relative magnitudes in the two figures are values expressed in money and not physical quantities..." (Hayek 1935: 51). This is a somewhat surprising direction for Hayek to take. After all, in the capital debates he partook in with Frank Knight earlier, Hayek's stress laid on capital as a physical good (heterogeneous by definition), whereby Knight defended capital as a homogeneous fund in monetary terms. See Boettke and Vaughn (2002).

Once we exit an equilibrium-always framework and entertain the possibility of entrepreneurial error occurring, the conception of capital must be altered drastically. For it is no longer sufficient to view capital in homogeneous terms (i.e., reckoned as a monetary aggregate). Capital in this equilibrium setting has the advantage that every element of capital is a perfect substitute for any other. Indeed, as there is no possibility for entrepreneurial error resulting in loss, or conversely, no possibility for any advantages to be made through combining capital in heretofore undiscovered ways, all capital is returning the same net yield (i.e., the originary rate of interest), and is of a homogeneous use. However, with the possibility of entrepreneurial error and success, a possibility provided in a disequilibrium setting, we must change our conception of capital drastically. As Lachmann (1947: 159) discusses the shift that must occur:

> In a homogeneous aggregate each unit is a perfect substitute for every other unit, as drops of water are in a lake. Once we abandon the notion of capital as homogeneous, we should therefore be prepared to find less substitutability and more complementarity. There now emerges at the opposite pole, a conception of capital as a structure, in which each capital good has a definite function and in which all such goods are complements.

Circulating capital now must take on a role decidedly different than before. As planned production may be discovered to be incorrect after progress in the production process has been made, capital complementarity comes to the forefront. As circulating capital that has not yet completed its production process is discovered to be erroneously undertaken, it must be shifted to its next best use. This implies that capital must have some degree of complementarity regarding capital that was once planned to be used in one production process, as it must be found to have a new use in another line – *it must have an element of complementariness to have use in another production process*.

Capital as a monetary aggregate may be suitable for use in an equilibrium construct excluding error from occurring, however, once an allowance is made for this negative eventuality, capital's heterogeneous features (including complementarity) must be accounted for.

As an example, we may assume an economy in equilibrium producing one good – an automobile. As the equilibrium defines that the circulating capital goods are all demanded equally, they are all perfect substitutes for one another. Now if we introduce a shock that breaks the equilibrium, we may see the necessity for a capital defined in different terms. If we assume a shift in consumers' demands, such that automobiles are demanded less, and now motorcycles are demanded in some relatively higher degree than before, we will find that the ability of producers to mitigate the detrimental losses from this demand shift stem from the degree of capital complementariness that allows the shift in production from automobiles to motorcycles. Circulating capital that could not be complementary to a different production process would be lost. Hence, in this example, we can see that some parts may be salvaged from the erroneous automobile production, and used for production of the new motorcycles.

Once the eventuality of entrepreneurial error has been introduced, we must shift our conception of circulating capital from a homogeneous fund (i.e., one reckoned in monetary terms for example), to one that necessarily includes the concepts of complementarity and substitutability. It is only in this way that entrepreneurial errors may be rectified, and the move towards a new equilibrium re-established.

Circulating Capital at Transition to Consumers' Goods

Consumer's goods prior to becoming consumers' goods are circulating capital, as we have previously discussed. The production process implies that they must be undertaken in the past in order to undergo a process which modifies the nature of these goods in order that they may become of a value that consumers demand. The temporal element becomes important, for, as we have seen time is one of the sources of true uncertainty for the entrepreneur (the other source being the inability to repeat actions for which we know the risk factors involved). The previous section looked at production of these goods in an equilibrium setting, and identified the problems inherent when entrepreneurial error was introduced. The temporal passage almost guarantees that entrepreneurial error will obtain, as losses are incurred by some entrepreneurs and are rectified by others.

Hence, at some past time a production process has been undertaken to supply a consumers' good for some time in an unknown future. If we think back to the structure of consumption that was previously developed, we find four factors that must be forecast concerning the consumers' good, and one concerning the production process.

First is that the entrepreneur must decide what type of want satisfaction the consumers will desire at this future date. This could range from the desire for transportation, a sporting good to play in their leisure time, a food for snacking, or any multitude of desires. It may be that consumers themselves do not even know that the possibility of having this desire satiated exists, or that the desire exists in any form. Therefore it is up to the producer to create awareness of this deficiency in their knowledge set, and therefore increase their ability to discover for themselves what desires they wish fulfilled.

Second is the *degree* of want satisfaction that they will wish satisfied by a given *type* of consumers' good. We have previously seen that an individual may desire the want satisfaction of transportation to go to the park on a Sunday afternoon, but the degree of want satisfaction can be altered significantly through the type of consumers' good they use – be it public transport, a bicycle, walking, driving a Chevrolet, or being chauffeured in a Mercedes-Benz. Hence, the entrepreneur must not only correctly gauge the want that is desired, but the level, or degree, of this satisfaction that they wish to have desired as well. As there is a general tendency for more want satisfying consumers' goods to incorporate more complex production structures, there will also be a need for more complementary goods combined in order to create these goods. Their production must be planned in advance, leaving greater leeway for erroneous decisions concerning higher want satisfying consumers' goods than their lower want satisfying counterparts.

Third is the duration of serviceableness of these consumers' goods. We had seen that time preference will dictate the trade-off that will exist between renunciating more in the present (more typically referred to as paying a higher price) to obtain a longer-lasting good, or renunciating less for a good of shorter serviceableness. Hence, it may be that consumers would prefer to have goods of shorter serviceableness, and then in a future date, look towards purchasing replacement goods as the need arises. As an example, a consumer may purchase a suit from *Georgio Armani* of sufficiently high quality that it will be expected to deliver serviceableness over an extended period of several years. Alternatively, they could purchase a suit of lower longevity from *H&M* which may only last in a pleasing condition for one season. A plethora of options exist in the middle ground between the two, and the degree of time preference an individual has will dictate which option is desired more highly. The entrepreneur must foresee this desire, and adjust their production process accordingly.

Fourth, and lastly directly concerning the good produced, the entrepreneur must decide if the consumers' good desired is of a durable or non-durable nature. Some goods will be desired to have a degree of durability that implies they will require little allowance for depreciation in the future, while others will desired that will depreciate quickly and be replaced. Again, the trade-off between the present amount of renunciation and and amount of expected future renunciations required to continue the serviceableness of the good in question. As a general tendency, goods that depreciate less, or that require less continual allowance for depreciation, will also require an increased amount of renunciation upon purchase. A Toyota may be essentially the same auto as a Fiat, however, one which will require continual maintenance throughout its useful life to retain value, while the other will continue functioning for a much longer period without the same issues arising. Again, the time preference trade-off of the individual will primarily dictate to what extent the choice between these extremes is made.

These four factors must be foreseen by the entrepreneur and accounted for correctly in advance. They concern the specific type of good, or want that is desired to be satiated. Failure to correctly foresee any one of these factors will arouse the need for circulating capital to be used in a different avenue as the consumers' good it is trying to create may change in nature before the good is complete.

There is one additional consideration that the entrepreneur must decide prior to undertaking the production process. That concerns the quantity that will be demanded by the consumers when the final good is complete. This will affect the amount of capital that the producer must utilize in the production process, as well as the amount of output that must be created along the way. An overestimation of the quantity that will be demanded will flood the market with goods, therefore placing downward pressure on prices even if the good is *exactly* what is demanded in quality by the consumers. An under-estimation of quantity demanded will place upward pressure on the final selling prices, but this may not be enough to offset the profit that could be made through additional sales if the price were lower (provided the quantity was available to justify this lower price).

These five factors work conjointly to determine that the quantity *and* quality of consumers' goods is consistent with the desires of consumers concerning the satisfaction of their wants. In equilibrium, all five of these factors must be in-line with the desires of consumers to ensure there are no excess, nor insufficient, quantities or qualities demanded or supplied.

Conclusion

The Hayekian Triangle is a wonderful heuristic device, which focuses almost exclusively on homogeneous circulating capital (reckoned as a monetary aggregate) as it progresses throughout the value adding stages of production. Used in a comparative equilibrium analysis, there is no significant difficulty with viewing capital in this way. However, once the possibility of entrepreneurial error is introduced in a disequilibrium setting, we find that capital must be viewed in its more heterogeneous reality to account for changes in production plans.

Furthermore, we have seen that entrepreneurs must forecast in advance five factors concerning the consumers' good in production, and the demand that derives thereof. The exact want that is projected to need satiation, the degree to which this want is desired to be satiated, the duration of serviceableness the consumers' good will provide for and the degree of depreciation and allowance for depreciated value all factor to require forecasting by the entrepreneur in advance so that the production process may work towards creating these goods. An ancillary consideration that must be foreseen in advance is the quantity of the consumers' good in question that will be desired. Taken together, these five factors all determine whether an equilibrium will be reached once the circulating goods are transformed into consumers' goods, or whether dis-equilibrating forces, with their resultant profit and loss opportunities, will result from the time-consuming process.

2. Fixed Capital

If we wish to remain within the framework provided by a Hayekian triangle to analyze changes in the structures of consumer and capital goods, some considerable additions will have to be made. Up to this point, we have off-set the traditional production oriented structure with its partner focusing on consumption goods (thus returning the analysis to its Jevonsian roots). The previous section set forth the conventional triangle, with its focus on circulating capital goods as they undergo the time-consuming process to be transformed into consumption goods. Furthermore, we have seen how changes in the structure of consumption (reflective of changing consumers' preferences) evokes changes in the structure of *circulating* capital as well. As entrepreneurs adjust their existing production plans to account for these shifts in consumers' tastes, the flow and uses of capital goods must also undergo an offsetting shift.

Unlike circulating capital goods, fixed capital is distributed in a much more complex process as it is produced. Circulating capital has the distinction of proceeding in a solely unidirectional process along the time-line of production. As value is added in each stage, circulating capital progresses onward to the point where it alters its nature to cease being considered a capital good, and enters the realm of available consumers' goods. Circulating capital may only regress in the production structure through entrepreneurial error. That is to say, if circulating capital progresses to a point where it is realized that entrepreneurs will not value it fully as a consumers' good in its expected future state, or, if an entrepreneur realizes that the profit rate will be greater if a change is applied to an existing circulating capital good, it may return to a prior stage of the production process and incur further value-adding steps. We find that in equilibrium, circulating capital will, as a result, be forever progressing through the productive structure, never regressing. However, owing the possibility of entrepreneurial misjudgment, in disequilibrium circulating capital may very well regress *although this will not be the norm*.

Fixed capital serves the unique role in that it too progresses forward through the structure of production, however, once complete, never progresses past the realm of capital. All circulating goods progress forward to become consumers' goods (provided they were not created incorrectly through entrepreneurial error), however, fixed capital are the produced means of production that return to the capital structure once complete. Hence, upon completion, these goods change direction and progress in what seems to be a backward sequence through the capital structure.[289]

Heterogeneity of Fixed Capital

In equilibrium there was no significant setback to viewing circulating capital as a homogeneous fund, reckoned in monetary terms á la Hayek. One of the main contributing reasons for this was that the goal of this production was homogeneous in nature – want satisfaction. Fixed capital has a somewhat different characteristic as it serves to produce significantly different goods. The assumption of homogeneity has analytical appeal when dealing with circulating capital as we can *imagine* our wants being satisfied with only one homogeneous consumer good. However, such a a level of abstraction is impossible when dealing with fixed capital. The reason being is that fixed capital exists to create circulating capital, which eventually are used as consumption goods. If fixed capital were homogeneous, two eventualities would occur.

The first is that there would no longer be any need to make a distinction between fixed and circulating capital. If all fixed capital were identical, interchangeable, and perfectly substitutable, the product of combinations made from them would also be. Hence, any distinction we could make between fixed and circulating capital disappears if we assume homogeneity.

Second that if fixed capital were homogeneous, and even if we could make a distinction between it and circulating capital, there could be no benefit from combining different capital units together. If fixed capital existed in a fully substitutable form, then the only advantage of combining x amount of fixed capital A with y amount of fixed capital B would be to get an amount $(x + y)$ of fixed capital (either of type A or B, it wouldn't matter as they are perfect substitutes). Hence, the only advantage of combining fixed capital of perfect homogeneity would be to create differing quantities thereof, which is far from what we are looking for when speaking of the production process.

Fixed capital must exist as a heterogeneous order of goods, one which will necessarily be better suited to be viewed in terms of individual goods. It is easy to see that fixed capital exists as separate machines, tools, books, etc, that all serve to work together in differing degrees of complementarity to

289 Nurske (1935: 236n1) demonstrates diagrammatically that fixed capital differs from circulating capital in the direction of its services. Nurske errs in thinking that fixed capital may be produced at any point on the productive structure, and then be redistributed either forward or backward. His confusion lies in thinking that any capital good may be considered complete at any given point on the structure. At this point it is helpful to return to the concept of stages of production, as stressed by Garrison (2001). If we take Garrison's final stage – retail – to imply the selling process of any produced good, we see that it applies equally to circulating capital as it does to fixed capital. In fact, the market for fixed capital (commonly referred to as the B2B market) is in many cases many times larger than that available for consumers' goods.

create a vast array of separate goods of circulating capital. Hence, the fixed capital structure is a series of interrelated goods, each fixed together to continue the production process further.

Equilibrium and Fixed Capital

In equilibrium, there are two cases we may think of. The first is where fixed capital growth is held steady, thus economic output is not increasing. The alternative is that due to continual investment, fixed capital grows beyond its minimum for depreciation allowance, and hence, over time economic activity increases. Each must be looked at in turn.

We may observe that all capital is of a finite life, and also that all exhibit a tendency to depreciate and lose value over time. Hence, some degree of capital replenishment is continually needed if the existing capital stock is to continue producing at a steady level of output. In this case, then, we notice that fixed capital *must* grow at a rate that is equivalent to its depreciation. For example, if machine A depreciates at *x*% per year, than production of machines A must each year add *x%* to the stock of existing fixed capital. In this way net depreciation is zero, and productive capacity stays level.

This implies that of total productive capacity along the structure of production, in equilibrium, two portions of production will serve two ends. The first is that the production of circulating capital will fulfill consumers' needs and replenish the depreciation of consumer durables, as well as services of zero duration. Second, we find that productive capacity in equilibrium will remain steady, giving rise to the need that fixed capital production must be maintained along the structure of production. Hence, a portion of capital production must be dedicated to a fixed capital depreciation allowance. The corollary to both these points, equilibrium between consumption wants and consumers' goods produced and production capacity needed and continual replenishment of the capital stock, implies that any consumption-production equilibrium must have an off-setting equilibrium between the savings and investment rates. Indeed, as Myrdal (1939: 22-23) notes the equilibrium relationships as:

> One has to divide income into *saving* and *consumption demand*, and similarly production into *investments* of real capital and *production of consumption goods*. In the combination of those four quantities arises Wicksell's new statement of the problem of monetary theory. The underlying idea is that one cannot assume an identity between demand and supply of consumption goods except in a state of static equilibrium. This proposition should seem

> obvious to the unsophisticated mind, since decisions to buy and sell a commodity are made by different individuals. Similarly, one cannot assume that capital (investment) demand and capital (saving) supple are identically equal; for they, too, originate with non-identical groups of individuals. To treat supply and demand in these cases as being *identically*, rather than *conditionally* equal, would involve a highly unreal and abstract concept of equilibrium.[290]

Hence, any equilibrium must satisfy both these conditions. However, as Myrdal correctly points out, the separation of savers and investors gives little chance that equilibrium will obtain. In the *aggregate* it may be noted that investment may not exceed *real* savings, hence, an equilibrium between the two will be quite likely to obtain in any sort of aggregate setting. The problem with the equilibrium framework is dependent on producers fully matching the consumers' needs. For not only will production of circulating capital need to be perfectly foreseen to account for production of consumers' goods, but this in turn will necessarily require fixed capital production to be correctly imputed and produced to reach any sort of equilibrium, or even a proximal equilibrium.

The interest rate therefore becomes crucial to keeping the productive activities consistent with the consumption needs of the consumers. We had previously seen that the interest rate we see manifest on the market stems from originary interest directly dictated by the time preference scales of individuals. Therefore, the trade-off between consumption and saving on the structure of consumption *dictates* the rate of originary interest that will reverberate throughout not only this same structure (of consumption), but also be transmitted throughout the whole of the structure of production.

Producers receive the savings of the consumers, thus creating a quasi-equilibrium between the savings and investment that is necessary to sustain the productive activities of the economy. However, the complicating factor is that entrepreneurs must then correctly foresee and produce according to what consumers *future* consumption patterns are expected to obtain by the time the production process is complete. As was noted earlier, this entails five factors: 1) the exact want that is projected to need satiation, 2) the degree to which this want is desired to be satiated, 3) the duration of serviceableness the consumers' good will provide for and the degree of depreciation and 4) allowance for depreciated value all factor to require forecasting by the entrepreneur in advance so that the production process may

[290] We may actually see this concept – equilibration between ex ante savings and ex post investment – was alluded to prior to Myrdal in Rosenstein-Rodan (1936: 273-274) where he notes the two sides of equilibrium. The first is the equilibration of income – between savings and demand for consumers' goods. The second is the equilibration of production – between the investment in fixed capital, and the production of consumers' goods (i.e., circulating capital). Hence, any interest rate functioning to aid in this process must equilibrate simultaneously the two separate realms – return on investment, and homogenize savings with the rate of consumption.

work towards creating these goods. And, additionally, the expected quantity of a good demanded must be forecast correctly. Even if an equilibrium obtains between savings and investment, there is still ample opportunity for disequilibrium from any of the aforementioned five reasons. To the degree that entrepreneurs can correctly foresee and produce according to these expected conditions, the productive process will continue with no significant disequilibrium disruptions, endogenously produced through entrepreneurial error.

Interest Charges and Fixed Capital

Much literature concerning Austrian Business Cycle Theory to date has concerned the manipulation of the market rate of interest to become a non-faithful representation of the underlying time preferences of individuals. While it is outside the scope of the present work (and not its purpose) to look into subsequent consequences of a non-market generated alteration in this rate, it will prove instructive to look at what will obtain on the structures of production and consumption owing interest rate changes (market generated or not).[291]

We have already looked at changes in time preference rate regarding changes to the structure of consumption, which may be summarized below in table 1.

Time Preference	Effect on Structure of Production
Increase	Shorten duration of serviceableness
Increase	Increase present want satisfaction
Increase	Decrease degree of durability of goods
Decrease	Increase duration of serviceableness
Decrease	Decrease present want satisfaction
Decrease	Increase degree of durability of goods
Table 1	

When analyzing changes in the structure of production, however, we must treat circulating and fixed capital separately. The general neglect of fixed capital in the framework of the Hayekian triangle has generally been ignored, with the consequence that some not wholly correct conclusions have been drawn.[292] As early as 1935, Smithies had worked within the Hayekian framework to add fixed capital to

[291] Interested readers may read Hayek (1935), Mises (1953), Rothbard (1962; 1963), Garrison (2001), or Huerta de Soto (2006) for looks at interest rate manipulation and the business cycle.

[292] Hayek (1935: xi-xii) realized in the preface to the second edition of his *Prices and Production* the problem this neglect

the analysis. Although giving new insights into effects of interest rate changes on capital productivity and output, he (1935: 121-122) concludes that: "The consideration of durable instruments, then, does not in any way impair the formal validity of the Austrian analysis for the purposes for which it was originally designed." In fact, following Machlup's (1932) analysis, we see that alterations on the interest rate have negligible effects on existing production processes. The reason is that the costs for these processes have already been sunk, therefore that any change in the interest rate could only apply to that portion of cost which has not yet been paid for, and also has not been subject to a locked-in repayment rate. Furthermore, Machlup (1935) reckons that interest rate reductions, in this example, will not stimulate production of fixed capital. This stems from the fact that the effect is quite negligable on the costs to be incurred compared to that on labor or circulating capital. Hence, interest rate changes do little to change the *existing* structure of production, however there are two avenues where significant changes do occur.

The first is in the production of circulating capital, which will now be increased in quantity due to an increased expectation of demand. Decreases in the rate of interest under normal conditions (caused by a decrease in time preference) imply a decrease in consumption spending. This signals to producers that one or any combination of three adjustments must be done. First is a reduction in consumers' goods production. A decline in the production of circulating capital must result. Second, an increase in the production of fixed capital must be undertaken. This is necessary as the decrease in time preference has signaled that consumers are demanding either: 1) goods with greater duration of serviceableness, or 2) goods of greater durability. As was previously discussed, these two options may generally only be undertaken through a greater use of fixed capital. Hence, any reduction in the rate of interest signals to producers that a decrease in production of consumers' goods is required, with an offsetting shift into the production of fixed capital to create longer lasting, and more durable consumers' goods.

The second is an increase in *new* projects as the lower interest rate makes the minimal profitability of these projects much easier to obtain. Hence, as the shift towards an increase in production of fixed capital commences, projects that were marginal before now become more attractive

---

wrought:

> It is impossible to assume that the potential services, embodied in a durable good and waiting for the moment when they will be utilised, change hands at regular intervals of time. This meant that so far as that particular illustration of the monetary mechanism was concerned I had to leave durable goods simply out of account. I did not feel that this was too serious a defect, particularly as I was under the—I think not unjustified—impression that the rôle which circulating capital played was rather neglected and accordingly wanted to stress it as compared with that of fixed capital.

as the "hurdle-rate" of profitability given by the cost of borrowing funds to finance them has been reduced. Hence, as projects are undertaken which yield fixed capital, those with a lower degree of profitability will be available to be attractively financed then before the reduction in the borrowing rate occurred. However, also notice that projects in the higher stages see a relative increase in their profitability due to the length of capitalization involved. As these projects are financed over many years, the reduction in the interest rate gives rise to a greater reduction in the cost of borrowing as the life-time of this financing options will be longer. Hence, a decrease in the cost of borrowing (as reflected in the rate of interest) will result in more projects to produce fixed capital undertaken (ones which were considered only submarginal before) and also that a relative shift to the higher stages of production will occur simultaneously.

Hence, we see that changes in the interest rate need not effect circulating capital or existing projects that have been paid for in any significant way. Instead, the bulk of the alterations to the structure of production will occur in those areas producing fixed capital, particularly in the higher stages more sensitive to interest rate fluctuations.

Conclusion

We have seen that fixed capital must necessarily be viewed in terms of heterogeneous capital goods – it is conceptually *impossible* to think of this order of goods as being of a homogeneous fund. Additionally, equilibrium will obtain on this structure of production when savings are matched to investment, *and* production is undertaken given two caveats. The first is that circulating capital is produced at the proper proportion to fixed capital to satisfy the wants of consumers. Second is that fixed capital is produced to such a degree that it perfectly offsets its own rate of depreciation, thus keeping the productive capacity of the economy steady.

Alterations in the interest rate have a more significant effect on fixed capital than on either consumers' goods or circulating capital. The reason is that fixed capital is required relatively higher in the stages of production, thus in regions more sensitive to changes in the rate of interest for financing. Decreases in the rate of interest signal to producers that two alterations to the structure are necessary. The first is that consumers are now demanding fewer consumers' goods, which requires a decrease in the circulating capital along the structure. Second is that an offsetting increase in fixed capital becomes necessary. This is so as while consumers are saving more, they are requiring goods of expected greater duration of serviceableness or durability. Both of these factors require that a relative increase in fixed capital be undertaken in order to produce the longer-lasting, greater durability goods that consumers are now demanding.

However, at the same time that this build-up in fixed capital is undertaken, a simultaneous increase occurs in the allowance for depreciation on these goods. This depreciation creates difficulties of its own, and must be looked at separately within the structure of production.

3. Depreciated Capital Allowance

All goods depreciate due to *use*. This finite life-span gives rise to the need for continual allowance to be made for depreciation in these assets. Therefore, as was discussed previously, in equilibrium, an amount of fixed capital will be required to be produced in order to offset depreciation on existing capital and maintain production at a steady level. Two factors arise from the need to replenish depreciated assets. The first is that the interest rate prevailing will provide incentive and direction to entrepreneurs as to what degree of durability of fixed capital should be produced. The second factor results from this durability of capital, and concerns the amount of fixed capital production that must be undertaken to replenish this depreciating stock. Both these factors work together to determine the minimum amount of fixed capital production that must be undertaken to spur growth in the economy.

Durability and Fixed Capital

Just as a direct trade-off will exist between durable consumers' goods and the cost of purchasing these goods, a trade-off exists between fixed capital durability and its cost to produce, and hence its general purchase price. At low interest rates, it becomes relatively more cost effective to obtain fixed capital through financing agreements, instead of outright purchases. Low interest rates make it possible to finance over a short time span only a portion of the cost of the greater asset, thus incurring a relatively lower cost than would result from purchasing the asset in full. The low interest rate originates with individuals having a low degree of time preference (saving more and consuming less). Hence, these situations characterized by individuals demanding fewer consumer goods in the present result in fixed capital being financed at these lower rates instead of purchased in full.

The result of this tendency is that fixed capital will be produced with a lower durability, a result of the shorter time period it will be demanded for before being replaced. Hence, as these goods are expected to be financed for a short period at a low interest rate, the expected duration of serviceableness is decreased accordingly. The lower interest rate originates with consumers' consuming less relative to their savings, and entices producers to utilize equipment of less durability.

Conversely, we see that at times of relatively higher interest rates, producers are enticed to fully capitalize fixed capital by purchasing it in full. A desire for more durable capital will prevail in this case, as businesses wish to utilize these goods for as long as possible before their useful life has been depreciated away.

Just as in consumers' durables we noted that the move towards greater durability would incur greater capital utilization and hence, an increased cost of production, we see the same result with durable fixed capital goods. A secondary effect of capital build-up occurs, then, whereby the lower interest rates not only increase the amount of fixed capital within the structure of production as producers shift from circulating to fixed capital accumulation. The additional build-up comes from the fact that low interest rates also entice lower durability fixed capital to be utilized in production. As it becomes more cost-effective for firms to continue rolling over capital into newly financed capital, the need for high duration of serviceableness decreases. As a direct result, much production must be dedicated to producing fixed capital to offset the depreciation of this quickly obsolete fixed capital that is being continually rolled over into new fixed capital. Hence, not only does a reduction in the interest rate, *ceteris paribus*, cause a decrease in the general durability of fixed capital, the structure of production must now dedicate more productive capacity towards replenishing this more quickly depreciating capital stock owing to this occurrence.

A reduction in the interest rate causes a primary, and a secondary, increase in production of fixed capital. The primary increase stems from the relative shift that occurs from circulating capital into the production of fixed capital. Simultaneously, the same interest rate which entices this shift as it is lowered relative to its previous level also causes the general durability of fixed capital to decrease. This occurs as the option to finance at a lower rate of interest becomes more attractive relative to purchasing fixed capital outright, and gives rise to an increased use of this option. As financing will generally be for a shorter time period than the alternative option – outright purchase of fixed capital assets – the desired durability of these assets decreases. As their useful life will be reduced, a need arises for a greater allowance to maintain them, supplied through the new production of fixed capital. Hence, a secondary increase in production occurs directly as a result of the need to continue replenishing the lost, or depreciated fixed capital assets, originally created due to the reduced interest rate.

4. Conclusion

In effect, we have seen that not one singular structure of production exists, but three intertwined structures all fixed together to create a balance with the structure of consumption.

Typically, the structure of production has focused on circulating capital – that capital destined to be transformed to consumers' goods. Although this is a very important part of the structure, it is only *a* part of it. Production of this capital has five important caveats that entrepreneurs must foresee and meet. The first is that the actual want that consumers will desire satiated in the future must be foreseen in advance, so that an appropriate good to satisfy this want may be produced. Second, the degree of want satisfaction must be assessed. Consumers will desire their wants satiated in differing degrees. In particular, we saw that the desire to have transportation to go to the park on a Sunday can be satisfied by many different consumers' goods – bicycle, metro, motorcycle, Fiat automobile, or a more luxurious Mercedes-Benz – or even that the want could be satisfied with no use of an actual good, only with the walking power of an individual. Third, the duration of serviceability of these consumers' goods must be forecast. To the extent that a trade-off exists between renunciation of goods in the present (or the price to attain something), and the duration of serviceability of a good, we see that an individual's time preference determines this choice. Fourth, the degree of maintenance on a consumers' good will also depend on their time preference scale. All goods depreciate through use, requiring some allowance for maintenance so that their value may be maintained. Higher quality goods that require less allowance for depreciation will incur a higher initial cost at purchase, thus creating a trade-off between the two opposite possibilities. All four of these possibilities concern the actual good to be produced. As the production process that creates them will differ accordingly depending on what the consumers' desire, and given that the production process is a timely progression of steps, correct forecasting of these four factors must be made well in advance of the actual consumer attaining the good. There is one final consideration that producers must factor for and that concerns the quantity of consumers' goods to be produced. Production processes must have capacity to allow for the full delivery of these goods when (or *if*) demanded, and hence, the structure of production must be aligned so as to allow for the proper quantity of consumers' goods to be produced (via circulating capital).

The neglect for fixed capital has proven a problematic missing of the typical Hayekian analysis of the structure of production. Fixed capital involves a large amount of the total productive capacity of the economy, supplementing the more noticeable circulating capital which has hereto now taken up so much of the discussion. Newly produced fixed capital is the type of good most susceptible to alterations in the rate of interest. As interest rates experience relative declines the impetus has been the shift from consumption to saving by consumers. Hence, while consumers' are signaling they require fewer consumers' goods, producers use this signal to realize they should produce less circulating capital. To off-set this shift, they commence producing greater amounts of fixed capital. This occurs as consumers begin demanding goods of higher durability, or of lower maintenance requirements. These goods which we typically refer to as "higher quality" are produced through more capital intensive production processes, necessitating the need for increased production of fixed capital.

The final part of the structure of production is that component which is itself a part of the total amount of fixed capital, that which is dedicated to the allowance for depreciated fixed capital. As all goods have a finite life concerning their ability to produce value, we see that they all also must continually lose a certain amount of value. This gives rise to the requirement to produce an amount of fixed capital solely for the purposes of maintaining production at a given level. A reduction in the rate of interest has the effect that more fixed capital is produced. A secondary effect is that less durable fixed capital is also produced. The reduction in interest charges entices a movement towards purchasing capital to capitalize over an extended period of time, to financing capital at lower interest rates for shorter periods of time. As the capital is utilized with the expectation that it will only be used for a short duration, the demands on producers are such that capital of lower durability is correspondingly produced. Hence, two effects of lower interest rates manifest. First the increase in production of fixed capital, and second, the increase in production of fixed capital to allow for depreciation of this increased utilization of fixed capital.

All three components work together to comprise the structure of production, which is directed towards an equilibrium by the interest rate. In particular, the equilibrium between savings in the structure of consumption, and investment on the structure of production, is necessary to avoid over-production or over-consumption. Secondary to this, an equilibrium must obtain between the desires that consumers want met, and the goods producers supply to meet this end. As a temporal lag will exist between the two – production must be commenced today to yield consumers' goods at some future date, we see that a great possibility for entrepreneurial error exists. To the degree that the rate of interest prevailing is indicative of the underlying time preferences of savers (and hence, an accurate approximation of the consumers' goods they will demand), the trend towards an inter-temporal co-ordination between producers and consumers will tend to obtain over time.

VII. LENGTHENING THE STRUCTURE OF PRODUCTION

One of the main factors effecting value growth of circulating capital is an increase in the length of the structure of production. This is a somewhat "fuzzy" concept, which poses problems depending on how exactly the structure is thought to be "lengthened."

The first way we can think of this process is through a temporal lengthening of the structure. That is to say, the time period between when a productive process commences and when it ends is extended, and this is thought to lead to higher value growth. Although this may make intuitive sense – the longer something takes to produce the more value it must have – it misses the crux of the issue as to why a temporal lengthening has occurred in the first place.

Alternatively, the lengthening of the structure may not be a temporal process, but a physical one. As a relative shift in production between circulating and fixed capital goods occurs, the structure is lengthened in the sense that more production is dedicated to production of the "produced means of production." Hence, the lengthening that occurs in the structure need not necessarily be of an absolute temporal dimension, a relative shift between the two types of capital is sufficient.

The use of Hayekian triangles to diagrammatically render the productive structure has again led to a flawed outlook concerning this lengthening process. By reassessing this situation, we can better determine what this process implies, and how it properly functions.

1. Temporally Lengthening the Structure of Production

Owing the fact that a decrease in interest rates makes *new* investment in the higher stages of production more profitable, it is commonly assumed that this same shift in interest rates causes the structure of production to lengthen. Hence, while at t_0 an interest rate of 10% may cause the equilibrium temporal length of production to be a theoretical 10 years, if at t_1 this rate of interest is reduced to 5%, the new equilibrium length of production is now assumed to be increased, for example to a theoretical 12 years. Reductions in the interest rate are assumed *ceteris paribus* to cause corresponding increases in the temporal length of the production sequence.

There are several flaws with this that need be addressed. First is the concept which Böhm-Bawerk had tried to define of the "average period of production." For if there is an actual lengthening in the period of production, it follows that there must be some sort of definable period with which we may compare this temporal lengthening. The second point is that this assumes that longer production processes are necessarily more productive. Although this may find support in some cases, there are problematic exceptions.

By reassessing both of these points we find that the concept of a temporally lengthened period of production need not be indicative of more productive processes. Instead, when we speak of a lengthening of the structure of production we must look elsewhere.

The Average Period of Production

Böhm-Bawerk originally posited that "roundabout" production processes would always lead to greater productivity. We can see that owing our innate time preference, individuals will always prefer a want satisfied sooner rather than later. Production processes that take a longer production process would therefore only be chosen if they were expected to yield a greater amount of want satisfaction upon completion than a shorter process. However, a great deal of bifurcation occurred as Böhm-Bawerk tried to further explain what exactly he meant by these roundabout production processes. Indeed, even though the concept has been granted general acceptance by the community of economists (see, for example, Samuelson 1980: 559) this acceptance may not be based on valid grounds. As Northrup Buechner (1989) explains, Böhm-Bawerk's real definition of roundaboutness has been misinterpreted primarily through his own problems in explicating what he exactly meant.

Böhm-Bawerk first lays out his theory in terms of to what degree production can be regarded as being aided by capital as the means for greater productivity. Hence, he (1889: 120) states: "The more capitalistic production is, the smaller the proportion of total original resources available in a given year whose product is consumed in that year." This seems to clearly imply that the definition for roundaboutness is in terms of capitalistic production process, and *not* directly related to time. He also, however, bifurcates between these two definitions – time and capitalistic processes – and refers to both occurring simultaneously.[293] Finally, in his final volume of *Capital and Interest*, he notes that while these roundabout processes are generally more time consuming, this need not be the necessary relationship. More roundaboutness does not translate into more time in all cases, in fact, he goes so far as to conclude: "Exceptional cases may occur in which a roundabout method is not only *better* but *quicker*" (1921: 82).

If we wish to speak of any lengthening of the structure of production, it becomes clear that it must not be a lengthening in terms of its temporal element. This troubling feature of the structure of production has reappeared in numerous articles and books, with detrimental repercussions as a result. Hayek (1941: 69-70) seems to refer to both types of lengthening in the following passage:

> [Economists] should be able to speak of changes in *the* 'period of production', or *the* 'length of the process as a whole', as a short way of referring to changes in the investment periods of the various factors used. In fact, however, most of the changes in productive technique are likely to involve changes in the investment periods of different units of input to a different degree and perhaps in different directions. This raises all kinds of difficulties which we shall have to consider later. In particular it makes it impossible to use the terms 'changes in investment periods' and 'changes in the length of the process' of 'changes in *the* period of production' synonymously.

It becomes clear that Hayek, deriving from his Böhm-Bawerkian foundation, uses the concept of lengthening the structure of production temporally as he applies the concept with his triangles.[294] More modern Austrians have been equally unclear in the aggregate as a community as to where they stand on the issue. Skousen (1990: 23) interprets Böhm-Bawerk to be reckoning a lengthening of the structure in

[293] For example, Böhm-Bawerk refers to the "[m]ore time-consuming roundabout ways" (1889:12), "the more time-consuming and roundabout combinations" (12) and "the lengthened roundabout method of production that consumes more time" (6). All of these definitions seem to imply that time and roundaboutness are two sides of the same coin.

[294] This is troublesome as Hayek (1936) criticizes Böhm-Bawerk's concept of roundaboutness as a temporal dimension.

terms of more capital goods, not an increase in the time needed for production. Garrison (1999), while writing a biographical sketch of Böhm-Bawerk implies that expansion of the capital structure is not an equiproportional increase of capital at each maturity class, instead it is the reallocation of capital between the maturity classes. Hence, Garrison seems to state here that the structure of production is altered by rearranging capital among the stages, not by any temporal lengthening. However, when fixed within the confines of the Hayekian triangle as an heuristic device, the structure is actually lengthened by adding new stages to production, with the result that production processes become more productive through increasingly capitalistic processes coupled with their necessary temporal element.

It becomes clear that although Böhm-Bawerk's concept of more capitalistic production as being a source for higher economic growth is well-founded, the reasoning behind this is murky. Any concept of lengthening the structure of production must center around the additional buildup of capital, not the temporal lengthening of the time needed for a consumer good to be produced. In fact, Hill (1933: 601) may have summarized this idea most succinctly:

> A lengthening of the average production period does not necessarily mean any lengthening of the time between the first and last applications of resources (the absolute production period), but merely an increase in the proportionate expenditure of resources at the earlier states, or, in other words, an increase in capital used.

Böhm-Bawerk's second and related fallacy lays in his concept of the "average period of production" (1889: 3-15, 85-88). Hence, this backward looking estimation of the length of time necessary for production was posited to be used to determine how capitalistic given production processes are. There are several flaws in this concept, primarily, the idea that the focus is shifted to the past, instead of the future where the decision is really oriented towards (Mises 1949: 488-489). Of course, Knight was the most ardent critique of this concept, using it in his debates with Böhm-Bawerk's successors (see Knight 1934: 501; 1935: 88-90; 1941: 419-420). As Huerta de Soto (2006: 298) makes clear: "[T]he first stage of production begins precisely at the moment the entrepreneur conceives of the final stage in the process (a consumer good or a capital good)." The implication then is that time is not the measure to be used when viewing the structure, but the actual production process is defined by the subjective interpretation of the expectation that an entrepreneur believes will prevail as the production process progresses. Longer production processes need not necessarily lead to increased productivity.

Paradoxically, we must realize that greater productivity will only result from the production of greater value in less time. Hence, the idea that productivity is created through the temporal passage ignores the very basis of what it means to produce something at a "greater level of productivity."

Owing the fact that we always desire our wants satisfied sooner rather than later, we know that we prefer the shortest production processes to those that are temporally longer. Hence, as Rothbard (1962: 546-537) tells us:

> The first processes to be used will be those most productive (in value and physically) *and* the shortest. No one has maintained that *all* long processes are more productive than *all* short processes. The point is, however, that all short *and* ultraproductive processes will be the first ones to be invested in and established. Given any present structure of production, a new investment will not be in a *shorter* process because the shorter, more productive process would have be chosen first.

Hence, we may conclude that any production process of longer duration that is undertaken at any moment will satisfy either of two conditions. First, it must be more productive in creating value than a process which is shorter. Second, it may be the case that the specific want that the process will satisfy has not yet been noticed, and hence, the production process will yield for the first time satisfaction of a want previously unexploited. In either of these cases, we see that production processes over time must continually shorten and maintain a constant productivity, or lengthen and become more productive. As an accumulation of capital is one method to attain higher want satisfaction in the future, we see that when we refer to lengthening the structure of production, it it not a temporal lengthening that we are primarily concerned with. Instead, the focus is on a relative shift from producing circulating capital to fixed capital.

The reason that this becomes a truism is when we see the link that savings has with investment. Higher rates of savings are channeled to investment activities that increase future productivity. This increase in savings necessarily implies a decrease in consumption – which in turn implies a decrease in production of circulating capital goods which are destined to fulfill this role. As a shift between fixed and circulating capital occurs, savings become a self-fulfilling prophecy of sorts, whereby the engine for growth is linked to the production of the capital that will increase future want satisfaction.

2. Associated Problems with Lengthening a Hayekian Triangle

The way that Hayekian triangles are rendered, it is easy to fall into the trap of thinking that lengthening the production structure involves a necessary temporal elongation and not just a shift from circulating to fixed capital. This is so as, if we take Garrison's exposition as an example, the triangle's x-axis may only be lengthened by adding additional stages of production. Hence, perhaps a new research stage is added necessarily preceding mining and distribution. Even with the structure of production as we have defined it earlier, of a sequence of value adding activities, will see its lengthening not through physical accumulation (as it should) but through the addition of more value adding activities that add a temporal extension to the structure.

One of the resultant problems with this viewpoint is the way that value growth is illustrated with a Hayekian triangle. As value growth progresses through the stages, the slope of the hypotenuse can be used to roughly approximate the changes in rates of profit and originary interest. This occurs naturally as the decrease in the rate of interest causes a temporal elongation which is reflected in a lower slope of the structure, which in turn brings us back to the lower rate of interest. However, the x-axis does not have to physically lengthen.

Typically, the angle of the slope of the hypotenuse is given as the rate of originary interest in an equilibrium economy lacking profits. Hence, as the rate of originary interest decreases, this is reflected in the altered sloped of the Hayekian triangle, which itself results from a lengthened production process. The issue is that no such temporal lengthening need occur. The "lengthening" that occurs is really a shift in the ratio between circulating and fixed capital.

Now, it becomes important to incorporate depreciated fixed capital into the analysis. As fixed capital is constructed of greater durability, there is less need for an allowance to be made to account for the depreciated portion of the capital. As a result, the ratio of circulating to fixed capital can be reduced (resulting in the more typical lengthening of the structure) *only by using more durable capital*.[295] As less new fixed capital production need be dedicated to replacing depreciated fixed capital, a constant ratio of circulating to fixed capital will result in the shift necessary between the two types to result in the lengthening of the productive structure.

Hence, what we normally have referred to as a lengthening in the structure of production may occur in one of two ways. First, we see that investment in fixed capital may be increased relative to that

[295] Nurske (1935: 238) was the first to notice this relative phenomena.

in circulating capital, with the result that the productive structure becomes more capitalistic. The second way occurs as fixed capital becomes more durable in nature. Under this scenario, it becomes less necessary for production to be dedicated to producing fixed capital to maintain output, hence, it is equivalent to a decrease in the ratio between circulating and fixed capital. Either of these two ways will result in a lengthened (i.e., more capitalistic) production structure.

3. Conclusion

Since Böhm-Bawerk originally developed the concept of roundabout production methods much misunderstanding has surrounded what exactly this concept implies. Although having wide support today, much of this support stems from an incorrect definition of the concept, with problematic results consequently gleaned.

Böhm-Bawerk himself was quite unclear in many ways on this topic, bifurcating between two definitions which he eventually declared had a general tendency to function together. The first is roundabout production through the use of more capitalistic production processes. Second, and ancillary, was the conclusion that these processes also entail a lengthening of the temporal element necessary to complete these more capital intensive processes. Böhm-Bawerk was particular enough to note these two possibilities only occurred simultaneously as a coincidence, not as a constant fact.

However, many of Böhm-Bawerk's followers have bifurcated the issue, without realizing they are doing so. Hayek, for example, in his now famous Hayekian triangles, seems to imply that any lengthening, or capital accumulation that occurs actually does result in a temporally longer production method. Garrison, following in Hayek's footsteps verbally notes the difference between more capitalistic processes and those that are only temporally longer, but refers repeatedly to the x-axis of the Hayekian triangle as a temporal-axis. Hence, for these two influential writers the emphasis is on the temporal aspect of an elongated production structure.

Instead, we have seen that changes in interest rates affect not the temporal length of the production process, but instead are reflected through relative changes between the proportions of circulating and fixed capital produced. Previously we have seen that declines in interest rates stimulate production of fixed capital, and increases in interest rates initiate the opposite effect – increases in the production of circulating capital. As a result, if we use the ratio between circulating and fixed capital as a comparative measure of this effect, decreases in the interest rate cause an increase in the ratio, with increasing interest rates having the opposite effect.

However, not only is the degree of capitalistic processes in the production structure affected by the interest rate, it can also be altered indirectly through changes in the degree of durability of fixed capital goods. As durability increases, less production is required for maintenance to allow output to maintain a sustained level. This is equivalent to a shift towards producing more fixed capital as it is now necessary to produce less fixed capital to maintain the productive structure's capacity. As less fixed capital is dedicated to maintenance on depreciated capital, the ratio between circulating and fixed capital also declines making for more capitalistic production processes.

VIII. CONCLUSION

Book I of this work analyzed how decision-making under uncertainty occurs on a micro-scale. The present book has tried to apply this to a more macro analysis. In particular, we have tried to outline the relationship that exists between consumption, and production. In the process, some new insights have been gleaned, and a more complete picture of the production (and consumption) process has been produced than has been apparent before.

Action has three categorical forms that it may fall into – consumption, production or exchange. This trichotomy is exhaustive, with no other possibilities along these lines. Consumption is that action which directly satisfied our wants and desires. Productive are those actions that aim at satisfying the ends desired – consumptive ends. A special class of action exists in between these two categories – exchange – with interesting characteristics. The first is that we may partake in exchange, but not actually produce anything from that exchange. For instance, exchange is not a necessary component of satisfying our wants, although production is necessary in all cases prior to this same outcome. Second is that some productive actions are not successful – they get us no closer to the want satisfaction in question. As a result, production is only known to be productive in the sense we wish to assign to it – productive action successfully creating consumption. We may never know when we undertake an action if it will be successful or not owing to the uncertainty of the future.

As a result of this trichotomy, we find that three types of goods also exist – production, consumption and exchange goods. Although we typically define money – the medium of exchange – as an exchange good, we see that given our trichotomy of actions, there are other goods that fall into this category. In fact, any good used with the intent that it will not be used directly to satisfy our wants could be termed as an exchange good. Furthermore, owing the ex ante difficulty (or rather impossibility) of identifying whether productive action will be successful or not, a necessity for it to belong to this category of action, the category of production goods likewise becomes difficult to identify. Of course, ex post no such difficult distinctions exist, as actions can be said with certainty to have been successful or not.

As a result of these actions and goods, we have two processes which partake in consumption and production. The first we call the structure of consumption. We have seen how this structure is effected by the time preference each individual holds. The trade-off between consumption and saving creates a feedback which reverberates throughout the structure of consumption, and eventually to its counterpart, the structure of production.

Increases in time preference – the desire to consume more in the present – has the corollary that what is desired is not only more consumption in the present, but also less renunciation for consumption in the present. Consumers' goods differ from each other depending on three characteristics. The first is the degree of want satisfaction they are capable of providing in the present. Second is the duration that this is expected to continue for. Last, we see that since all goods deteriorate over time, some amount of maintenance will be necessary to continue deriving value from any good over time; this is no different with consumers' goods. As a general tendency, goods that are more want satisfying in the present, those that have a longer duration of provided services, and those that require less allowance for depreciation of value will tend to be priced higher in the present than counterparts that do not share those aspects. As time preference increases, there will be a tendency for any one (or combination thereof) of the following results: 1) goods will be demanded in higher want satisfaction in the present, 2) goods of shorter duration will be demanded, and 3) goods of lower quality, or what we would normally refer to as goods which require a greater allowance for depreciation will be demanded. These three features together act to affect the structure of consumption through our actions and choices.

Production exists solely to satisfy the needs of consumers. We see, then, that the structure of production is inextricably linked with its consuming counterpart. As goods go through a process of action adding steps, capital goods are transformed to consumption goods that will directly fulfill our desires. The analysis up to this point has suffered a fatal set-back, as the emphasis has been placed almost exclusively on circulating capital at the expense of fixed capital and its depreciation allowance. Circulating capital is the easiest to illustrate within the confines of the Hayekian triangle, and as a result has dominated the discussion on capital structure.

Circulating capital is an important component of this intertwined structure, but it is far from the only one. Fixed capital too progresses from early stages to be transformed into those capital goods which are necessary to add to the capitalistic processes they are made from. Circulating capital will eventually become part of the consumption realm as its ownership, and role, shifts from the production side to the consumption side of the process. Fixed capital progresses through the same stages of production, but never makes the transformation that would result in its exodus from the production realm. Instead, once completed it finds itself shifted back to within the structure of production, to assist with more production processes, much like the ones it was just created through.

Production of fixed capital has five important caveats that entrepreneurs must foresee and meet.

The first is that the actual want that consumers will desire satiated in the future must be foreseen in advance, so that an appropriate good to satisfy this want may be produced. Second, the degree of want satisfaction must be assessed as consumers will desire their wants satiated in differing degrees. Third, the duration of serviceability of these consumers' goods must be forecast. To the extent that a trade-off exists between renunciation of goods in the present (or the price to attain something), and the duration of serviceability of a good, we see that an individual's time preference determines this choice. Fourth, the degree of maintenance on a consumers' good will also depend on their time preference scale. All goods depreciate through use, requiring some allowance for maintenance so that their value may be maintained. Higher quality goods that require less allowance for depreciation will incur a higher initial cost at purchase, thus creating a trade-off between the two opposite possibilities. All four of these possibilities concern the actual good to be produced. As the production process that creates them will differ accordingly depending on what the consumers' desire, and given that the production process is a timely progression of steps, correct forecasting of these four factors must be made well in advance of the actual consumer attaining the good. There is one final consideration that producers must factor for and that concerns the quantity of consumers' goods to be produced. Production processes must have capacity to allow for the full delivery of these goods when (or *if*) demanded, and hence, the structure of production must be aligned so as to allow for the proper quantity of consumers' goods to be produced (via circulating capital).

As the structure of production becomes more capitalistic, that is to say, more fixed capital is employed relative to circulating capital, a greater need for fixed capital production is required to offset the depreciation of the existing capital stock (*ceteris paribus*). Hence, for a greater amount of capitalistic production to be maintained, more allowance must be made for this depreciated capital stock.

Lastly we looked at what is commonly referred to as lengthening the structure of production. Much confusion abounds this topic, with a bifurcation between whether this means a lengthening of the amount of time necessary to complete a production process, or whether it means the accumulation of a greater degree of fixed capital in the structure. An increase in productivity can only result from the ratio between circulating capital and fixed capital production being decreased so as to experience a relative increase in the productive stock of fixed capital within the structure. This process is difficult to illustrate using the conventional Haykeian triangles, which give the appearance of a temporal lengthening of the structure, at the expense of a loss of understanding just how the ratio of capital *within* the structure is altered.

With a clear elaboration of how the two structures co-exist, we may now look deeper into where value stems from within this structure. It may be noticed that the structure of production has been illustrated up to this point as a series of interlocking actions. No consideration has been made for institutional concepts such as firms. In the next and final book, we will delve into the theory of the firm, and determine where value comes from as these organizations function to provide the want satisfaction desired by consumers.

REFERENCES

Arena, R. and S. Gloria-Palermo. 2008. Menger and Walras on Money: A Comparative View. *History of Political Economy* 40(2): 317-343.

Barnett II, William, and Walter Block. 2005. Money: Capital Good, Consumers' Good, or (Media of) Exchange Good? *The Review of Austrian Economics* 18(2): 179-194.

Barnett II, William, and Walter Bock. 2006. On Hayekian Triangles. *Procesos de Mercado: Revista Europea de Economía Política* 3(2): 39-142.

Barnett II, William, and Walter Bock. 2007. Saving and Investment: A Praxeological Approach. *New Perspectives on Political Economy* 3(2): 129-138.

Boettke, Peter J., and Vaughn, Karen. 2002. Knight and the Austrians on Capital and the Problems of Socialism. *History of Political Economy* 34(1): 153–175.

Böhm-Bawerk, Eugen von. [1884] 1959. *Capital and Interest. Volume I: History and Critique of Interest Theories*. South Holland, IL: Libertarian Press.

Böhm-Bawerk, Eugen von. [1889] 1959. *Capital and Interest. Volume II: Positive Theory of Capital*. South Holland, IL: Libertarian Press.

Böhm-Bawerk, Eugen von. 1906. Capital and Interest Once More I: Capital vs. Capital Goods. *Quarterly Journal of Economics* 21(1): 1-21.

Böhm-Bawerk, Eugen von. Capital and Interest, Vol. III: *Further Essays on Capital and Interest,* (trans.) Hans F. Sennholz, Capital and Interest. South Holland, IL: Libertarian Press.

Clark, John Bates. 1888. Capital and Its Earnings. *American Economic Association Monograph* 3(2). Baltimore: American Economic Association.

Foldvary, Fred E. 2006. The Complete Taxonomy of the Factors: Natural Resources, Human Action, and Capital Goods. *American Journal of Economics and Sociology* 65(3): 787-802.

Garrison, Roger W. 1978. Austrian Macroeconomics: A Diagrammatical Exposition, in *New Directions in Austrian Economics,* Louis M. Spadaro, (ed.). Kansas City: Sheed Andrews and McMeel.

Garrison, Roger W. 1995. Linking the Keynesian Cross and the Production Possibilities Frontier. *Journal of Economic Education* 26(2): 122-130.

Garrison, Roger W. 1999. Eugene von Böhm-Bawerk: Capital, Interest, and Time. In (*ed*.) Randal G. Holcombe, *Fifteen Great Austrian Economists*. Auburn, AL: Ludwig von Mises Institute.

Garrison, Roger W. 2001. *Time and Money: The Macroeconomics of Capital Structure*. London: Routledge.

Garrison, Roger W. 2004. Overconsumption and Forced Saving in the Mises-Hayek Theory of the Business Cycle. *History of Political Economy* 36(2): 323-349.

Gilbert, J. C. 1953. The Demand for Money: The Development of an Economic Concept. *Journal of Political Economy* 61(2): 144-169.

Hayek, Friedrich A. [1928] 1975. *Monetary Theory, and Investment*. New York: Augustus M. Kelley.

Hayek, Friedrich A. [1935] 1967. *Prices and Production, 2nd Ed.*. New York: Augustus M. Kelley.

Hayek, Friedrich A. 1936. The Mythology of Capital. *The Quarterly Journal of Economics* 50(2): 199-228.

Hayek, Friedrich A. [1939] 1975. *Profits, Interest, and Investment*. New York: Augustus M. Kelley.

Hayek, Friedrich A. 1941. *The Pure Theory of Capital*. Chicago: The University of Chicago Press.

Hayek, Friedrich. A. 1976. *The Denationalisation of Money*. London: Institute of Economic Affairs.

Herbener, Jeffrey. M. 2002. After the Age of Inflation: Austrian Proposals for Monetary Reform. *The Quarterly Journal of Austrian Economics* 5(4): 5-19.

Hill, Martin. 1933. The Period of Production and Industrial Fluctuations. *The Economic Journal* 43(172): 599-610.

Holcombe, Randall G. 2001. The Two Contributions of Garrison's *Time and Money*. *The Quarterly Journal of Austrian Economics* 4(3): 45-50

Horowitz, Steven. 2000. *Microfoundations and Macroeconomics*. New York: Routledge.

Howden, David. forthcoming. Money an a Medium of Exchange and Its Evolution: An Elaboration on Mengerian Monetary Economics. *Procesos de Mercado: Revista Europea de Economía Política*.

Huerta de Soto, Jesús. [1992] 2005. *Socialismo, Cálculo Económico, y Función Empresarial*. Madrid: Unión Editorial.

Huerta de Soto, Jesús. 2006. *Money, Bank Credit and Economic Cycles*. Auburn, AL: Ludwig von Mises Institute.

Huerta de Soto, Jesús. 2008. *The Austrian School: Market Order and Entrepreneurial Creativity.* Cheltenham, UK: Edward Elgar.

Hülsmann, Jörg-Guido. 2000. A Realist Approach to Equilibrium Analysis. *The Quarterly Journal of Austrian Economics* 3(4): 3-51.

Hülsmann, Jörg-Guido. 2001. Garrisonian Macroeconomics. *The Quarterly Journal of Austrian Economics* 4(3): 33-41.

Hülsmann, Jörg Guido. 2002. A Theory of Interest. *The Quarterly Journal of Austrian Economics* 5(4): 77-110.

Jevons, William Stanley. [1871] 1965. *The Theory of Political Economy*. New York: Augustus M. Kelley.

Jevons, William Stanley. 1876. *Money and the Mechanism of Exchange*. New York: D. Appleton and Company.

Kirzner, Israel M. 1966. *An Essay on Capital*. New York: Augustus M. Kelley.

Knight, Frank H. 1934. Capital Time, and the Interest Rate. *Economica, N.S.* 1(3): 257-286

Knight, Frank H. 1935. Professor Hayek and the Theory of Investment. *Economic Journal* 45(177): 77-94.

Knight, Frank H. 1941. Professor Mises and the Theory of Capital. *Economica* 8(32): 409-427.

Lachmann, Ludwig M. 1947. Complementarity and Substitution in the Theory of Capital. *Economica* 14(54): 108-119.

Lachmann, Ludwig M. [1956] 1978. *Capital and Its Structure*. Kansas City: Sheed Andrews and McMeel, Inc.

Lewin, Peter. 1999. *Capital in Disequilibrium: The Role of Capital in a Changing World*. London: Routledge.

Machlup, Fritz. 1932. The Liquidity of Short-Term Capital. *Economica* 37:271-284.

Machlup, Fritz. 1935. The Rate of Interest as Cost Factor and as Capitalization Factor. *The American Economic Review* 25(3): 459-465.

Machlup, Fritz. 1943. Forced or Induced Saving: An Exploration into Its Synonyms and Homonyms. *The Review of Economics and Statistics* 25(1): 26-39.

Menger, Carl. [1871] 2007. *Principles of Economics*, (*trans*.) James Dingwall and Bert. F. Hoselitz. Auburn, AL: Ludwig von Mises Institute.

Menger, Carl. [1883] 1963. *Problems of Economics and Sociology*, (*trans*.) J. Dingwall & B. F. Hoselitz. Urbana, IL: University of Illinois Press.

Menger, Carl. 1892. On the Origins of Money. *The Economic Journal*, 2, 239-255.

Menger, Carl. [1909] 1977. *Money*, chapters V – XIV, (*trans*.) A. H. Zlabinger. As Prepared for "Carl Menger Symposium" at Claremont Men's College. May 26-28.

Mises, Ludwig. von. [1953] 1971. *The Theory of Money and Credit,* (*trans*.) H. E. Batson. Irvington-on-Hudson, NY: The Foundation for Economic Freedom Inc.

Mises, Ludwig von. [1932] 1990. The Position of Money Among Economic Goods, reprinted in *Money, Method, and the Market Process*, (*ed*.) Richard M. Ebeling. Norwell, MA: Kluwer Academic Publishers.

Mises, Ludwig von. [1949] 1998. *Human Action*. Auburn, AL: Ludwig von Mises Institute.

Myrdal, Gunnar. 1939. *Monetary Equilibrium*. William Hodge and Co. Ltd.

Niehans, Jürg. 1978. *The Theory of Money*. Baltimore: John Hopkins University Press.

Northrup Buechner, M. 1989. Roundaboutness and Productivity in Böhm-Bawerk. *Southern Economic Journal* 56(2): 499-510.

Nurske, Ragnar. 1935. The Schematic Representation of the Structure of Production. *The Review of Economic Studies* 2(3): 232-244.

Patinkin, Don. 1956. *Money, Interest, and Prices*. Evanston, IL: Row, Peterson & Co.

Reisman, George. [1990] 1998. *Capitalism: A Treatise on Economics*. Ottawa, IL: Jameson Books.

Rosenstein-Rodan, P. N. 1936. The Coordination of the General Theories of Money and Price. *Economica*, *N. S.* 3(11): 257-280.

Rothbard, Murray N. [1962] 2004. *Man, Economy, and State: A Treatise on Economic Principles*. Auburn, AL: Ludwig von Mises Institute.

Rothabrd, Murray N. [1963] 2000. *America's Great Depression*. Auburn, AL: Ludwig von Mises Institute.

Rothbard, Murray N. (1976). The Austrian Theory of Money, in *The Foundations of Modern Austrian Economics*, E. G. Dolan, (*ed*.). Kansas City: Sheed and Ward Inc.

Rothbard, Murray N. 1978. Austrian Definitions of the Supply of Money, in *New Directions in Austrian Economics*, Louis M. Spadaro, (*ed*.). Kansas City: Sheed Andrews and McMeel.

Rothbard, Murray N. [1995] 2006. *An Austrian Perspective on the History of Economic Thought, Volume II: Classical Economics*. Auburn, AL: Ludwig von Mises Institute.

Salerno, Joseph T. 1987. The "True" Money Supply: A Measure of the Supply of the Medium of Exchange in the U.S. Economy. *Austrian Economics Newsletter* 6(4): 1-8.

Samuelson, Paul A. 1980. *Economics*, *11th ed*. New York: McGraw-Hill Book Company.

Say, Jean-Baptiste. [1803] 1880. *A Treatise on Political Economy or the Production, Distribution, and*

Consumption of Wealth, (*trans*.) C. R. Prinsep & C. C. Biddle. Philadelphia, PA: Claxton, Remsen & Haffelfinger.

Sechrest, Larry J. 2001. Capital, Credit and the Medium Run. *The Quarterly Journal of Austrian Economics* 4(3): 63-77.

Shackle, G. L. S. 1967. *The Years of High Theory: Invention and Tradition in Economic Thought 1926-1939*. Cambridge: Cambridge University Press.

Shostak, Frank. 2000. The Mystery of the Money Supply Definition. *The Quarterly Journal of Austrian Economics* 3(4): 69-76.

Skousen, Mark. 1990. *The Structure of Production*. New York: New York University Press.

Smithies, A. 1935. The Austrian Theory of Capital in Relation to Partial Equilibrium Theory. *The Quarterly Journal of Economics* 50(1): 117-150.

Strigl, Richard von. 2000. *Capital & Production*, (*trans*.) Margaret Rudelich Hoppe and Hans-Hermann Hoppe. Auburn, AL: Ludwig von Mises Institute.

Tullock, Gordon. 1957. Paper Money – A Cycle in Cathay. *Economic History Review, Second Series* 9(3): 393-407.

Walras, Leon. [1898] 1992. Etudes d'économie politique appliquée. In *Oevres économiques complétes*, Vol. 10. Paris: Economica.

Wicksell, Knut. 1951. *Lectures on Political Economy: Volume I*. London: Routledge and Kegan Paul.

Wu, C.-Y.. 1939. *An Outline of International Price Theories*. London: George Routledge and Sons.

Yeager, Leland B. 1979. Capital Paradoxes and the Concept of Waiting, in *Time, Uncertainty, and Disequilibrium*, (*ed*.) Mario J. Rizzo. Lexington, MA: D. C. Heath and Co.

Book III
Action and Its Value

I. INTRODUCTION

Having established the specific forms of action that occur within a market setting, we may now set out on the task of establishing where value sources from concerning these actions. Until this point we have placed a heavy emphasis on individual actions within the market. This adherence to methodological individualism has provided the strong foundation necessary to continue. However, in reality, we see that many occurrences on the market are performed through individual actions within the context of an institutional setting.

The firm, for instance, is the source of almost all productive activity of any well developed society. Until now the firm has been a relatively overlooked area of economic analysis. A recent surge in literature concerning the "entrepreneurial theory of the firm" has spurred a revival in this area of research, however, only small steps have been taken towards identifying the true essence of the firm. As will be shown, firms are attempts to synthesize *pure entrepreneurs*, much along the lines we have previously developed. Hence, entrepreneurs typically fall into one of two classes – risk-mitigating (i.e., efficient) or uncertainty-bearing (i.e., forward looking). Both of these qualities are necessary for a firm to successfully function. Only having one of these qualities will not be enough to successfully engage in any productive activity. Hence, by incorporating both functions in one entity, benefits can be experienced.

Once we see the firm as a synthesized pure entrepreneur, we may delve deeper into the value adding qualities that are inherent in it. Mrydal's (1939) monetary equilibrium demonstrates the equilibrating tendencies that exist and most dominate in the long-run. First, an equality between savings and investment must obtain. Second, an equality between profit rates of the individual stages of production will have a tendency to manifest. Of course, sharp entrepreneurial foresight will, at any given time, allow the possibility of firms to exist exhibiting above average profitability, but in the long-run, competition will ensure that these are equalized across all sectors.

By incorporating the structure of production into investment analysis, we may also provide the micro-foundations necessary to assess two ancillary concepts – vertical and horizontal integration. Hence, firms that are "optimally" integrated in both directions will have a higher value placed on them than firms that are too narrow, or spread out too broadly along the productive structure. Hence, added complexities may make it more cost effective for a firm to contract across different stages of the productive structure. Alternatively, it may be more profitable for a single firm to internalize multiple stages within its own boundaries. When Henry Ford opened the Rouge River plant, it was one of the first large-scale attempts at internalizing multiple stages in one firm (and one location). The reason that it worked so well was that competitors did not have the other productive processes that Ford needed available to be contracted. The only option for Ford was to internalize them and manufacture them within its existing firm boundaries. Likewise, horizontal integration involves ensuring there is an adequate supply available to meet the demands of consumers. Conversely, it is possible to have too great a productive capacity and hence, have value drained as it is shown to be not beneficial for the firm. High-tech firms with Internet cable capacity after the bubble of the late 1990s are a prime example where value has been removed from companies by being too horizontally integrated.

Lastly, we can see three specific determinants of value concerning financial assets. The first is the dividend stream expected to be received by an investor. As this stream is dependent on firm profitability, the onus falls on the firm's risk-mitigators to produce as profitably as possible. Secondly, the future growth rate of the firm is dependent on the uncertainty-mitigating entrepreneurs directing the firm into new profitable directions. Hence, we see on a micro-scale that much firm value is determined through the individual employees at work within it. Lastly, we will see that the expected length of time that an individual will hold a financial asset for will significantly affect its value. As the holding period is increased, a larger stream of dividends will be earned. However, this duration (which is the temporal dimension of saving) will be tempered by two factors. First are the personal attributes of the investor. Age, for instance, will dictate to some degree how long an individual will renunciate consuming before they commence again. Institutional factors such as investment companies or foundations may override this limitation, however. Hence, many foundations do not plan solely for an individual's life-span, but look to the future and satisfying many other temporal goals.

By linking the theory of the entrepreneur, the structures of consumption and production, and the temporal aspect of saving, we can begin to shed light on the specific sources of value for financial assets. In this way, the building blocks of micro-foundations for finance theory are laid, with hope that a larger edifice may be built.

II. THE FIRM

The structure of production, as was formulated in Book II, demonstrated that a series of interconnected productive actions would forward circulating capital goods toward a state where they are transformed into consumption goods. A problematic outgrowth of presenting the structure of production as a Haykeian triangle has been a bifurcation as to what this x-axis should be labeled as. Some regard it as a strictly temporal axis, although add the caveat that it is measured in "production time." Others regard this axis as representative of "stages of production" – industries that are aligned separately to represent the flow of capital along the path to production. Hence, under this view, capital begins life at a higher stage of production (i.e., mining) and moves through stages closer and closer to the point where it is transformed to a consumers' good. Capital is mined, goes through a period of manufacturing, is distributed to warehouses, and finally retail stores sell these goods to consumers – such is the flow of capital under this reckoning.

Of course, one the failings of this focus has been the lack of a corresponding structure of consumption to off-set the production component. However, there has been another unfortunate consequence as viewing the structure as pre-defined stages of production has overlooked other insights. One of the most important of these is the theory of the firm. Until now we have defined the production process in terms of individuals each adding value with additional actions along the productive structure. This has allowed us to create a structure of production based on the same principle – consumption is that series of actions which remove value from consumers' goods. However, the reality of the situation is such that we see individuals partake in social cooperation to produce goods *together*. It is true that there still exist artisans and independent workers operating in relative isolation, but the great bulk of production is marked by a high degree of collective effort. The grouping of labor to make a common product is commonly referred to today as "the firm."

Much like Coase (1937) originally reckoned in his famous article *The Nature of the Firm*, the existence of firms creates three pressing questions that must be answered.[296] First, we must answer why

[296] Although Coase (1937) may have marked the beginning of economic research directed at answering the origin and structure of the firm, Berle and Means (1932) had previously paved the way, raising the issue of separation of control between shareholders and company managers. This question as to why and what effects a separation of control has are very similar to the questions why and to what effect can production be centralized within a firm. An additional pioneering study was provided by Hall and Hitch (1939) who showed that managerial decisions were not based on Marginalist principles, such as those which all post Marginalist revolution economics have been based upon. The disparity between economic theory, and firm activity, spurred the need for further research into the functioning of the firm.

it is that firms exist, that is to say, why are all actions not undertaken by individuals working independently of each other. Second, where is the boundary between the firm and the market, and how is it that it is located where it is. Hence, there is the trade-off between performing actions within a firm, and performing them outside of a firm's organization – why is it that the firm's boundaries extend in the way that they do? Last, we must ask why it is that firms are structures in the way that they are – how do the formal and informal relationships function together within the firm's setting? Demsetz (1995) points out that questions pertaining to the firm's origins, and those concerning its organizational structure, are two separate problems to be dealt with independently. For the present work, we will be concerned with answer the first question – why is it that firms exist if production is able to be done by individuals.

Two traditional camps have looked at these questions in hopes of finding their answers. The first are economists following Coase's lead who try to explain a firm's existence through transaction costs. Hence, differences in the costs incurred between inter-firm exchanges and intra-firm exchanges create incentives for further market growth, and also provide disincentives when firms grow to large to effectively function as independent entities.[297] A secondary approach, pioneered by Penrose (1959) seeks to answer the questions as to the firm's existence by delving into its specific internal capabilities. Hence, this capabilities based approach sees firms as bundles of skills that are shared and developed between the human assets that constitute them.[298]

More recently, a competing theory of the firm has been gaining importance. The entrepreneurial theory of the firm seeks to answer the previous three questions, as well as one additional one – what is the role of the entrepreneur within the firm's setting? By synthesizing the entrepreneur into the firm, differing skill sets are linked together to create core competencies within the firm.[299]

While all three of these viewpoints will be critically assessed, we will see that none is the definition of the firm that we wish to use for our case. Instead, we see a more full integration of the entrepreneur is necessary, more so than even these entrepreneurial theories can provide. Thinking back to Book I, chapter V, we saw that entrepreneurs comprise two different sets of skills. One is the bearing

[297] This approach may be better characterized as a broader "contractual" approach, whereby firms exist to contract resources with an optimal incentive structure. See Alchian and Demsetz (1972), Williamson (1975; 1985), and Hart (1995) for pioneering research in this direction.

[298] See also Richardson (1972) and Nelson and Winter (1982) for early literature on this approach.

[299] There are two additional theories of the firm which deserve comment, though have little application in the problem we wish to address. The first views firms as the result of technological necessity – there are some production processes which could not be undertaken absent the firm's scope. The other is the denial that firms have a distinct nature from the market – they are just a special type of contractual method (see Alchian and Demstez 1972; Jensen and Meckling 1976: 311; Fama 1980; Cheung 1983,). Neither of these theories addresses the necessity of explaining why firms exist distinct of a market as a *general theory*, with Williamson (1985; 1995) arguing the technological argument is not a theory of the firm. These theories are summarized in Alchian and Woodward (1988), Holström and Tirole (1989) and Sautet (2000).

of uncertainty that is given either: 1) statically through non-repeatable decisions, and 2) through future events yet to obtain. Alternatively, entrepreneurs may also exhibit the skill of risk mitigation. That is to say, there are some people who excel at efficiency, waste reduction, or any number of roles that require no degree of uncertainty (i.e., they are purely risky), but require changes in the current state of affairs to mitigate and reduce risk. The common link between these two entrepreneurial roles was the drive towards increased consumer want satisfaction. We have seen that this can be achieved in two ways. The first is through satisfying needs in the future that are not known to exist today (or satisfying them better). The second way is by reducing the cost that consumers must renunciate in order to obtain goods to satisfy their wants. This second category may occur with no degree of uncertainty, as producers can increase consumer want satisfaction merely by reducing the cost to obtain the goods that serve this end.

As a perfect entrepreneur was previously defined as one who: 1) mitigates all risk to the maximum extent, and 2) bears all uncertainty perfectly, we find that this occurrence – the combination of both skill sets to the highest degree – is an unlikely occurrence in the real work. We could even go so far as to exclude it from consideration as it is so unlikely it is impossible. At this point the theory of the firm becomes instrumental.

Firms combine entrepreneurs of differing skill sets to try to synthesize the perfect entrepreneur. Hence, firms strive to fulfill consumers' demands to the highest possible degree by perfectly foreseeing all uncertainties, and mitigating all ensuing risks. To the extent that firms combine different individuals together, they are trying to synthesize this perfect entrepreneur.

With this foundation, it will become clear that the structure of production is really a combination of a few autarkic individuals working to produce consumers' goods, mixed in an interconnected market of firms producing these same goods. As the individuals are what defines the firm – both in its emergence and boundaries – understanding their role in production becomes instrumental to understanding the greater firm they create. To the extent that we have previously established the micro-foundations of action, we may now turn our attention to a more macro-institutional framework, such as provided through the firm.

1. Theories of the Firm

Contractual Theories of the Firm

Coase (1937) opened the modern emphasis on the study of the firm within economics. In his pivotal article, there are costs associated with using the market price mechanism (i.e., transaction costs), which may be eliminated if they are internalized within a firm's boundaries. Hence, in place of using the price mechanism for allocative issues, decisions may be directed through an entrepreneur via contracts determining how production should proceed. Costs are incurred, however, by internalizing these directions as market conditions dictate less directly the necessary adjustments that must be made. Hence, a firm will exist where the costs associated with using the internal direction of resources is less than that of using the price mechanism on the market. Within this approach we may notice two distinct sub-fields.

The first are those which focus on the measurement of costs associated with the internalization of decision-making. Following Alchian and Demsetz (1972) we find the emphasis laid on issues surrounding the costs associated with administering, directing or monitoring teams. In contrast, Williamson (1985) looks at the post-contractual governance structures that influence the issues at hand – moral hazard, bounded rationality and opportunism, for example.

One significant problem with these theories of the firm is the ignorance towards the entrepreneurial function that is displayed. Sautet (2000: 41) points out that both Coase and Williamson operate within an equilibrium framework, hence, their explanation of the firm must exclude any entrepreneur from being the prognosticator. Instead, any changes in the firm (including its origin) must result from bounded rationality issues, or opportunism. Boudreaux and Holcombe (1989: 147) reach a similar conclusion concerning Coase, noting "the Coasian firm emerges only after markets exist: it engages in management rather than true entrepreneurship."

Perhaps the largest misgiving we find with the modern contractual theories of the firm are the treatments of uncertainty, and its counterpart – ignorance. Sautet (2000: 70) stresses that a deficiency lies in the way that these authors assume that the cost to overcome *known* ignorance is attainable. A neglect for the serious effects of true uncertainty becomes manifest.

The idea that a trade-off is consciously made between the costs of using the market price mechanism and the costs of using internal direction of resources depends crucially on the definition one has of uncertainty. For, as we have seen, the determination of these prices ex ante is an impossibility owing to the uncertainty surrounding their attainment. Economists writing in this tradition have generally neglected this possibility, a direct result of working within an equilibrium framework which necessarily excludes true uncertainty from arising. As a result, these theories see little more than a fallacious link between entrepreneurship and the firm.

Capabilities Theories of the Firm

Penrose (1959) opened a new line of research concerning the firm by assessing the internal qualities that comprise its capabilities. Hence, under this capabilities theory of the firm, the individual skill sets that individuals have are what explain the firm's success. Firms emerge and grow dependent on the quality of the employees that comprise its boundaries, hence, as Penrose (1959: 3) summarizes the view:

> There surely can be little doubt that the rate and direction of the growth of the firm depend on the extent to which it is alert to act upon opportunities for profitable investment. It follows that lack of enterprise in a firm will preclude or substantially retard its growth.

The rate of growth, or even the *possibility for growth*, of a firm depends on the abilities of the entrepreneurs to move the firm into new unchartered territories. Indeed, for Penrose, economics must make a distinction between the entrepreneurial activities that move a firm forward, and the managerial theories that we wish to study from the vantage point of an established firm.[300]

This capabilities-based approach has seen some contributions from writers usually associated with the previous camp of contractual theories of the firm. Hence, Williamson (1985: 131-162) defines two types of incentives which motivate the individuals within a firm. The first type – high-powered incentives – are those that exist in the market, and may theoretically continue forever without end. The second type – low-powered incentives – are those that characterize firms, and have known limits (i.e., an employee has a set salary). Managers typically become exposed to high-powered incentives through performance-based remuneration, while lower-level employees face the known low-powered incentives. The importance of these incentive structures becomes known once we see, as Penrose

[300] For an early attempt at melding these two viewpoints – the theory of the entrepreneur and the theory of the firm – see Foss and Klein (2002).

(1959: 35) does, that: “[T]he managerial competence of a firm is to a large extent a function of the quality of the entrepreneurial services available to it.”

The influence of this capabilities approach has been broad, but more particularly it remains isolated within business schools, used in courses such as management, or entrepreneurial studies, while being mostly excluded from the realm of pure economic theory. Much like Kirzner lamented that the *pure* entrepreneur is a difficult concept to study within the framework of economics as it is conceptually based upon individual characteristics, so to does this capabilities theory suffer the same fate.

As Sautet (2000: 97) demonstrates, there are two consequences of this approach to the firm. First is that it has led the theory of the firm to be contained within the boundaries of previously established neo-classical doctrine of economics. Hence, by viewing firm growth in terms of *known* capabilities, the problem becomes one of optimizing these distinct skills to maximize an outcome – profit, output, sales, etc. These theories fail to account for change that occurs in a disequilibrium world, instead focusing on how these skills-sets are maximized within an equilibrium setting of a known firm. Second is that a new theory of change and growth has been created, different than the then-existing analysis. For example, the way in which management can use and organize different resources becomes crucial to understanding how the firm grows, as well as what new products may be made available given *existing* capabilities. This becomes compatible with an open-ended world in some regards, however the use of known capabilities limits this open-endedness to some degrees.

Capabilities are an important aspect of the firm, but, as Foss (1996: 24) states: [I]t is possible to construct propositions about economic organization that do not directly turn on considerations of morally hazardous behavior and incentive alignment.” The incentive issues that Williamson raises add to our understanding of firm behavior, but do little to add to our knowledge concerning the emergence of such institutions. Instead, as we shall see, the Penroseian insight into the specific qualities of entrepreneurs *cum* employees requires a further integration into the theory of the firm than has previously been afforded.[301]

An Entrepreneurial Theory of the Firm

[301] Knight (1921: 268) alludes to this necessary but not sufficiency aspect of capabilities thus: “When uncertainty and the task of deciding what to do and how to do it takes the ascendancy over that of execution, the internal organization of the productive group is no longer a matter of indifference or a mechanical detail. Centralization of this deciding and controlling function is imperative, a process of 'cephalization,' such as has taken place in the evolution of organic life, is inevitable, and for the same reasons as in the case of biological evolution.”

A recent outgrowth within the literature concerning the theory of the firm is an emphasis on providing a link between the oft neglected entrepreneurial function, and the firm which often this individual works within. Many different aspects have been raised, each with its own validity and necessity concerning the firm. However, as we shall see, there is a missing feature in each contribution – an emphasis rooted in the true entrepreneurial function, much like we have outlined in book I, chapter V.

In perhaps the most developed and complete entrepreneurial theory of the firm, Sautet (2000) contends that intertemporal coordination is problematic, especially concerning labor, in the absence of firms. This is due to the "unexploitability thesis" which states that in the absence of firms, the exploitation of an opportunity discovered through labor cannot take place as labor is both a specific and non-specific asset. Hence, individual A might not be able to exploit their opportunity, as in the absence of a firm (with its implicit and explicitly contractual obligations), a different individual – B – could bid away A's labor (Sautet 2000: 74). As a result, a firm is necessary to implement a series of long-term contracts that supersede the pricing mechanism on the market.

Under this view, we find that firms are non-price planned coordination centers that mitigate errors stemming from ignorance of each others' abilities on the market. As an entrepreneur coordinates this activity, a centralized decision-making process creates a firm that acts as an "island of planning that serves as the locus of exploitation of a discovered profit" (Sautet 2000: 76). When intertemporal uncertainty is prevalent, firms are required for two reasons. The first is to mitigate the uncertain nature of the market process. The second is to specify the inputs necessary for the exploitation of a profit opportunity by a common group of individuals. "The firm is the pulling together of entrepreneurial activity by a central entrepreneur: the promoter" (Sautet 2000: 83).[302]

Loasby (2002) stresses the role that firms serve as collaborators of knowledge. As knowledge exists in only an incomplete, fragmented, and tacit manner, organizations provide the structure necessary to reflect and shape the available knowledge in a direction that effectively achieves or aids its own ends. Langlois (2002) views firms from a more "modular" standpoint. As they partition decision-making rights and differentiate between residual claimants of profit, firms are able to modulate in order to optimally exploit benefits of team-based production and asset specificity. Indeed, this view becomes more important once the realization is made that firms result in the generation of knowledge that exceeds that of its individual components (Aoki 1990). Synergies occur between individuals working

[302] See also Casson (1982) for this emphasis on the entrepreneur as the coordinator within the firm.

together which generate an increased knowledge base than would exist in the absence of this collaborative approach. Likewise, Garrouste (2002) points out that as knowledge is a process of learning, internalizing this process within a firm secures the use of its benefits in the future.

One of the shortcomings of previous theories of the firm is their implicit assumption that markets precede their emergence. Hence, a market exists, and a firm is formed to exploit this market. We find that in Coase's firm (and followers of this contractual theory tradition), the trade-off between the cost of internalizing decisions and that associated with using the pricing mechanism in the market presupposes that a market exists in the first place. However, we see that many markets are the result of firms. Indeed, as Casson (1997) stresses, firms act as market-making intermediaries which continuously monitor changes in supply and demand to adjust price and output accordingly. In this way we see that firms are integral to any concept of the market we wish to utilize.

There are some issues with these theories, despite the obvious advantages gained over the contractual and capabilities theories previously discussed. Sautet, in the heretofore most extensive expression of what the entrepreneurial theory of the firm is, focuses on the intertemporal coordination issues, but at the neglect of coordination problems that occur only in the present. The emphasis of the Misesian entrepreneur, looking to the future continually to bear uncertainty, is partly responsible for this emphasis. As has been assessed in book I, chapter IV, there are many uncertainties that exist in a timeless (i.e., static) setting as well. Likewise, although he is keen to focus on the problems arising from labor as both a specific and non-specific factor of production, he neglects what would happen in circumstances when labor is either one or the other. In a general sense it is true that labor embodies both these contradictory facets, but the ends of action are concerned with labor in its *specific* applications. Obviously labor cannot be both specific and non-specific at the same time concerning a specific action, the result being *contradictio in adjecto*.

Likewise, Loasby's (2002) stress on the collaboration of knowledge within a firm cannot explain why the same cannot happen outside a firm (or by an individual). Garrouste's (2002) emphasis on internalizing the learning process inside a firm is valid, but begs the question: “Why would this factor mark the origin of the firm?” Indeed, this internalization of knowledge benefits can be seen as neither a necessary, nor sufficient, factor to explain the emergence of the firm. As Huerta de Soto (1992) demonstrates, the maximum benefits of knowledge are experienced through the unbounded allowance for their influence to spread, and occurrence sure to be hampered within the defined boundaries of a firm.

Many of these problems stem from an erroneous, or incomplete, theory of the *entrepreneur*. As a theory of an entrepreneurial firm relies on this fundamental concept, it will prove beneficial to commence from that point as we rebuild the entrepreneurial theory of the firm in light of this consideration.

2. A Pure Theory of the Firm

Introduction

Earlier, in book I, chapter IV, we looked at forming a new theory of the entrepreneur. One important piece that has been missing in theories up to this point is the lack of *direction* concerning the entrepreneur's efforts. While Mises' entrepreneur can be seen to look to the future with the eyes of an historian, we see little emphasis placed on why exactly they are doing so (Mises 1949: 56; 1957: 320). Likewise, although Kirzner's (1973) entrepreneur primarily is concerned with discovering latent opportunities, there is little attention paid to the specific reason why this effort exists. Huerta de Soto's (1992) entrepreneur as a discoverer, collaborator, distributor and interpreter of knowledge demonstrates much concerning one of the central benefits of the entrepreneurial function, but lacks in its explanation as to why the entrepreneur undertakes this role.

We have corrected this missing facet by drawing attention to the end goal that entrepreneurship is concerned with – the increase in consumer want satisfaction. All entrepreneurial activity is directed toward this singular end. While this may seem to be compatible with all three of the aforementioned theories, at the same time they ask too much and too little. Too much in that they assume that only actions that fall under their respective categories can be said to be true entrepreneurship. They ask too little by failing to address the specific necessity that entrepreneurship fulfills – consumer want satisfaction.

We have rectified this by showing that entrepreneurship can function in two distinct ways to increase consumer want satisfaction. The first is through bearing uncertainty and delivering want satisfaction that were previously unknown to exist. This is achieved in two ways. The first encompasses the stress Mises placed on future uncertainty. Hence, there are situations unknown to exist in the future, and through the entrepreneur's actions, these are brought to light; the entrepreneur creates the future. Alternatively, there are also what we have termed static uncertainties, which are those decisions that must be undertaken in the present that we have no logical explanation for the result – it is fundamentally uncertainty. We saw that Shackle's (1949) stress on "non-divisable, non-seriable" events gives rise to a new class of uncertainty – those belonging to actions which are unique events. In fact, we see that there are many decisions which are made in the present which have typically been classified as "risky" (i.e., those with known historical probability distributions) for which these probability distributions do nothing to annul our uncertainty. For what would be the case if we had the choice of:

> … drawing a card from a pack containing 25 red cards and a black one, or from a pack containing 25 black cards and a red one; and if the drawing of a red card were destined to transport him to eternal felicity and that a black one to consign him an everlasting woe, it would be foolish to deny that he ought to prefer the pack containing the larger proportion of red cards, although from the nature of the risk, it could not be repeated... [S]uppose he should choose the red pack and draw the black card. What consolation would he have? He might say that he had acted in accordance with reason, but that would only show that his reason was absolutely worthless. And if he should choose the red card, how could he regard it as anything but a happy accident? He could not say that if he had drawn from the other pack he might have drawn the wrong one, because an hypothetical proposition such as 'If *A*, then *B*' means nothing with reference to a single case.' (Shackle 1952: 110-111)

Hence, entrepreneurs can bear uncertainty through two aspects. First is bearing future unknown events, and the second is the bearing of uncertainty in the present of known events whose outcomes cannot be assigned any meaningful outcome distribution.

If entrepreneurship is defined as increasing consumer want satisfaction, we find an additional method to achieve this besides the more common act of uncertainty bearing. Consumption always involves the element of renunciation for attainment. Hence, we must always exchange something in order to have our wants satisfied. Entrepreneurs can increase our want satisfaction by reducing the renunciation part of the exchange process. This is a distinctly different type of entrepreneurship than the above type as it *may* involve no degree of uncertainty. Instead, the focus is shifted towards increased efficiency in want satisfaction, not in satisfying heretofore unknown wants. It may be, then, that a production process has an established demand and known inputs – in the short-term it exists within a closed-ended system. Hence, there is no uncertainty element involved with this production, but there is much risk which requires mitigation. As an example, it could be that the production process has 5 unsaleable outputs for every 100 that are produced. This 5% rate of unsaleable goods produced represents an element of risk which can be mitigated further. A reduction of this rate to 3 per 100 produced goods would decrease this risky element of production. As a result, prices can be lowered for consumers, and hence the renunciated portion of their own exchange for the good is reduced commensurately; consumer want satisfaction has been increased.

A pure entrepreneur, that is to say, one who brings a theoretical equilibrium into existence, is one who bears all uncertainty (future and static) perfectly while also mitigating all risk to a maximum extent. The perfect entrepreneur is a theoretical construct, in reality we find that the existence of one is an impossibility. This arises from the limited cognitive abilities we have as actors, which brings forth issues such as imperfect knowledge as well as risk and uncertainty into our actions. If the pure entrepreneur fails to exist as a singular being, is there a possibility of synthesizing this actor in hopes of increasing consumer want satisfaction? As we will see in the next section, this activity is the essence of what the firm undertakes at its origin.

The Pure Entrepreneur as a Firm

> "Why do entrepreneurial firms exist? When I first heard the question, it seemed an easy one to answer. "Entrepreneurship" and "firm" were part of the prose I had been speaking all my academic life. After a bit of thought, however, it became clear that this was not exactly the question I had been trying to answer for more than 20 years. My question has really been Coase's question: why does the firm exist?" (Langlois 2005: 1)

Two simple concepts – entrepreneurship and firms – are used on a daily basis, but little is understood about the interaction between these two concepts. However, many economists have tried to answer the original Coaseian question "why do firms exist?" without a basis of entrepreneurship. The results have been muddled and theoretically lacking.[303]

Entrepreneurs exist as individuals with differing skill sets. We find that in the general categories of risk and uncertainty, there are individuals who show distinct ability to excel in each independently of

[303] Although, as Foss and Klein (2006: 6) note:

> [A]lmost since the inception of economics, many economists have had things to say about the firm. However, the recognition that a theory of the firm is necessary is – seen in the perspective of the history of economics – a recent recognition. In the same way, it is in fact possible to say much about, for example, earthquakes without arriving at a theory of earthquakes in the sense of a theory that explains the reasons for the occurrence of earthquakes. By the same token, much can be said about firms from an economic perspective (e.g., the size of the individual firm, size distributions of firms, market behavior, market power....) without saying anything theoretically grounded about the reasons for the existence of firms, the reasons for their size, boundaries, internal organization, incentive mechanisms, etc. in terms of, for example, causal explanation or lists of sufficient reasons. The relatively detailed account of firms' production and selling decisions found in intermediate microeconomics may exemplify a theory about, but not of, the firm.

the other. For example, it is easily seen that there are individuals who can work on an assembly line and perform their job with the utmost efficiency, while if placed in the position of CEO of the company could not navigate the future with any degree of certainty. Conversely, there are heads of companies who are skillful at seeking new markets, new products and expanding revenue streams for their respective firm, who lack the basic ability to file their own expense reports. Being able to navigate uncertain waters makes no presupposition of being able to mitigate its inherent risks. Likewise, being able to efficiently minimize risk makes no statement as to dealing with future risk that will develop at an unknown time due to uncertainty. These are two distinct concepts – risk and uncertainty – as are the methods individuals use to deal with them.

However, while individuals all inherently have differing degrees of ability to comprehend and effectively deal with risk and uncertainty, it is through *combinations* of individuals that a *unified* whole can be achieved to more optimally manage these two factors. Firms become the method that individuals implement in order to more effectively use and have control over each others' resources regarding risk and uncertainty bearing. This concept may be quite similar to the Knightian firm, whereby individuals unwilling, or incapable of bearing uncertainty will join those who can and form an organization. In Knight's (1921: 269) own words, the firm is "the system under which the confident and venturesome 'assume the risk' or 'insure' the doubtful and timid by guaranteeing the latter a specified income in return for an assignment of the actual results." However, there are important differences which bear attention.

Knight was quite astute in noting the inherent differences individuals have concerning their personal abilities to manage risk and uncertainty. After all, individuals differ:

> … in their capacity by perception and inference to form correct judgments as to the future course of events in the environment... [There are differences in] men's capacities to judge means and discern and plan the steps and adjustments necessary to meet the anticipated future situation. (1921: 241)

As Langlois and Cosgel (1993: 461) point out, these inherent difference between individuals gives rise to specialization within their respective skill sets. Hence, Knight (1921: 268) refers to a cephalization process, which results in the firm:

> When uncertainty and the task of deciding what to do and how to do it takes the ascendancy over that of execution, the internal organization of the productive group is no longer a matter of indifference or a mechanical detail. Centralization of this deciding and controlling function is imperative, a process of 'cephalization,' such as has taken place in the evolution of organic life, is inevitable, and for the same reasons as in the case of biological evolution.

This process, in turn, implies that the "result of this manifold specialization of function is is *enterprise and the wage system of industry*" (Knight 1921: 271). The uncertainty that fogs the world results in the "tendency of the groups themselves to specialize, finding the individuals with the greatest managerial capacity of the requisite kinds and placing them in charge of the work of the group, submitting the activities of the other members to their direction and control (Knight 1921: 269). Hence, the Knightian emphasis is on specializing based upon judgment skills and managerial capacity, and not necessarily risk aversion.

One may, at this juncture, argue that the literature since Coase (1937) has demonstrated that in the absence of transaction costs, there is no reason why a theory of specialization would result in a pure theory of the firm – the market could contract this specialization without a firm's structure. However, Knight makes clear that judgment is a fundamentally non-contractual ability (Knight 1921: 311).

One problem with Knight's conception of the firm is his emphasis on uncertainty, at the neglect of risk. Hence, in his view, uncertainty bearers are the entrepreneurs who are sought after to command control of groups of individuals and direct them accordingly. What Knight fails to understand is that uncertainty bearing entrepreneurs also require risk mitigating abilities if they are to achieve their goals. The two functions are necessary for *continued* success within an organizational setting.[304]

An entrepreneur could correctly see all future uncertain contingencies, however, lacking any sort of efficiency through risk-mitigation in the present moment, would be unable to move towards this envisioned future state. It is only through contracting the abilities of an individual who can mitigate this risky element that an uncertainty-bearing entrepreneur may realize their goals of delivering greater amounts of future consumer want satisfaction. Alternatively, there could be a risk-mitigating entrepreneur who requires the assistance of an uncertainty bearing counterpart in order to foresee the

[304] Menger (1871: 160) makes four points concerning the role of the entrepreneur. His final point, that they exist to supervise the execution of the production plan comes close to approximating what we wish to convey about the emergence of the firm. Indeed, as he writes concerning firms: "In small firms, these entrepreneurial activities usually occupy but an inconsiderable part of the time of the entrepreneur. In large firms, however, not only the entrepreneur himself, but often several helpers, are fully occupied with these activities." The joining together of these "helpers" and entrepreneurs is, in Menger's own words, the idea we wish to convey concerning our own theory of the firm.

future directions that they must direct their efforts. This is, in fact, the process we see every day on the market.[305]

No single individual can be a pure entrepreneur, there is, however, a method that can be used to synthesize this ability. By contracting these services together, risk-mitigation and uncertainty-bearing, the advantages of both parties are internalized *and* exclusive to one another. For this part becomes the critical point. One may raise the argument, at this point, as to why in a world of no transaction costs would these parties not contract each others' services to one another, and perform their tasks via the market mechanism, and not through a firm.[306] But this argument would neglect several vital points.

The first point that we wish to stress is the exclusivity arrangement that forming a firm entails. Indeed, many times service with a company commences with an exclusivity contract signed ensuring that one's services are offered only to the contracting firm. This is so for two reasons. The first is that the firm does not want competitors to have access to what they deem the best resource of its type that they can obtain at the given time. Firms are continually in competition with one another, not only competing for the more conventional physical resources, but also for talent through our human capital. As a firm discovers an individual that they expect to deliver the most optimal services they can discover, they will have an incentive to prohibit other firms from contracting this individual as well.

Second, there is an advantage to be had by keeping knowledge created by the firm internal to the firm. A competitive advantage is often had through keeping proprietary knowledge internal to the firm. What good would it do to have a competitive advantage over other competitors, only to have the same employees contracted out to those same competitors and able to share this knowledge with them?

An example may serve to illustrate the problem. Assume many there is a group of individuals who know how to play football. Likewise, there is a set of coaches who know how to manage teams. The players are very good at mitigating risk, they can work within the closed-system of the pitch, on a field of predetermined size, with players of known abilities and numbers, and play against each other accordingly. For all intents and purposes we assume they bear no uncertainty. This task, in contrast, is undertaken by the coaches. They are the individuals who determine the qualitative aspects of the

305 In this way we go one step further than Simon (1991: 27) who noted that a theoretical Martian viewing the world through a telescope would see organizations, such as firms, as the "dominant feature of the landscape", and not dispersed markets. We see that not only are firms the dominant organizational form, but firms themselves are dominated by interconnected entrepreneurs attempting to form a *pure entrepreneur*.

306 Of course, Simon (1951) argues that employment contracts are always incomplete when employers offer wages in return for the employee to accept direction from the employers. The two parties are incapable of writing an enforceable contingent contract fully specifying what the employee must do. Hence, the employee must accept an open-ended contract which has its acceptance based on the extent the employee feels the future required work will fall within their own "zone of acceptance."

players' abilities, combine players to create optimal teams, and deal with the unforeseen consequences (i.e., injuries, trades, new skills, develop new techniques, etc). We see that individuals that comprise these two sets require individuals from the opposite set to fully develop their services. Coaches require players in order to fulfill their future plans (i.e., teams to win championships). Players require coaches to train them, make them aware of contingencies, and help them realize their unforeseen potentials.

First, assume a case where there are no transaction costs, and as a result players are contracted freely for each game – there is no restriction on the mobility of labor. One result will be that players will be contracted for games with a potential myriad of teams. As all players will be able to learn from all coaches, competition will become quite limited concerning the results of the games. Players and coaches will no longer be limited to a knowledge of their own team's abilities. Instead, as individuals are shifted around there will be a general increase in the homogeneity of knowledge that each will have. There will no longer be an exclusivity of knowledge element involved with individual teams. The process of competition will become limited as all will have access to the same knowledge set.

Now assume that exclusivity of knowledge is granted through the internalization of this factor within a firm's boundaries. Garrouste's (2002) emphasis on internalizing the benefits of learning become evident. As teams are now formed of defined sets of players, unable to switch in the short-run, there become clearly defined strategy advantages. Knowledge can be shared, and approaches developed which would be unable to achieve through a market-based system actively competing for each others' players and coaches. Furthermore, as Wilson (2009) points out, the amount of future uncertainty that will prevail is determined by the learning that occurs on an individual level leading up to this uncertainty. Hence, the learning process cannot be known in advance to dissipate the uncertainty, which creates the need to internalize this process within a firm's boundaries to increase the possibility of a more certain future. Competition would be increased as there is now a possibility that private information may be produced and others' excluded from using it. Benefits from learning are increased, as they are internalized within the firm's controlling sphere.

We see that firms are attempts by risk-mitigators and uncertainty-bearers to form the pure entrepreneur. To the extent that each needs the other, they each form an essential piece of the puzzle that both sides will seek with equal effort. How they affect firm performance will be different, however, as we will assess now.

Firms as Profitability Centers

In modern neoclassical economics, the firm is often referred to as a profit center. Its role is commonly stated as the pursuit of profits through the production of goods or services. In fact, many business text books would list the firm's explicit goal as profit maximization. This viewpoint is very important in one respect, but creates problems when viewed in isolation. It is true that firms, to some regard, maintain a *continued existence* through profits, however, without viewing what profits entail the emphasis is shifted from profitability from sustainability.

For firms to profit in the present, they require the ability to deliver a product at a price higher than the costs of production. Risk-mitigating entrepreneurs are the primary vehicle that generate this profitability for the firm. Indeed, through efficient production methods, that is to say, through the mitigation of all risk that is inherent in the production technique, profitability can be increased *in a static sense*.

For example, a firm may exist to produce crystal wine glasses. There is a known demand, and the production process is established such that there is a breakage of three glasses per 1,000 produced. *Statically*, profitability may be increased through the reduction in waste during the production process. In this way, if an employee can reduce the breakage rate from 0.3% to 0.2% (a reduction of 33%) then the profitability of the firm will be increased. This arises as less costs are attributed to inputs while output prices and sales remain constant. This type of profitability increase may only occur within a closed-system.

However, we see that in a closed-system the profitability increases are eventually self-defeating. In the next step, the breakage rate will be reduced from 0.2% to 0.1% (a 50% reduction!), then from 0.1% to 0.025% (a 75% reduction). Finally, there will be a point reached where no further risk may be mitigated from the process. For every 1,000 wine glasses produced, none are wasted in the production process. No further increases in efficiency can be wrought through decreases in risk – all possibilities are exhausted. Of course, this is complicated through competition. As other firms attempt to mitigate risk better than others, there is a continued diminution in the risk inherent in a process. These two effects work together to eliminate risk to the maximum extent in all production processes.

Firms therefore have a very important aspect in that they are profitable in the present through their ability to produce with a minimal degree of risk. Through the elimination of this factor they can produce at a lower cost and hence increase consumer welfare through a diminished cost of renunciation. As we saw earlier, consumer welfare can be increased in two ways. The first is through having more wants satisfied, and the second is through lowering the cost of servicing a want (through a reduction in the cost of renunciation to attain want satisfaction). Risk-mitigating entrepreneurs serve an important role in being the center of efficiency in the firm at any given time. In fact, we may go so far as to say that *at any given time* this type of entrepreneurs – risk-mitigators – are the *sole* source of firm profit. However, this is a self-defeating endeavor. Risk, by definition, may only be eliminated to a certain point. As risk involved with any process must sum to unity, it may only be decreased to a point of zero. As a result, although this type of entrepreneur serves a very important role – the creation of present profitability – they require an additional input in order for this factor to be sustained.

Firms as Growth Centers

For firms to grow they must move to unchartered waters. Continued growth and profitability may only stem from reaching into new, unexploited territory to discover new sources of profits. As we saw previously, profitability is exhaustible – lacking fresh risks there will eventually be a point reached where no further exploitation of efficiencies will reap continued profits. Hence we see the point in a firm's life where the uncertainty-bearing entrepreneur becomes so crucial. In order to allow risk-mitigating entrepreneurs to continue producing profitable activity, new avenues will require to be introduced in order to bring forward new risks. The importance of the uncertainty-bearing entrepreneur is thus two-fold.

The first point we notice is that consumer wants are not given, but will exist in a future that has yet to be discovered. The stress of the Misesian entrepreneur, looking to the uncertain future to discover yet unseen wants becomes apparent. As these wants are identified, the future is created. Risk-mitigating entrepreneurs will continually be delivered to fresh, risky situations where they can flourish by allowing their abilities to be put to full use. Continued profit opportunities are opened up as new situations and avenues for growth continually feed a fresh source of risk into the production process which will be rewarded accordingly. Hence, Mises' entrepreneur provides that continual source of novel events that are necessary in order for the risk-mitigating entrepreneur to remain relevant, and also for continued profitability to be ensured as a *potential*.

Kirzner's entrepreneur focuses almost exclusively on undiscovered opportunities. These exist in the present, but represent fundamentally uncertain elements as well. In fact, as we have previously seen, Shackle's stress on non-divisible, non-seriable events implies that those events which we are allowed only one opportunity to exploit are fundamentally uncertain to the actor. Hence, the Kirznerian entrepreneur provides secondary avenue for continued growth, and hence, profitability, within the firm. Fresh production possibilities will be unleashed by this secondary type of uncertainty, which will require new insight from the risk-mitigating entrepreneurs in order to maximize efficiency of the order. As a result, continued profitability remains a possibility provided that the Kirznerian entrepreneur can continue discovering new outlets to be exploited, and hence, bring about new viewpoints of risk into the production process. As Shackle (1952) makes clear: "[T]he point which I have tried to keep conspicuously in view [is] that the enterpriser often *does not wish* to eliminate risk, or rather, uncertainty: on the contrary, it is uncertainty which allows him, and only uncertainty which can allow him, to entertain as possibilities his highest levels of imagined success."

Both of these entrepreneurial types – Misesian and Kirznerian – are essential not for direct firm profitability, but for the continued *possibility* that profitability will be available. Uncertainty-bearing entrepreneurs thus provide the boundary that risk-mitigating entrepreneurs function within. Indeed, without providing this boundary, in effect a closed-system, the risk-mitigators could not effectively manage the risk involved in any production process. This is so as risk, by definition, requires a closed system concerning its outcome set. As these are continually exploited, it becomes necessary that the uncertainty-bearing entrepreneurs take the firm into new risky directions and hence, allow for the possibility of continued profits.

Future growth is a direct result of the entrepreneur's ability to discover and steer the firm to new vantage points. Lacking this ability, two outcomes will occur.

The first is that, at least for a short time period, lacking any fresh directions resulting from uncertainty, profitability may actually increase. This is a result of the fact that lacking any external changes in their world, risk-mitigating entrepreneurs may continue to mitigate the risk inherent in any process to the maximum extent. Thus, profitability, at least in the short-term, suffers from the new directions that the uncertainty-bearing entrepreneurs steer the firm to. This process disrupts the continual elimination of risk (through increased efficiency) that is occurring continually within the firm, and hence, decreases the potential profitability that may occur at any one moment. Hence, the elimination of this uncertain factor will result in an increase in profitability in the short-term which will extend to that point where either all inefficiencies have been exhausted, or until demand has switched to a previously uncertain degree, and the value placed on the old process is now diminished.

Secondarily, as alluded to above, the elimination of any new uncertainties breeds the conditions for an eventual decline in profits. Consumer wants are not only constantly shifting in often unpredictable ways, but they are developing in as yet unknown ways. Without uncertainty-bearing entrepreneurs correctly steering the firm to provide satisfactions aligned with consumer demands (both known and as yet unknown) there will be no continued profit opportunities to exploit. However, as there will forever be a class of entrepreneurs seeking out new growth directions, this does not necessarily mean that consumer wants will go unsatisfied. Instead it means that a firm that fails to move into these new directions will fail to have the potential to maintain its future profitability and hence, growth. As a result, profitability in the present will begin to suffer accordingly, despite production processes which may have fully eliminated risky occurrences and hence are operating at full efficiency.

Hence we see the importance that the uncertainty-bearing entrepreneurs have concerning the firm's growth potential. In fact, we can say that this type of entrepreneur is the true engine of growth within the firm. To the extent that they steer the firm to new, uncertain territory, they continue allowing risk-mitigating entrepreneurs the ability to continue maintaining, and increasing, profitability within the firm as a result. These two roles combined become important when viewed in light of these two interrelated factors – growth and profitability.

Remuneration and Role

Wages for employees are typically defined as being conditioned by employers bidding up these prices to the point where marginal revenue product is still positive after accounting for wages paid. While determining what exactly an individual's marginal revenue product is a problematic reality, our new theory of the firm, coupled with our dichotomy of entrepreneurial actions, sheds new light on how firm specific wages are assigned. For it may be noticed that there is a general tendency within firms for what are commonly referred to as high-level managers to earn more in wages and benefits than their more lower-level counterparts. This general tendency is, however, just that – a tendency. For we see there are examples where those employees who we would typically refer to as lower level earn in excess what higher-level managers, or even owners, are afforded. We may make two distinct comments on this generalization.

The first is that, when we look at upper level managers, we are really speaking of a specific manifestation of the uncertainty-bearing entrepreneur. These are the individuals who shape the future direction that the firm is heading, and hence, its future opportunities for growth. As a result, the impact that these individuals have on *potential* firm profitability is immense. They drive the firm into new directions and set the boundaries that lower-level employees may profit within. As a result, when viewing the remuneration that these individuals are paid, their marginal revenue product is dependent on the growth potential they provide the firm with. As the firm is considered a continuing entity, this can be a substantial amount, as it will compound (although at a diminishing rate owing to time preference) over many years into the future.

However, when we look at lower-level employees, what we are mostly referring to are those risk-mitigating entrepreneurs. While their marginal revenue product is more easily determined than upper-level managers, there are several caveats that we may add. The first is that they will have a maximum established marginal revenue product, which is defined as the value they can contribute which is defined at any given time by the amount of uncertainty that the uncertainty-bearing entrepreneurs can introduce to their daily tasks. Hence, they are bound by the absolute profitability that the upper-level managers may steer the firm towards. Second, we see that their wages are typically confined to a shorter period of time than uncertainty-bearing entrepreneurs. The average worker may only affect profitability in a static sense – through mitigating the risk that exists at any one time – and hence, differs from the managers who bring the firm continually into new directions. As a result, the wages paid to this lower-level class of employees will be commensurately smaller to account for this difference in time periods that will be under assessment concerning profitability of the firm.

We may see one important point becomes clear when looking at the differential between employee wages. Risk-mitigating entrepreneurs are necessary for the firm to remain profitable in the short-term while the uncertainty-bearers move into new, unchartered territory. As a result, the future viability of the upper-level managers depends on the profitability provided by the lower-levels in the short-term. Wages will reflect this accordingly. Take an example of a start-up company of only one entrepreneur. This individual has a vision of a product for the Internet, which they expect offers high levels of growth potential. The problem that arises for this individual is that, although they excel at looking to the future and seeing opportunities for growth, they have few of the skills to actually produce a product as they envision is demanded. They are ill-equipped at mitigating the risk within the closed-ended system that they have created. As a result, they have hired a computer programmer to create the program as needed. We may see that the computer programmer will often times, in cases such as this, earn a wage in excess of that of the initial entrepreneur. It is evident that the programmer is shouldering no uncertainty, and that they stand to lose no capital from investment. However, their role is crucial as without their input, the uncertainty-bearing entrepreneur would never be able to realize a profit in the future, let alone have an opportunity for their growth potential to be unleashed.

We see that both roles of entrepreneurs are necessary within the firm, and that their wages are affected accordingly. Uncertainty-bearing entrepreneurs will generally be paid a higher amount as their wage is conditioned by the future growth potential that the firm will experience as a result of their efforts. In contrast, risk-mitigating entrepreneurs are paid that portion of the potential profit that they actually realize. As their role is mitigating static risk, they are remunerated accordingly as a reflection of the static contribution they make to the firm's profits. However, as uncertainty-bearing entrepreneurs depend on the risk-mitigators for longevity of the firm, we see that the crucialness of the lower-level entrepreneurs can become a significant portion of relative wages, as the ability they have to keep the firm profitable in the short-term will allow upper-level managers the ability to seek new avenues for growth in the future.

Conclusion

The firm is a synthesis of the two fundamental entrepreneurial functions undertaken in an attempt to form the pure entrepreneur. As the risk-mitigation qualities of some entrepreneurs are mixed with the uncertainty-bearing attributes of others, we see that a composite entrepreneur can be created contractually as a firm. No one individual can be endowed with these qualities – the ability to manage risk and the foresight to navigate uncertainty – to a perfect extent. This gives rise to the advantages that can be had through combining these in a firm-like setting.

Some may argue that this cannot explain a firm at its emergence. Although it is true that these qualities may explain combinations of individuals that would benefit a firm after its initial emergence, as Coase (1937) demonstrated, in the absence of transaction costs, these parties could contract each others' services on the market as needs arise. This is an erroneous assumption to make as we find that the world we live in is one marked by transaction costs, however, there are further arguments that demonstrate why this argument fails to apply to our firm.

Knight (1921) reckoned that the services offered on the market by entrepreneurs were of a non-contractual nature. Hence, there arose an impossibility to solve these coordination problems through the market directly, giving rise to a firm to internalize these contracts and avoid the contractual problems. More recently, Sautet (2000) has argued similarly in more general terms, pointing out that human labor is non-contractual owing to its unique ability to be both specific and non-specific at the same time. Garrouste's (2002) argument as to the internalization of learning benefits offered by a firm further strengthen our case. As learning is a process, benefits accruing from this process would be lost if a series of market based contracts replaced internal contracts that structure a firm. By having employees work together, the learning process may flourish further, and the benefits accruing from this may be kept private by the firm.

Thus, our theory of the firm has some specific advantages over prevailing theories. First, no presumption is made as to technological considerations. Firms are not seen as necessary developments to economize on production processes, nor are they seen as merely profit centers. By doing away with these assumptions, we may make a more general statement about the firm's origins. At the same time, we make no prior assumption as to the existence of markets prior to a firm's emergence. Firms and markets depend on each other in a complex of ways. Instead, the pure theory of the firm outlined here relies only on the assumption that there are heterogeneous agents of differing abilities to navigate uncertainty and mitigate risk. At the same time it is assumed that there is *a* consumer want which is currently unsatisfied. These two assumptions are based on the reality that we see every day through our own actions. Life is defined by a continuing series of wants to be desired, and as Salerno (1994: 114) points out, heterogeneous agents are what enable our society to progress continually. Contra Friedman (1953), our realistic assumptions aid in developing a theory that explains a reality that exists, not one which exists only within the confines of our theory. Transaction costs can carry one's theory so far when explaining the firm's existence, but, when we realize that non-negative transaction costs are a reality of the market, they become a *non sequitur* concerning the theory of the firm. For what difference would it make to demonstrate that the firm would not exist in a world of no transaction costs if we realize at the same time that such a world is an impossibility?

The entrepreneur has been a long neglected factor concerning the theory of the firm. Recent literature has done much to rectify this, but has been based upon an incomplete definition of the entrepreneur. Only by viewing entrepreneurial ability as a dichotomy of functions – risk mitigation and uncertainty bearing – coupled with the impossibility that these may be embodied fully in any one individual, it becomes clear that a firm is an attempt to synthesize these two functions together to create a *pure entrepreneur*.

III. STRUCTURES OF ACTION

1. Introduction

When we left book II, there were two structures of action that were developed and that were shown to rely on one another to function. The structure of consumption was that series of actions within which individuals use consumers' goods and services to satisfy their wants and desires. Owing our time preference, the structure of consumption has four trade-offs that temper its behavior. The first is the commonly cited trade-off between consumption in the present vs. consumption in the future. Consumption that is offset until a future date is referred to as savings typically. The second feature was that some wants can be satisfied to a greater magnitude in the present via some consumers' goods than others. The example was provided of the want of transportation, which could be satisfied through any myriad of ways – from foot, to bicycle, to car. Third, we saw that some wants can be satisfied by goods with a long duration of serviceableness. Hence, there may be two consumers' goods, which provide equivalent services in the present, but which are expected to have differing *durations* of serviceableness or useful lives. Lastly, we looked at the rate of depreciation on consumers' goods. As all goods will depreciate in value with use, we find that the rate of depreciation can be altered, and hence, be manifested in the structure of consumption.

These four points become crucially important when we realize that the goal of all action is want satisfaction. Hence, when we assess the structure of production, we must realize that changes in that structure only result, or are *derived from*, the structure of consumption. The consumptive actions, and the goods that they are manifested through are the alpha and the omega of all action. They place in motion the series of actions which produces the consumers' goods desired. Later they end that relevant action as they actually consume the goods that were previously produced.

The structure of production was reassessed and redefined in terms of series of actions. This replaced the logical inconsistencies that resulted from using the structure as a series of "stages" of production. However, the reality of the situation is that while production does fundamentally boil down to a series of individual actions, as we saw in the previous chapter, these actions are amassed together within an institutional framework called "the firm." Taken this way we may notice that the structure of production may, as a simplifying step, be viewed as a series of individual firms intertwined in a structure of production. Taken even one more step removed from the tenet of methodological individualism, we may say that these firms are broadly grouped together into common industries, which we may see as roughly parallel to the "stages of production" that is so popular today.

It may be asked why we abstracted back to the fundamental individualism of production if only to shrug it off several chapters later. Two comments would bear making at his junction. The first is that the refinement of the structure of production was a necessary step to open up the possibility that an off-setting consumptive structure may exist. It was only be reverting to our individualist roots that we could realize the full potential that the structures of action have. Secondly, we will see that we are not shrugging off the structure of production as a series of individual actions. In fact, if the previous chapter was any indicator, we are adopting the individualism of action even more fully than before, and will remain doing so throughout the analysis. It was only by looking at the particulars of entrepreneurial action that we were able to identify the true source and nature of the firm. In the same way, it is only by keeping in mind the individualism of action that we can continue building a theory of pricing financial assets that have long been removed from any concept of methodological individualism. Hence, we may continue building the micro-foundations of finance through this manner.

We have showed that there is a trend towards equilibrium prevailing in the market. Mrydal (1939: 22-23) showed that monetary equilibrium must rest on the intersection of saving and consumption being in equilibrium, and simultaneously having an equilibrium between production and consumption. Although Mrydal is commonly given the credit for introducing these conditions for "monetary equilibrium" to obtain, it was Rosenstein-Rodan (1936: 273-274) who had earlier identified the issues with greater exactness. While equilibrium is shown to stem from both the "income" side of the equation (savings equates with sustainable consumption), and the "production" side (production of consumers' goods must be offset by production of capital necessary to produce these goods), these conditions are fulfilled through the *monetary* interest rate. As the monetary rate is the visible manifestation of the originary interest rate that ultimately determines it, we find that it is the rate that individuals and entrepreneurs must use to co-ordinate their plans – both concerning production *and* consumption. Hence, Rosenstein-Rodan demonstrates two fields that must be analyzed to determine if equilibrium will obtain from a given interest rate.[307] The first is through the production function – the marginal productivity of capital must equate with the *natural* rate of interest. The second is on the capital market itself – where investment must be offset by the demand for savings (also at the market rate of interest). As both of these equilibrium situations depend on a market rate of interest equating the

[307] Actually, he had identified three fields necessary for equilibrium. The third was the stabilization of the consumer price level, a condition which has since been abandoned by all neo-Wicksellians (see Wicksell 1936: 274).

natural rate, there are obvious opportunities for misalignment to occur. However, when we couple this knowledge with the prevailing tendency towards equilibration of these rates, we see that although these factors may be in disequilibrium, sometimes even for extended periods of time, that there will always be a trend eventually over-powering these dis-equilibrating forces to cause these conditions to converge.

2. Savings and Investment Equilibrium

The first condition that must obtain is the eventual equilibrium between savings and investment. In a world dominated by direct exchange, there would be no misalignment of these terms. As goods would be exchanged against one another, savings and investment would be related one-to-one. Hayek (1941: 409) shows us that money acts as a "loose-joint" that can disrupt this equilibrium. Indeed, as Shackle (1967: 290) tells us:

> [I]n a barter system, or one where money serves only as a *numéraire*, knowledge is *effectively* bound to be perfect. For nothing can be sold except by the concomitant purchase of some other resource-embodying thing. Without money, we cannot put off deciding what to buy with the thing we are in the act of selling. If we do not know precisely what use a thing will be to us, we are compelled nevertheless, by an absence of *money*, to override and ignore the ignorance. It is *money* which enables decision to be deferred.

Hence, in a barter system (one which is defined by a moneyless general equilibrium setting) money may serve no role other than as a unit of account – a measurement device. However, the appearance of money allows not only for decision-making to be deferred, as Shackle brings to light, but also to undertake decision-making before it is possible. Money, in is wider-definition to include credits or fiduciary media, allow investment to be undertaken prior to the savings being available for such actions to occur. Hence, not only may an equilibrium be off-set by deferring investment to a later date through money, but investment may be undertaken at a *sooner* date than would otherwise be possible.

A second dimension of this is the time-period that savings are available to be invested for. Savings are made available in an absolute quantity in the present, but they are also offset for a definite *period* of time into the future before they are expected to be either used for consumption, or repaid. Hence, there is a time dimension that must be matched in order for savings to equal investment.

Any equilibrium that obtains concerning savings and investment must rely on these two dimensions. In the present moment, savings must be available to wholly finance investment. This means that the natural interest rate must be equivalent to the more visible monetary interest rate. Second, we see that the time period of savings must be equal to the time period of investment. The temporal structure of savings must coincide with the necessities of the investment world. Once these two conditions are met equilibrium will obtain concerning savings and investment. However, given the more likely occurrence of a monetary rate of interest continually diverging from the natural rate, these two factors will instead be more likely found to continually trending towards one another *in the long-run*, despite what may be substantial short-run divergences.

Present Savings Equals Present Investment

Total investment may only be sustained by an equivalent amount of total savings. If we were to take a situation of direct exchange, we would see that only by renunciating through exchange a good could another good be purchased in its place. Hence, the first renunciated good is the savings part of the equation (by electing to exchange it you have forgone consuming it), and the second good obtained is the investment part of the equation. Money, acting as its loose-joint, allows for this simultaneity to be disrupted. Through the use of credit we see that present investment can exceed present savings. In Myrdal's terminology, we find that the two magnitudes – savings and investment – need not be equal but only conditionally so. The condition is that in the future there must obtain a trend towards these two magnitudes to equilibrate.

In fact, in equilibrium models we see that this condition is always met. As money is excluded from appearing, there must always be prevalent an exchange economy where present goods are saved and invested in equal magnitudes.[308]

However, there is a temporal dimension to savings that needs to be assessed at the same time. This structure of savings, then, becomes the important factor that requires attention, not the static or absolute measure that exists at any given time. Future liquidity is what concerns the monetary economy, not present equilibration between savings and investment.

The Temporal Dimension of Savings

Savings are never just blindly made, they are always made for a reason. Garrison (2001) notes that savings would be better addressed by the term "saving up for something" (SUFS).[309] However, although noting that savings *implies* this temporal limitation (i.e., savings only exist for a finite time period), Garrison's analysis concerns savings in a very static way. Hence, for example, he employs a

[308] See Davidson (2002: 77) on this point.
[309] This concept was, however, elucidated 100 years prior by Böhm-Bawerk (1901).

loanable funds model solely concerned with the static availability of savings, but ignoring the future availability thereof. In fact, Garrison uses the concept as illustrative to show the incentive that producers have for increasing productive capacity when savings are high – high savings rates signal to producers that consumers are only "saving up for something", and that this something is increased future consumption.

SUFS is a useful concept when we look at the yield curve on savings. Time preference dictates that some individuals will undertake savings in the present for a shorter projected period of time, and others for a more extended period. Hence, we may see on a yield curve differing time preferences ranging anywhere from the short-term (i.e., one month or less) to very extended savings periods of (conventionally) thirty years.

Banks are generally in the business of finding parties of mismatched maturities, and acting as intermediaries in this process. Hence, we may see that there are individuals who generally prefer to save for the short-term, and at the end of this term "roll-over" or reinvest their savings continually. Alternatively, there are individuals who would like to borrow for longer-term projects, and these requirements may not be met by savers who have similar wants concerning their savings durations. Banks may intermediary and take on the risk of reinvesting funds to ensure that, for example, one-year loans may be continually rolled over thirty times to cover a thirty year investment project.

Although there are many considerations that need to be heeded concerning this role (see Bagus and Howden 2009a; 2009b) we may assume a dynamic equilibrium setting where there is no mismatching of security duration – i.e., all short-term borrowing is able to be rolled over continually with no ensuing mismatching issues arising.

In such a case, we see that there still is a possibility for investment to exceed savings in the present and be sustainable in the long-run. This is achieved through economic growth.

Economic growth is the result of a positive return on investment. More investment can be undertaken in the present than current savings allow for by allowing for a greater *future* amount of savings through growth. Hence, let us assume the following one-period example. There exists \$10 of savings of one year duration and that the expectation, and reality, of the future is that each year \$10 will be rolled over as savings. At the same time, due to credit, there is \$15 of investment undertaken which is expected to last for one year. A bank acting as an intermediary has loaned \$15 for one year, while only having \$10 available this year. However, it becomes clear that a discrepancy has developed. For, investment will occur using more savings (\$15) than actually exists in the economy (\$10). Savings become negatively affected as this higher than warranted investment utilizes resources that are not really available in the economy. That is to say, resources are used which ordinarily would not be determining the pure originary rate of interest. At the end of the year when the loan is to be repaid, there will be a funding short-fall. This unsustainable boom is a characteristic feature of what is typically referred to as an Austrian business cycle.[310]

However, there is an offsetting process which is often forgotten. For the credit expansion is problematic *only* (bounded by some conditions as we will see) if the credit expansion has made use of more resources than will be made available through successful investment of these real savings can return over their lifetimes. Hence, a one-period investment may be undertaken utilizing more savings (via a lower interest rate) than would be the case otherwise. This is unsustainable except in the case that the investment of these excess funds can yield a growth rate higher than the original inflation rate. Hence, numerically: if an investment project is undertaken with loaned funds, at an inflation rate which implies there is 5% more real savings available in the economy than reality would suggest, then at the end of this period, the unsustainability will only become problematic *if* the investment project cannot counter this inflation rate through its own return. That is to say, if economic growth can be maintained above the rate of inflation, than no unsustainability need occur.

This statement may seem controversial, and requires some qualifications. It becomes apparent that what is required is not mere economic growth to offset the inflation rate, but a savings rate that is increased to provide real savings to offset this creation of credit. Hence, the rate of economic growth (g) must increase such that:

$$g_i > \Pi_i - s$$

whereby: s = savings rate in the economy,

Π = rate of inflation prevailing over period i

It follows that provided growth is high enough to offset the rate of inflation of the same period, that the effects can be mitigated. Having a high savings rate will necessarily aid this process, as the higher the savings rate, the lower the necessary rate of growth in the economy will be required to off-set inflationary effects. Of course, the savings rate assists in a secondary way as higher savings are also the

[310] On Austrian business cycle theory, see Bagus (2007), Garrison (2001), Hayek (1933; 1935), Hülsmann (1998), Lachmann (1956), Mises (1934; 1949), and Rothbard (1962; 1963).

path to high rates of economic growth.

The preceding analysis may seem as though the effects of inflation can be fully mitigated. However, this would avoid problems arising from Cantillon effects that would prevail under these inflationary periods. As distribution issues arise during the inflationary period, systematic unsustainable imbalances become apparent throughout the economy. For instance, if, as follows from our example above, producers are the first to receive the fresh inflated money supply, they will be the first to spend it at prices reflective of the old money supply. As a direct consequence, they will gain at the expense of those who will receive the money later (i.e., typically fixed income earners and the non-investing working class). This process will sow the seeds of its own bust in due time, but will be mitigated and delayed through the economic growth afforded through successful use of the inflated credit-derived funds.

Conclusion

Hence, for a monetary equilibrium to obtain between savings and investment, two conditions must be met. The first is that savings equates investment in a static sense. That is to say that we need production to be undertaken to be consistent with the consumption/savings preferences of consumers. Hence, consumers choose to forgo present consumption and save, which makes available funds to be invested. This is the typical assessment that is given through the loanable funds model popularized by Garrison (2001).

Secondly, we see there is a temporal dimension to savings as well. Savings do not exist in a static sense, but individuals choose to make the fruits of their renunciated consumption available for a period of time. This gives rise to a time dimension of savings which must be satisfied. For production cannot be undertaken in the present relying on a level of future savings which may not obtain. If a case exists where present investment exceeds present savings, we have an issue where the long-term sustainability can be jeopardized as it is unknown whether the future savings will be available when needed to meet the savings requirements. There are two ways that this process can be made *more* sustainable, although, as noted earlier, issues such as Cantillon effects will give rise to distribution problems which may prove unsurmountable to overcome.

The future rate of savings may increase which would provide for the looming short-fall that will occur from investing more in the present than real savings allows resources to sustain. Alternatively, if the investments can yield a return high enough to offset the inflated money supply, negative effects may be avoided. Under this scenario, provided that economic growth can allow for a high enough *future* rate of savings, there will be no disequilibria between these two variables – present and future savings availabilities.

We may generalize the conclusion as follows. Investment cannot proceed at a rate higher than savings under static conditions. However, once we introduce the temporal elements, both problems and solutions develop. One problem is that through credit inflation, more monetary savings may be introduced into the economy than real savings can sustainably maintain. This implies that investment may exceed *real* savings in a dynamic environment. However, this situation may prove sustainable given the after-consumption rate of economic growth is high enough to offset this credit inflation rate. This is to imply that if future savings (after consumption has been allowed for) are high enough through economic growth to offset the rate of inflation (that is, the rate of "savings" made available through the credit market in excess of what "real" savings exist in the economy) any future unsustainability may be mitigated. Given ancillary issues such as Cantillon effects, it remains unlikely that this would succeed for any extended period of time. However, over short-time periods, we cannot exclude the possibility of economic growth given rise to exogenous increases in the savings rate.

3. Inter-Production Equilibrium

The structure of production was shown to be fundamentally based upon individual value-adding actions modifying capital to produce consumers' goods. Realistically, however, we have seen that production may be seen as occurring at a firm level trying to incorporate individual capabilities to synthesize a *pure* entrepreneur. The typical way of looking at the structure of production has been in terms of "stages of production". We would now like to return to this concept of stages to see what equilibrating trends exist within the structure.

Inter-stage Profit

Specific stages of production each can be seen as having distinct profit rates. If we focus within a specific stage – retail for example, we see that firms all obtain differing profits, but that they generally group around a rate that is similar depending on the specific stage. Hence, we may notice that utility companies (i.e., electricity, water, etc.) have an average profit rate of, for example, 5%/year, and that biomedical research companies yield 7%/year. As we look towards any type of equilibrium obtaining, we see that these profit rates will tend to converge towards one another over time. In the ERE, as we saw in part I, chapter VII, all stages, indeed all processes, will have an equivalent profit rate of zero.

We may use Garrison's use of five specific stages as instructive for showing this equilibrium, and its tendencies. Hence, five stages – mining, refining, manufacturing, distributing, and retailing – all will be seen to have the same rate of profit in the ERE. Also assumed at this point is that each stage is the same temporal length, and that the rate of interest has also reached an equilibrium where it is equal for all temporal lengths. We may summarize these results in the following table 1.

Profit and Interest in Equilibrium			
Stage	**Monetary Profit**	**Interest Rate**	**Entrepreneurial Profit**
Mining	7.00%	5.00%	2.00%
Refining	7.00%	5.00%	2.00%
Manufacturing	7.00%	5.00%	2.00%
Distributing	7.00%	5.00%	2.00%
Retailing	7.00%	5.00%	2.00%

Table 1

Hence, all stages result in the same monetary (i.e., accounting) profit, and given an equal interest rate across all time horizons, the net entrepreneurial profit is identical across all stages. This results in the removal for an individual to shift their efforts from one stage to another as all are yielding the same entrepreneurial profit (i.e., that made in excess of the interest on capital).

It is this entrepreneurial profit which becomes so important at this point. If an individual took their money and did nothing directly with it, they could save it and earn the pure rate of interest. These savings would then be used by other entrepreneurs who borrow it to yield a higher return in excess of this pure interest rate, which we call the entrepreneurial profit rate. The differing spreads in these entrepreneurial profit rates demonstrate to alert entrepreneurs where production can be undertaken to yield a higher return on investment.[311]

In the real world, however, we find that the interest rate and the monetary profit rate are never constant. In fact, there is a general tendency for the interest rate to increase as the time horizon increases also. This result occurs as increased uncertainty surrounding repayment of the original sum of savings forces borrowers to pay a premium for borrowing funds of longer duration. However, despite these differing interest and monetary profit rates, it is still possible that the entrepreneurial rate can exhibit equilibrium. Look to table 2 for one such example where this equilibrium condition is met.

Entrepreneurial Profit in Equilibrium			
Stage	**Monetary Profit**	**Interest Rate**	**Entrepreneurial Profit**
Mining	8.00%	6.00%	2.00%
Refining	7.00%	5.00%	2.00%
Manufacturing	6.00%	4.00%	2.00%
Distributing	5.00%	3.00%	2.00%
Retailing	4.00%	2.00%	2.00%

Table 2

In this example, we may notice that the interest rate declines as the time period declines. The stages of production on the right side of the table are in descending order from those taking the longest time to complete, to those requiring the least. Hence, we see that the interest charge for these longer production

311 As Huerta de Soto (2008: 52) notes, it is possible that entrepreneurs can be earning accounting profits, while simultaneously making entrepreneurial losses, "if accounting profits fail to reach the amount necessary to exceed the implicit gross market rate component that applies to the resources capitalists invest during the financial year."

processes (i.e., mining) is higher than that on shorter production processes (i.e., retailing). However, while the monetary profit rate on mining may be 100% higher than that on retail activities, we will note that entrepreneurs will make no rush to change their operations from the lower-order retail industry, to the higher-order mining industry. The reason for this is that the entrepreneurial profit rate is equivalent at each stage. Entrepreneurs can not seek higher pure returns by shifting from one sector to another; inter-stage equilibrium has obtained.

However, we may note a problem with the above example. For in it we have assumed that the applicable interest rate is the interest rate approximately prevailing for the given stages of production. However, what is applicable is not the prevailing rate, but rather the *spread* between the rates that gives the same duration. Hence, it is not the interest rate on a ten year loan which is important. Instead, the proper interest rate to use is the spread between the nine and ten year loans (assuming we are looking at the profit rate for the last year of a ten year production process).

Hence, assume the following positively sloped yield curve is available today, and that the yield curve will not change over time.

Interest Rate Spreads and the Yield Curve		
Year	**Interest Rate**	**Interest Rate Spread**
1	2.00%	2.00%
2	3.00%	1.00%
3	3.50%	.50%
4	3.75%	.25%
5	3.80%	.05%
Table 3		

We can see that there is a significant cost involved with borrowing in the short-term, but this diminishes as the term structure increases. Hence, to lend funds for 4 years in our example will earn 25 basis points more than a corresponding 3 year loan, but only 5 basis points more than the equivalent 5 year loan. The reason is that felt uncertainty increases as time also increases (the future is unknowable), but the difference in these increases diminishes. For example, it is easy to see the large difference between the uncertainty of one year and the certainty of the present. This is what gives the largest spread to the shortest dated loan – the increase from a certain situation to an uncertain one. However, increasing from an already uncertain situation to a more uncertain one entails less commitment on the side of both the borrower and lender. As a result, the spread in interest rates from, for example, a 4 to a 5 year loan will be lower than that from holding cash and a 1 year loan.

Hence, using the same monetary profit rate, constant across all industries, as 8% and the interest rate yield curve as shown in table 3, we can see a much different picture then before.

Entrepreneurial Profit in Disequilibrium			
Stage	**Monetary Profit**	**Interest Rate Spread**	**Entrepreneurial Profit**
Mining	8.00%	3.80%	4.20%
Refining	8.00%	3.75%	4.25%
Manufacturing	8.00%	3.50%	4.50%
Distributing	8.00%	3.00%	5.00%
Retailing	8.00%	2.00%	6.00%
Table 4			

Some may be tempted to say that an equilibrium is obtaining based on the equivalence of the monetary profit rates. However, by focusing on this aspect the true nature of entrepreneurial profit is masked. Entrepreneurial profit is, after all, that which occurs in excess to what may have been done by doing nothing. In these cases, doing nothing would be saving the funds instead of investing them and earning the rate of originary interest.

The case that exists in table 4 is distinctly one of disequilibrium. Hence, entrepreneurs in the lower-order industries (retailing and distributing) are earning a much higher profit rate than those in the higher-order industries (mining, refining and manufacturing). In fact, not only are these higher-order industries earning less than the lower-order counterparts, but they are earning less than the economy average of 4.79%. As a result of this disequilibrium, an entrepreneurial shift will occur that will see competition leaving the higher-stages, and enter the lower-stages. A result will be that the "entrepreneurial profit curve" will level off as these profits equilibrate. This can be seen in table 5 below.

Entrepreneurial Profit in Equilibrium, Monetary Profit in Disequilibrium			
Stage	**Monetary Profit**	**Interest Rate Spread**	**Entrepreneurial Profit**
Mining	8.80%	3.80%	5.00%
Refining	8.75%	3.75%	5.00%
Manufacturing	8.50%	3.50%	5.00%
Distributing	8.00%	3.00%	5.00%
Retailing	7.00%	2.00%	5.00%
Table 5			

The shift into the lower-stages of production has caused the monetary profit rates to change drastically. Now there is a higher gross profit to be made in the higher-stages, at the expense of the lower ones. In the end there is no difference made, however, as these monetary profit rates are only of prima facie importance. For the entrepreneur, what matters is the entrepreneurial profit rate they can make in excess of the prevailing interest rate of similar time duration. Hence, now we see that all entrepreneurs are rewarded the same profit rate of 5%.

This particular example has focused on changes in the monetary rate of interest to bring about equilibrium. Hence we may say that while entrepreneurial profits were brought to equilibrium, this was done at the expense of monetary profits, which are very much out of equilibrium. There is a second alternative that could obtain to bring equilibrium to both the monetary and entrepreneurial rates of profit. The interest rate yield curve could flatten to make profits equivalent at all durations. In table 6 we see just how this may occur.

Entrepreneurial and Monetary Profit in Equilibrium			
Stage	**Monetary Profit**	**Interest Rate Spread**	**Entrepreneurial Profit**
Mining	8.00%	3.00%	5.00%
Refining	8.00%	3.00%	5.00%
Manufacturing	8.00%	3.00%	5.00%
Distributing	8.00%	3.00%	5.00%
Retailing	8.00%	3.00%	5.00%
Table 6			

As the yield curve shifted from a positive curve to become flat, both profit rates became equivalent. This is important as it shows that equilibrium conditions concerning entrepreneurial profits can be brought around in two ways, or in a combination of both effects. The first is through a shift of the monetary profit rates to factor in the effects of a sloped yield curve (either positively or negatively).

The second alternative is that the yield curve may flatten to equate the two profit rates. In either case, the end result is always the same – the equilibration of entrepreneurial profit rates – the only difference is the mechanism through which this is achieved.

Trends Towards Equilibration

Entrepreneurs will always be operating with three goals in mind. The first is to achieve higher rates of profit than their competitors. The second is to move into industries, or businesses, that offer possibilities for these higher rates of profit. The last is the more hidden goal of seeking out profit from new, undiscovered opportunities. A less positive effect manifests when we that see the trend towards equilibration can be upset through exogenous changes in the interest rate. Each of these holds interesting repercussions when viewed in light of the equilibrating trends of entrepreneurial profit seekers we have just looked at.

When we view the structure of production as industry-based, much like more modern renditions have viewed it (see, for instance, Hayek 1935; Garrison 2001 and Huerta de Soto 2006). This is much as we have viewed the equilibrating trends in the entrepreneurial profit rates up until now. As some industries yield a higher rate of entrepreneurial profit relative to others, there will be a tendency for entrepreneurs to shift their efforts from one area of the economy to another. Of course, as this process progresses, there will be diminishing returns to be made. The profit rates will equalize, and the advantages of moving into a new industry will be minimized.

Second, entrepreneurs will continuously seek to earn more profit than other businesses, *in the same industry*. Not only is the equilibrating tendency aiming towards homogenizing profits inter-industry, but this is attempted to be achieved *intra*-industry as well. Therefore we may find two trends operating at differing degrees. The first is the shift that moves firms within other industries to move into more profitable undertakings. The second is that profit rates will be distinct within an industry, and only average out to a number comparable to other industries. As some businesses within an industry see higher profits are possible, a tendency exists for these entrepreneurs to seek out that which they know is attainable. As a result, less profitable businesses will, over time, move towards higher profit rates, and higher profitability businesses will tend to see this advantage reduced over time.

Of course, the above neglects the entrepreneurial role to a large extent that we have previously developed. For as we have seen, there are two sources of profitability ultimately determining the firm's standing amongst its peers. The first is the uncertainty-bearing entrepreneurs shaping the future and directing the firm into new opportunities. These directions effectively *bound* the risk-mitigating entrepreneurs, who then must work within this framework to maximize profits. This is much like the distinction that Huerta de Soto (2004) makes between the concepts of static and dynamic efficiency. Static efficiency is the more typically cited example of maximizing efficiency, through minimizing waste, for a closed-system. In distinction, dynamic efficiency is the outgrowth of new opportunities made available, and hence, an increase in the total possibilities for growth. Individual firms will enjoy prolonged periods of above or below average valuation dependent on these two entrepreneurial roles. In fact, in a static world, one defined with entrepreneurs of only the risk-mitigating category, all profits across all businesses *and* industries would be at equilibration. This is a result of the closed nature of the system that allows no room for new production possibilities to enter. However, once we allow for an open-system with the uncertainty-bearing entrepreneur, we see the possibility now available for prolonged periods of under and over-valuation as individual business conditions warrant.

However, despite these equilibrating tendencies, there is one alternative that we have not yet accessed. The role of interest rates until now have been endogenously assigned, implicitly through the time preferences of the individual actors. However, if we open the possibility for an exogenously determined interest rate, we see that the rates of entrepreneurial profit – those monetary profit rate adjusted for the corresponding interest rate – will be upset in the balance. For instance, a flattening of the yield curve may lead towards a flattening in the entrepreneurial profit curve as well (as was seen in table 6) however, this would be a non-realistic way of looking at the underlying opportunities. As the interest rate would not be consistent with the underlying time preference desires of lenders and borrowers, the ensuing profit rates would be unsustainable, and hence, irrelevant for judging where capital needs to flow to equilibrate the *true* entrepreneurial profit rate.

Hence, manipulations of the interest rate exogenously imposed on an economy serve to disrupt the profit signal given to entrepreneurs in a way not previously explored by economists. As the entrepreneurial rate of profit will be affected by such manipulation, changes in productive activities will result as entrepreneurs seek out sectors and firms in the economy that maximize their pure entrepreneurial profit. As a corollary then, to the extent that the monetary rate of interest is an accurate reflection of the originary rate, entrepreneurs will be able to direct their resources inline with the demand of the needs of the economy.

4. Conclusion

Equilibrium between consumption and production requires two sub-equilibria. The first was the equilibration between savings and investment. This is demonstrated on our structures of action as an equilibration between savings (on the structure of consumption) with investment (on the structure of production). However, due to money acting as a Hayekian "loose-joint" we see that these will not be equivalent but *merely* conditionally so. What this implies is that the conditional equilibrium will be found under the tendency for investment to equate after-growth savings. Hence, if the return on investment is positive the resultant economic growth may be enough to provide for high enough *future* savings to allow for sustainable production, *despite a higher rate of investment than current savings allows for.*

The second component of the monetary equilibrium was the trend towards inter-industry homogeneity in profit rates. This would satisfy Rosenstein-Rodan's (1936) criteria of an equilibration of the marginal productivity of capital. Entrepreneurial profit is the measure that guides this tendency.

Entrepreneurial profit has been defined as that which rewards the entrepreneur in excess of of the interest yield they could have made by loaning their available savings to others. Hence, firms make an entrepreneurial profit that will trend towards convergence across industries. A secondary effect is that business profits within an industry will also converge towards a common average value. This convergence will always be towards entrepreneurial profit, and not the more visible, or more commonly compared accounting profit. The true measure of profit for an entrepreneur can only stem from that which they have earned in excess of what would have been earned anyway. Only by looking at the yield in excess of the interest rate for a given time period can an entrepreneur's true contribution in adding value be assessed.

These two factors give rise to a tendency along the structures of action – consumption and production – to trend towards equilibration over time. Hence, in general terms we see that the rate of savings will equal investment (or more accurately, the rate of savings in the present plus that made available in the future due to economic growth must equal investment) and that *entrepreneurial* profit rates *across* industries and even of firms *within* industries must trend towards equilibration.

These viewpoints ignore the influence of uncertainty-bearing entrepreneurs who over time may consistently provide returns higher than the average suggests they should. Hence, excellent entrepreneurial ability may consistently yield returns beyond what the average states they should. It is only in a static state of that this profit eliminating trend would exist (or be eliminated by definition). The existence of the uncertainty-bearing entrepreneur brings new light to the theory of production by opening up the possibilities of extended periods of time with above average profit rates.

IV. VERTICAL INTEGRATION

One of the unfortunate consequence of using Hayekian triangles as they were depicted post Garrison (1978) is that that which is commonly referred to as vertical integration (VI) is diagrammatically shown on the triangle as a horizontal phenomenon, and that which we refer commonly to as horizontal integration has no specific place on the diagram. Hayek's original portrayal of the now-famous triangle was rotated 90 degrees vertical from what we commonly see today. This was important for many reasons. The first was that with time placed on the *y*-axis, it was given a role fundamentally shown to be outside the reach of entrepreneurs. With the temporal element shifted to the *x*-axis, an implicit assumption is that it is the control variable. Time, as we saw in Part I, chapter VI, is an element whose passage we must accept as a uni-directional progression, which we may only control to the extent that we can utilize resources sooner rather than later. The second important factor served by placing time on the *y*-axis is that VI, as we like to define it today as businesses integrating across a wider range of conjoined production activities, could be accurately depicted as *vertically* expanding the structure of production.

Vertical integration is a commonly used concept, but can be refreshingly developed further by incorporating the concept of the structure of production. This can be demonstrated in two ways.

The first is by showing that VI involves incorporating “stages of production” in the form of businesses that add value as an integrate whole along the whole productive structure. Hence, this essentially becomes the question of to what extent a firm's boundaries should extend. The ancillary question becomes: “What productive actions are best left to the external market, and which should be incorporated into the firm's scope of business?” Both of these questions can be answered within the context of our previously developed theory of the firm (see book III, chapter II), and illustrated by using the structure of production as an heuristic to see how these distinct resultant businesses integrate to form the larger structure.

Secondly, we can shed new light on the concept of “synergy” which is an oft-cited example of a method to add value to a firm. Firm value is often cited as being a result of synergies which occur by incorporating distinct functions together. Of course, this is a two-way street. Not only can value be added by joining previously separate production process, but aggregate value can be increased by separating processes which are better off independent of each other. Firm value can be increased or decreased as its degree of vertical integration is viewed by individual investors as being more or less optimal than it could be.

There are, of course, several limitations (or more correctly difficult issues to overcome) concerning the concept of VI. For there are two types of VI that can occur – one static, and one dynamic; the boundary between the two is, however, blurry. There is a productive structure that exists in a given state. There will be, therefore, opportunities to maximize its efficiency and return on investment by integrating into a single business any number of disparate productive processes. This is the essence of Coase's (1937) argument – firms internalize the costs of using the market pricing mechanism, and can instead undertake these productive processes within their own boundaries and forgo the cost of using the market pricing mechanism. Hence, to the extent that these internal costs outweigh the external costs of the mechanism, it will be advantageous for a firm to integrate these productive abilities. One specific failing of this viewpoint is, however, that it is very static in nature; ignorant of future changes. For it may be true that we may approximately determine the costs and benefits of using each respective production direction mechanism – via commands internally, or market prices externally – however, this assumes that all production processes are known in their entirety (or at the very least, determinable). This essentially boils down to a question of whether the risk-mitigating entrepreneurs can function with increased efficiency *in the aggregate* given a set of known potential firm-boundaries.

The dynamic alternative of looking at this process (indeed, by definition the only way of looking at *a process*) yields much more uncertain results. There are productive processes that are not known to exist yet; they exist merely to the extent that the entrepreneur can envision them. Vertically integrating then becomes a question of a firm undertaking a production process which may not necessarily exist in any extent along the productive structure. This is a glaring omission of Coase's original analysis, and the majority of transaction-cost theorists that followed him. This uncertain future level of integration becomes of question of entrepreneurial foresight; foresight through the uncertainty-bearing entrepreneur's eyes. This category of entrepreneurs will build a new level of integration which was not previously available within the structure of production.

Lastly, we may want to look at a method that individuals can use to synthesize firms of optimal vertical integration. By purchasing separate firms, individual investors can begin to synthesize a firm of optimal level of integration that is not possible on the market as a unique business. Of course, under this scenario the benefits of internalizing 100% of the decision-making process may be lost due to the ensuing costs of using the market. However, these may be reduced for the individual investor as any profits that exist in one of the firms in question from dealing with the other can be internalized by the investor's portfolio. Hence, the demand for a firm's products and the profits from its supply thereof can be held at the same time by a single individual. To the extent that an individual holds more than one security for this reason, it will indicate that they feel that they can vertically integrate a firm better than the management is currently doing.

Vertical integration has conclusions that can be not only illustrated through the structure of production, but extended when coupling this tool with modern asset pricing theory. As the advantages and limitations are fully explored, we come one step closer to developing a more thoroughly developed theory for valuing our actions.

1. The Structure of Production and Vertical Integration

Introduction

When we look at the structure of production, as is more commonly depicted with production time on the *x*-axis, we see that vertical integration implies a broadening of the scope of a firm's activities *horizontally*. Despite this terminological discrepancy, we will continue to use the more accepted term "vertical integration," or VI for short, even through we wish to illustrate it with a horizontal extension along our structure.

Originally, we had "repaired" the structure of production to account for the fact that when we refer to production, we do not imply previously defined "stages" but individual actions adding value to an intermediary capital good, also known as circulating capital, which will eventually yield to consumption as a completed consumers' good. However, we see the reality of the situation is that, although production is reducible to individual actions, these are often contained within a shared institutional setting – the firm. As a firm arises to combine the two entrepreneurial functions – risk-mitigation and uncertainty-bearing – we see that efficiencies arise due to internalized learning benefits that are an outgrowth, and impetus, of the firm's formation. The ultimate cause of the creation of the firm, even in a world with zero transaction costs contra Coase (1937), was the non-contractible nature of labor, much like Knight (1921) forwarded. Hence, these individuals must join forces in a distinct legal setting in order to combine the fruits of their labor meaningfully.

This theoretical contribution has done much to forward an answer to the question "why does the firm exist." However, it has done relatively little to answer the question, "what are the firms boundaries and how are they defined." by combining the concepts of the structure of production, with our previously developed theory of the firm, we see that there are distinct advantages by being integrated vertically to a specific degree. There are, however, also costs involved which need to be assessed. To the degree that the benefits outweigh the costs of vertically integrating, we will find a firm undertaking an increased number of production processes under the rubric of its own boundaries.

Stages of Integration

Businesses, or firms to use more conventional usage, are collaborations of multiple actions by distinct individuals that further the production process. However, it may have been implicitly assumed to this point that the combination of these actions must necessarily fall into similar categories of goods' production. For instance, a firm may be defined as a minimum of two individuals – one risk-mitigator, and one uncertainty-bearer. These two functions need not exist wholly in isolation of the other, each may be held in some degree by both individuals. However, owing Ricardo's law of comparative advantage, we see that two individuals with varying degrees of *relative* superior or inferior degrees of these abilities will be advantaged to combine forces to specialize in what each is *relatively* best at.

The simplest example of this two-agent firm can be provided as follows. There are two individuals: X and Y. X is a hairstylist who can cut hair with a high degree of precision in a way that is valued quite highly. However, X has no foresight concerning what styles are in demand, or what new directions should be taken to further, or at the very least maintain, her business. Agent Y knows no technical knowledge about styling hair. However, Y is very astute to looking into the future to see undiscovered possibilities. The two agents – X and Y – join together to form a firm where X ensures present profitability by styling hair as per a defined set of rules (styles), which are chosen by Y with the goal of reaching new territories and increasing profitability over time.

In this example it is easy to see that both parties are essentially engaged in the same set of goals – cutting and styling hair. However, it is possible that a more complex range of actions is demanded by each individual in order to increase their potential profitability. On the structure of production this example, styling hair, would comprise one of a plethora of broadly different stages – that stage concerned with cutting and styling hair. However, as we shall see the possibility is much broader than that.

The production process that we have just described could take on many more stages under the same umbrella of a unique "firm." Assume that the uncertainty-bearing entrepreneur who is the director of the firm's future direction decides that advantages could be had by expanding the scope of operations to other related fields. Hence, scissors are the most important tool that the hair stylist uses, and as a result, the decision is made to expand operations to hire an individual to make scissors. Now there is an additional risk-mitigating entrepreneur. As directions are given as to the type of scissors that would be optimal, it is this individual's task to make this good. On our structure of production, we can now see that an additional stage is occupied within the confines of one firm. Not only is the final consumer service stage being fulfilled by the firm, there is also a manufacturing stage which now is engulfed within the firms boundaries which adds to the complexity, as well as the final product.

There are multiple reasons why a firm may choose to vertically integrate to a higher degree. Two in particular are of particular relevance when we look at the concept within the framework of the structure of production. The first deals with the uncertainty that the actions performed in another stage, and which are essential for the firm in question, will be continually delivered as desired by the firm in need. Hence, it could be that the supply of scissors is uncertain for two reasons.[312] First, it could be that competitors threaten the free supply of these scissors which are vital for the firms' continued existence. Vertically expanding is one way in which it can be ensured that the supply of scissors will be used first by the firm for its own ends, and only that surplus production made available (or not) to other competing firms. Second, it could be that the quality delivered by existing suppliers of this good is inferior to what could be obtained if production was included by the same firm. Hence, advantages are to be had by incorporating this stage of production into the same firm's structure.

In fact, the situation may be taken further. We may assume that now the same issues arise concerning those stages of production that concern the production of scissors. Perhaps steel with which to make these scissors is deemed so important that *its* production is deemed essential or advantageous to internalize within the firm's boundaries. The process can continue along the whole structure of production as distinct processes are internalized within a single firm's boundaries.

Hence, we see that firms are stages of production which internalize the processes that the market could perform separately in a distinct manner. However, as we shall see, there are unique advantages, commonly referred to as synergies, which result from collecting distinct stages into one rubric.

Synergies and Value

Given the previous two main reasons for vertically integrating – control of certainty and increased quality of product – there are valuable insights to be gleaned that are mostly consistent with today's concept of synergies. Of course, these synergies can run both ways. There are advantages to having a multitude of joined processes under one firm's control, however, these advantages can soon turn to disadvantages if they extend beyond a firm's capabilities – namely, its human capital.

The Coaseian concept of transaction costs to using the market mechanism give distinct

[312] We use "uncertain" here as a synonym for "risk." True uncertainties we would not be able to identify in advance. Uncertainty is implied in this case as there is no way for us to determine the necessary probabilities that future supply problems will ensue or not.

advantages to internalizing certain productive processes within the firm's internal command structure. A brief numerical example may help to see the value added through integration.

Assume two companies exist, a retail store (R) and a pencil supplier (S). Store R purchases pencils from supplier S and sells them to the general public.

Profit from the pencil supplier is equivalent to the proceeds from sales less the costs of production:

(4.1) $$\Pi_S = Q[P_R - C_S]$$

whereby:

Π_S = firm S's profit
Q = quantity of pencil's sold
$P_{R\,=}$ price paid by firm R for each pencil
C_S = cost per pencil incurred by firm S

Firm R on the other hand will have profit defined as:

(4.2) $$\Pi_R - Z = Q[P_C - P_R]$$

whereby:

Π_R = firm R's profit
P_C = price paid by consumers for each pencil
Z = search costs of using the market price mechanism to determine P_R

In the example we have assumed the market clears, with the quantity of pencils produced equilibrating the quantity sold. As a result, two conclusions become clear in this example.

The first is that the two firms operating separately earn a collective profit of: Π_T= Π_S + Π_R- Z. The search cost incurred by firm R has been lost to the market (i.e., through advertising costs, or employee time in searching for a new supplier of pencils for example).If the two firms integrate into one firm spanning more productive stages, the cost Z of searching can be internalized and hence, total profits will increase. However, some might say, why would firm S agree to this, as their profits seem to be unaffected by the integration or not.

This question deserves three comments. The first is that firm S can be said to also incur search costs, although we have negated them for purposes of our example. Hence, firms actively engage in advertising to ensure their product has an ample market. By integrating Firm S will also eliminate the need for this marketing and advertising expense as all sales will now directly be made to Firm R, regardless of the amount of advertising. Second, we can see that even if Firm S incurred now ancillary marketing expense, Firm R may be in a position to offer more to purchase S than would normally be the case if its own profits are expected to be increased commensurately after the purchase through the increased synergies. Last, note we have placed a confining assumption on our model by assuming that markets clear. In reality, production and sales may not be aligned so clearly, and as a result Firm S may be in the position of losing money if they produce above the market demands, and have an unsaleable surplus. In this case, this form of risk is eliminated as if production takes place within the firm and no unsaleable surplus will be produced. Production of pencils will proceed as commanded by the higher managers, who would utilize all available resources as produced.

The market will place a higher value on firms that are "optimally" vertically integrated. As costs to using the market price mechanism (in a Coaseian sense) are greater than the costs to internalizing this same method of direction of production, a higher valuation will ceteris paribus be placed on that integrated firm as compared to an identical set of distinct firms that direct production through the market price mechanism.

However, there is an important corollary to this value-adding procedure. For if there was nothing but advantages to be gained from this practice, then the logical conclusion would be for a firm to continue integrating until it was the whole of the productive structure. Instead, we find that there are important costs to vertically integrating too broadly, which serve to slowly remove value from the overall firm. These costs come from two main directions. The first is from within the firm, as costs of management become greater than the savings of internalizing. The second comes from outside the firm, as the internalization of customers within the firm effectively removes a competitive component which may lead to reduced competition and leave the firm open to competing innovations from other firms.

Vertical integration stems from the benefits of an internal command-based direction of resources that has costs lower than using the external market price mechanism. These costs involved with using the internalized approach are two-fold. The first is that employees must be directed to produce optimally. That is to say, the risk-mitigators must be directed towards the fixed means-ends framework that they can optimize. This relies on managers being able to accurately identify the framework in question and inform employees about this effectively. Secondly, this assumption relies on a cost of information flow, essential for directing employees, to be lower within the firm than outside on the market. In markets operating with high degrees of efficiency, there is little reason why internalizing this information flow would result in cost savings over allowing the market to direct resources through its price mechanism.

The forces from outside the firm stem from the loss of "feel" towards the desires of the customers. As the "customers" in the traditional sense that existed on the market are now internalized, there is no longer a competitive bidding process directly acting to serve these needs. As a result, a loss of attention to the final consumers' needs *may* result as these competitive forces abate. The cause is the positioning of an internalized process into a fixed means-ends framework of the firm, which is drastically different from the open ended market which defines its actions as a separate entity, competing for customers. Hence, in our previous example, the sole customer for the scissor maker is the hairstylist. The scissor makers lone task is now to meet the needs of this one role. However, if the scissor maker is a separate entity, then they find themselves directly competing not only regarding prices on the open-market, but also in innovation and quality against the other competing scissor makers. This internalization of extra processes risks losing this attention to the the dynamic rubric of external needs of customers, in exchange for more optimally satisfying the static set of needs of the company's internal customers.

As a result of these considerations, firms will have an optimal level of integration *at any one time*. Those firms which are valued more highly than those counterparts that are non-optimally vertically integrated. The ultimate cause of this is that there will be cost savings to be obtained if a firm can integrate productive processes which lower the cost of using the direction mechanism to produce. The trade-off between the costs of the internal command-based direction mechanism, to that of the external market price-based mechanism, will dictate to what extent a firm should vertically integrate.

Capital and Vertical Integration

Previously we discussed two elements of capital that are encompassing within the structure of production – complementarity and substitutability. Both of these factors have an important role to play in the productive structure which are imperative for the concept of vertical integration.

We saw that complementary capital are those individual capital goods which are essential, or value-adding, when added together to produce a good that is one step closer to a saleable consumers' good. The concept of vertical integration is that exact process with which we take distinct capital goods that we perceive to have a positive degree of complementarity. Thus, as different capital goods are found to be producible via different patterns of complementariness, vertical integration structures and the extent of the firm's boundaries will be shifting.

More importantly is the concept of capital substitutability. This characteristic is only apparent in economies of disequilibrium. As equilibrium generally implies a lack of entrepreneurial error, there becomes no need to have substitutable goods, as there will be no requirement for this action. However, given the uncertainty dominating the productive structure, there is a positive degree of disequilibrium operative at any given time. Firms may vertically integrate to battle this uncertainty through substitutable assets. Hence, there is uncertainty concerning the supply of one good on the productive structure. To the degree that it is believed that the risk of this good not being available for the firm in need, the process to create it may be internalized to avoid the uncertainty that it will not be made available in the future. The uncertainty that the good will be supplied in the future is substituted for the risk that it will not be beneficial to integrate this process into the firm's productive structure.

Limitations of Static Integration

Vertical integration tries to maximize profits through reducing the cost of directing resources. However, this is a distinctly static way of looking at the process, which can yield some erroneous conclusions. For while comparing the costs of directing resources internally compared to using the external pricing mechanism, the future changes in both of these relative costs is ignored. The assumption that the production process is known at any given time excludes the possibility that future changes in this structure will obtain. Or, alternatively, it assumes that the productive structure may change but in a determinable way.

If we seek the answer the the question “what level of vertical integration is optimal in a firm?”, we must accept that if the method we answer the question is through a comparison of costs – those of directing resources via command versus prices – that the outcome will be highly static. The reason is that we will be forced to make one of two assumptions.

The first is that the production process is complete as is, and is not expected to change in the future. In this case, the question of vertical integration becomes one concerning at what level of performance the risk-mitigating entrepreneurs can best fulfill their duties. For if by integrating further they can continue to perform optimally, benefits will accrue as a result. The problem becomes one of mere maximization given known resources.

The second case is that the production process will change (admittedly more realistic than the first), but that it will change in a predictable manner (somewhat less realistically). It then becomes a similar problem to the above maximization issue, however instead of being based on known, objective data, the decision is based on entrepreneurial forecasts. To the degree that these forecasts are coherent with the actual world that obtains, the level of integration undertaken at any one moment may be approximately optimal. However, to the extent that true ignorance concerning the future state of the world is unknown, extreme difficulties will be encountered with vertical integration decisions.

Limitations of Dynamic Integration

We may see, then, that when we look at VI under uncertain and dynamic conditions, the uncertainty-bearing entrepreneur becomes essential. This was the omission of Coase's original analysis, as well as transaction-cost theorists that followed in his footsteps. It is the task of the uncertainty-bearing entrepreneur to see the future state of affairs that will exist and adjust the vertical integration structure accordingly. It is only by incorporating this entrepreneurial aspect into the decision-making process that the more realistic, and fruitful, result of decision-making under uncertainty can be obtained.

2. Portfolios of Synthesized Integration

Up until this point we have assumed that individual firms were the only entities that search for an optimal level of vertical integration. However, we see that there is an alternative method that this may be achieved. Individual investors may *synthesize* more optimally vertically integrated firms by purchasing equity in more than one firm. In this way, *some* of the benefits can be internalized by the individuals, while some will be lost. To the extent that the benefits are deemed appropriate by the individual, the decision to hold more than one company for this purpose will be undertaken.

Hence, we may find that there are two separate firms on the market. One is a hairstylist whose main factor of production is a pair of scissors. There is a separate company that sells scissors to the stylist. They are separate and distinct entities. These companies have individually decided that they must remain separate as it is in their own best interests to do so. However, what if an individual external to these companies believes that this is a wrong decision *concerning profitability*, and would like to see both companies join to realize the efficiencies expected through the integration.

The individual may synthesize such a company by purchasing equity ownership in both companies and holding a portfolio that combines the two together. In this way, at least some of the benefits can be internalized, while others will not be possible unless the companies physically extend their boundaries around one another (either thorough long-term contracts, or through a merger).

The main advantage that an individual sees by forming this portfolio is that the profits the scissor manufacturing firm earns by selling to the hairstylist now accrue to the individual. The hairstylist is the necessary demand, and the scissor manufacturer the supply of the productive factors. In this way, it is clear that the individual enjoys the benefits that stem from holding both determinants of the aggregate profits – supply and demand.

However, it should be clear that an individual will not be able to influence, or internalize, some potential profits just by merely purchasing the two firms' equity. For instance, there will still be search costs involved with using the market price mechanism. Advertising and marketing, for example, will still damper total profits that could otherwise be available. Likewise, the learning benefits that stem from having entrepreneurs legally functioning under the same firm (Garrouste 2002) will be lost under such a situation.

It becomes clear that the maximum potential advantages can only be fully exploited by a firm physically altering its boundaries. However, the drive to try to grasp *some* of these benefits is available through holding a portfolio of complementary firms. The end result that the portfolio holder is aiming for is to try to synthesize a perfect entrepreneur. Remember that the perfect entrepreneur is one who mitigates all risk perfectly, and bears all future uncertainty in advance. As no single individual is endowed with these qualities, it becomes the function of the firm to combine entrepreneurial talents to synthesize this pure entrepreneur. We have assumed that only company management has the ability to proceed with this process, and all others are left helpless to independently synthesize the entrepreneur to its full extent. However, we see that this is possible, in some degree, by combining companies that are suspected not of fully exploiting their opportunities to do so. As company management is not fully exploiting these opportunities, as perceived by company outsiders, the door is open for portfolios to be created to grasp these latent profit possibilities.

3. Conclusion

We have seen that the benefit of incorporating the structure of production is to assess and illustrate the concepts of vertical integration and synergy. As the structure of production is assigned an *x*-axis of productive actions, we see that vertical integration involves joining multiple productive actions into the same firm's institutional structure. Thus, the secondary question concerning how far a firm's boundaries should extend becomes a question more fundamentally seeking to answer the question: “To what degree should a firm vertically integrate?”

Synergies become apparent owing to the concepts of complementarity and substitutability of capital. Modern finance theory has already pointed out that synergies exist which make it profitable for firms to enact mergers or acquisitions. To the degree that a firm is vertically integrated to an optimal amount already, that is to say, it is using the proper mixture of complementary capital that maximizes profits, it will be seen that no gains exist through further mergers or acquisitions.

However, there is also a case where a firm has integrated too much. Hence, losses on profits are now occurring as a result of a firm's boundaries being stretched beyond its capabilities. In this case, a firm would see less value than the aggregate that would obtain if it were separated and distinct and independent companies resulted.

There are limitations to integrating, which provide several limitations which are intricately concerned with our theory of the entrepreneur.

The first case is that of static vertical integration. If it is assumed that the production process currently employed will also be used in the future, or that it will be altered in a pre-definable way, we can see that the question is a simple one of present profit maximization. Thus, given a set of employees and a known structure of tasks to be completed, the benefits and costs of integrating (or dis-integrating) can be calculated to determine an optimal integration level. Risk-mitigating entrepreneurs are, therefore, provided with a fixed means-ends framework (or at the very least, a fluid framework whose future state is known now), and the question is to what degree this class of entrepreneur can mitigate the risk to increase efficiencies within this framework. This type of integration becomes problematic when we view it within the confines, or lack thereof, of a dynamic world. In fact, this critique based on a static look of integration has plagued transaction-cost firm theorists ever since Coase (1937) commenced the line of research.

The dynamic alternative for vertical integration includes the uncertainty-bearing entrepreneur to a much larger extent. Integration becomes not a mere maximization problem in this instance, but is concerned with future unknown possibilities that require to be foreseen accurately. It is now that the concept of capital substitutability becomes important. As we open the future to unknowable change, a disequilibria setting takes over which makes the case for entrepreneurial error apparent. Substitutable assets will allow for a safety cushion to be used when the firm's production decisions turn out erroneous. Profitability can be maintained at the second-best use that the capital good in question has.

The last concept we looked at concerned the method that individuals could use to synthesize a firm of more optimal vertical integration – portfolio holdings. It was previously implicitly assumed that only firm management could design and strive towards an optimal firm through VI. Instead we see that individuals may disagree with this opinion, and through the purchase of multiple equities at once, synthesize a more optimally integrated firm. Although not able to deflect all the costs that would be eliminated under a singular firm, many of the advantages can be reaped by the individual. Hence, just as a firm aims to create the perfect entrepreneur by bringing together unique individuals with distinct capital assets through integration, an individual may go one step further and construct a portfolio to try to attain this same end.

The structure of production is able to more fruitfully expand our understanding of these two concepts – vertical integration and synergy – and provide a more thorough micro-foundation for the firm.

V. HORIZONTAL INTEGRATION

Although the structure of production, at least as is presented by a Hayekian triangle, has had some theoretical issues which have precluded attention to consumption and micro-considerations, the potential has been available for these to be discovered and added to the analysis. However, the Hayekian triangle has, by design, excluded quantity of production from being analyzed. With the focus being solely heeded to the dichotomy between production length and value, there has been no way in which the aggregate quantity of goods produced along the structure to be analyzed.

By incorporating a third dimension to the structure, that of quantity, we can bring two new theoretical considerations to the forefront. The first is that no longer will production be viewed solely in terms of its value-adding capability, but in its gross productive physical capabilities as well. Second, we will see that horizontal integration may be analyzed more fully.

The typical Austrian rendition of production shows goods passing through distinct "stages" of production, which we have refined to show their true microfoundations as individual actions, with each action adding value over time to create a consumer good. This focus on the value component of production ignores the physical component. For there are two dimensions of physical productivity that must be addressed. First is that not all economies share a similar aggregate size, but some specialize in one area of production (i.e., production) while others specialize in different areas (i.e., distribution, or mining). We previously saw that not all "stages" of production exhibit similar profits rates, and the same will hold true when comparing distinct economies for this same reason. Second is that, as we have seen, production is dominated by firms and not through autarkic individuals. Hence, there will be a greater or lesser degree of monopolistic production (in the sense of a unique producer or greater levels of competition) at any one place on the structure of production. By looking at the static physical structure at any one stage we may see to what degree inter-firm competition is occurring. Furthermore, the difference between potential and actual output will be able to be assessed.

Second is that we may then analyze conditions of horizontal integration. Firms that have a higher proportion of market share can be said to be more highly horizontally integrated. As we have seen in book II, chapter V, one of the goals of production is not only to provide the goods of the correct *qualitative* type that consumers demand, but also of the correct *quantitative* amount. Value can be added to a firm by have a level of potential production available to meet the needs of consumers that is expected to obtain in the future. Firms that have a potential level of output closer to the expected future demands will have a higher value placed on them than firms which operate under a divergence of these two factors.

The structure of production is not a bi-dimensional process, which exists in solely value related terms. Instead, we find that is has an important, indeed crucial, quantitative axis which yields fruitful insights. By developing this further, we may see where an additional source of firm value lies.

1. Quantitative Considerations and Stage Specificity

With the typical Austrian rendition of production as a series of value-adding steps furthering circulating capital forward until that point at which it becomes a consumers' good, a clear lack of attention has been given to that *quantity* of goods which exists at any one time within the capital structure. Two factors become of great importance when we view production in physical terms.

First is that differing extents of physical production will be taking place at distinct stages within the structure of production. This arises due to dynamic profits at different points along the structure, as well as the division of labor which gives rise to a specialization along the structure. In fact, if we were to compare different economies, we would see that although the value component of the structure of production would be relatively constant among all, the physical component would show wide variation. This results from increased globalization in the face of realized benefits through the division of labor.

Second is that a difference becomes apparent between the potential and actual production. As there is some degree of uncertainty concerning the future level of consumer wants, there will usually be a buffer zone of un-utilized productive capacity. This idle capacity will allow producers the ability to increase production as consumers demand greater *quantities* of want satisfaction.

Both these aspects will be assessed in greater detail

Value, Time and Quantity – Adding the z-axis

The two-dimensional structure of production widely used to illustrate the productive process is a wonderful heuristic tool to illustrate value added through human and technological interactions with circulating capital. However, it excludes the physical component, that which we wish to use to demonstrate the total productive output of the economy. In fact, we see that if we focus solely on the value-adding component, distinct economies will show a remarkable similarity. As the focus is placed almost exclusively on the human ability to add value to circulating capital, each economy will have the shared experience that goods are furthered along a structure that culminates with the production of a consumers' good.

However, there is a widely different story if we look at each economy's ability to produce these consumers' goods in terms of its *ability* to do so. In fact, since Ricardo's law of comparative advantage, we have seen that unique economies have an incentive to specialize in production avenues that they have a relative advantage over others at. In this way, we may notice that not all economies are created equally. This will be reflected on the structure of production as some economies will have a quantitatively large "manufacturing" sector, for example, while for others this may be non-existent. Some economies are marked as being primarily agrarian, while others see the bulk of their quantitative capacity in natural resources – mining or refining. By adding a *z*-axis to the conventional Hayekian triangle we may see these differences more apparently.

The importance of this *z*-axis becomes clear when we compare different growth rates of economies. Without this quantitative component, it must be assumed that each sector of production has equal importance, that is to say, has an equal weight with total production, than the other sectors. However, in reality we see that while each economy likely has the same sectors of production, the relative importance of each within its own economy will differ widely. Most African countries, for example, display highly agrarian oriented economies, with agriculture providing a large amount of total output. In distinction, Hong Kong or Singapore offer almost no agricultural activity, instead focusing their economies on banking, distribution and manufacturing.

Therefore, when we assess individual countries' growth rates, we must look not at the absolute number, but see where on the productive structure this stems from. Conversely, when we see a highly profitable sector, we must ask how important the sector is to the greater economy if positive externalities will become apparent eventually. As value growth of the larger economy becomes an important aggregate to look at when assessing valuable individual enterprises, we see that the microfoundation to assessing this growth is found in the structure of production. As we search for economies that will be increasingly valuable in the future, we must look at the individual productive structure to assess how these economies will be able to capture this future profitability in their productive structure.

Potential Production and Uncertainty

We may also notice that there is a difference between the productive capacity of an economy, and that level which it normally functions at. The difference is a cushion which allows individual firms to account for their uncertainty concerning changing future consumers' demands. As uncertainty about the future quantity demand increases, producers may be enticed to keep a larger amount of unused capacity available to meet this demand when the need arises. Economies operating under fairly certain conditions will be able to use little excess capacity as the quantities demanded are thought to be fairly well predicted.

The output gap that exists between actual and potential production exists owing to two factors: future uncertainty and expectations given this uncertainty.

As the production process always involves a temporal element, the fruits of our labor will not be available until some future date. Owing the uncertainty of the future, we can not be fully certain that the demand for future goods will be a definite quantity. Instead, we will have a expectation as to what this future quantity will be. As a result, the productive capacity will be created to allow for these expectations. This implies that at t_0 we undertake creating a productive structure that will yield the approximate quantity of goods we expect to be in demand at t_1. These expectations will generally not be a definite number, but instead represent a range of quantities. For instance, no entrepreneur would say "I expect to sell one hundred shirts next year." but rather they would be more inclined to recognize their ignorance and say: "I expect to sell more than seventy-five but less than one hundred and fifty shirts next year."

This ignorance manifests itself as a result of our uncertainty concerning future events. We do not know with any certainty what the extent of future demand will be. However, we do know that if we have too much potential capacity we will suffer reduced profitability from idle resources. Likewise, we realize that if there is excess demand we cannot satisfy we may lose customers, and hence, profits. The latter can at times be even more devastating than the static loss of profits as customers turn to other businesses that can supply their needs more consistently. It is for these reasons that firms generally operate under a potential productive capacity in excess of what their present forecasts suggest is necessary.

Lastly, as production goods require an allowance to be made for their continual renewal and maintenance, we see there is a sometimes great cost associated with having a large excess capacity. Production goods require this maintenance expense to be offset by present productive activities. As some degree of capacity lies idle, profitability will be sacrificed in the economy as a larger portion of depreciation allowance will have to be funded by a smaller amount of total productive profitability.

Conclusion

Hence, we find that adding a dimension to the structure of production allows us to gain new insights concerning the productive process. One the one hand, we find that growth can not be measured by looking at only the qualitative areas of the economy enjoying positive expansion. Instead, we must look to see how large a portion of the larger economy is affected by these growth and profit rates. The distribution of profits within the economy becomes important when we stress the relative importance of the differing sectors of production.

Second, we have seen that attention must be afforded to the difference between potential and actual production in the economy. Owing our expectations of future demands concerning the output of our productive processes, we will aim to have a certain productive capacity available to be used at a future date. Uncertainty plays a large role as we are never fully certain what this future will hold, instead a cushion is necessary in order to have capacity available to be used when it is needed. However, this creates a paradox. As idle capacity will still require maintenance, it becomes increasingly costly for a firm to have excess capacity not be used at any given point. The interplay, then, between uncertainty of the future quantity demanded, and profitability of the present conditions, will determine to what extent a firm will operate with excess capacity or not.

2. Horizontal Integration and Value

In the previous chapter, we looked at the case of vertical integration within the structure of production as a source of firm value. A similar case can be made concerning horizontal integration. As we have seen that firms must not only forecast the correct qualitative parameters of consumers' goods that will be demanded, but also the correct quantitative magnitudes of these future demands, we see that the degree of horizontal integration completes the production question.

We may define horizontal integration as the degree, or percentage, of production that a single firm has over a single market or good. The degree of monopolization that an individual firm has over production of a product can be equivalently stated as the degree of horizontal integration it has achieved. While there are benefits to be had from gaining a 100% market share of the good in question, there is also significant risk involved with this practice as well.

As profits are always fleeting in the face of competitive entrepreneurs searching for opportunities to exploit, there are costs involved with having 100% of any given market. The advantage of gaining a greater share of the market is a greater share of the resultant profits. However, profits will never be allowed to get out of line with other industries for two reasons. First is that entrepreneurs will continually be placing pressure on existing firms by trying to enter offering higher quality or lower-prices competing goods. Alternatively, on the consumption side of the equation, consumers may shift their consumption patterns to include different goods, at the expense of those that are now deemed to be overpriced.

Ignoring these possibilities, or assuming they are not of consequence, we may see a unique determinant of value stemming from a firm's horizontal integration. That is to say, if we take the demand for a good as given and not subject to significant change in the short-term, while at the same time assume that the firm will not exploit consumers by attempting to charge monopoly prices, the extent of integration will prove to be a major determinant of value.

As was previously explored, the amount of unused potential capacity creates an uncertainty buffer for a firm concerning future demand, but also entails a cost. As firms horizontally integrate, there are two sources where value can be made more apparent.

The first is through having a potential productive capacity close to that of what future expected demand will be. In this way, firms will be able to capitalize on the maximum amount of products available to them. However, this can also work negatively. For what if a firm has more capacity than what the future demand will dictate is necessary? In this case, there will be much idle capacity either being not used (non-productive), or un-maintained (anti-productive). In this case, a firm will see its total value decrease as investment has been undertaken in the past erroneously. It has expanded capacity beyond what will be needed, at the expense that future excess capacity will either need to be maintained (at a cost to profitability), or left to deteriorate (to realize the malinvestment of the past).

Additionally, a firm will be valued more highly if it is operating closer to its full potential capacity. In this way there will be a reduced amount of idle capital in need of allowance for its maintenance and depreciation, which would entail a negative hit on firm profitability. One of the hallmark's of Toyota's success against some of the other major auto manufacturers was its stress on just-in-time inventory. By utilizing this capital management technique, there was less idle capacity removing profitability from the firm. As a result, firm valuation was higher than its competitors.

This analysis has assumed that demand for a product will remain. However, with competitors constantly searching for ways to introduce alternative products, a firm always runs the *risk* that its product will become obsolete. In this way, we can see that becoming a monopolist by horizontally integrating to be the sole supplier of a certain good can be a very risky endeavor. If a firm has 100% of production of a single product and that product becomes obsolete, a significant decrease can be expected in its value. In such a case, we see there are limits as to the extent that value can be added to a firm through horizontal integration.

3. Conclusion

Hayekian triangles, with their sole focus on the value adding activities of production, have lacked attention regarding the magnitude of production. By incorporating a third dimension, a *z*-axis on the typical Hayekian triangle, we may illustrate production not only in value terms but also in magnitude of production in quantitative terms. As a result, production will not only pass through stages of higher to lower order, but will also involve a quantity produced at each stage. There are several benefits of taking such an approach.

The first is that we may now see that economies are wildly different. While almost all economies would exhibit similar structures concerning the individual stages that comprise the structure of production, there is wildly differing degrees of concentration in each stage compared to another. This is important as when we look at growth within an economy, we must identify not only those avenues that are growing and expected to continue to do so, but also the degree to which they are reflected in the larger economy. For what good would the increased affluence and positive externalities of a manufacturing sector growth rate of 20% a year have if your economy is only based on 5% manufacturing for its total output. In this way we see the importance of comparing the extent that individual sectors have within the economy.

Second, we see a trade-off that exists between potential capacity and actual capacity. Firms need to estimate future demand and ensure that productive capacity will be adequate to satisfy this. As the future is uncertain, usually a positive buffer level will be maintained, giving rise to a portion of unused capacity within a firm's output level. This output cushion has a negative result that capital is owned which requires maintenance on its depreciation, but which is currently not contributing to the profitability of the firm. It acts only as an insurance safeguard, to take advantage of future demand that may arise in excess of what is expected.

Although this may be the larger macro implication of including a quantitative aspect to the structure of production, there is a micro benefit as well. As we see that production is dominated by firms, each will exhibit a positive degree of horizontal integration. In this way, we see that firms will either produce all of the goods of their given sector, or a fraction of it. Firms will have value assigned to them depending on the extent to which they are optimally horizontally integrated. As full integration will ensure they take 100% of the potential profits, a high value will be placed on an individual firm that is able to collect on this potentiality. However, as the future demand of a good also has an uncertain aspect to it, there will be that possibility that the good will be obsolete in the future, or at least not valued as highly as was expected. In this event, we see that firm value would suffer negatively as their productive capacity is dedicated towards a good of reduced economic significance.

Adding the quantitative dimension to the structure of production allows us to not only delve deeper into production theory, but also to see ways in which this aspect can affect firm value. As firms struggle to obtain the "optimal" level of horizontal integration, investors will value them accordingly.

VI. TIME AND LIQUIDITY PREFERENCE

When individuals choose a portfolio of holdings, there is a two-fold trade-off that exists. In sequential order, the actor must first make a decision between that portion of income they would like to consume, and that which they would like to save. This is the basic time preference trade-off that Böhm-Bawerk (1889) elaborated on. Hence, individuals may choose to consume much in the present while leaving little disposable income available for future consumption. Such individuals represent those we refer to as having a "high time preference." In distinction, there is that class of individuals who prefer to save (or not consume) in the present, but instead aim to do so in the future. This class of individual we refer to as those who exhibit "low time preference."

An individual must first select that portion of their income that they wish to consume in the present, and that which will be saved for future consumption. A second choice faces the individual after they select the portion to be saved. An individual also exhibits a liquidity preference.[313] As their savings will then be available to be directed in a myriad of avenues, they will have to select portions to be assigned for each degree of liquidity. Typically, we see that three specific liquidity classes exist.

The first are those assets which are fully liquid. Money is this good *par excellence* as it is always expected to be redeemable at a value of unity regarding itself. That is to say, money always sells for the same price as itself on the market. Money serves a two-fold role in our investment decision. First, it is spent in the present as part of the consumption part of our time preference decision. Later, it is held as a guard against uncertainty concerning the future contingencies that may arise.

The second asset class which we hold are those we refer to as liquid. These are generally able to be sold on a well functioning market at a price which is approximately known in advance. Hence, we see liquid assets as primarily those which concern the financial exchanges – stocks, bonds and other financial products. These are, *under normal market conditions*, able to be sold quickly and with little risk of loss due to a lack of demand.

The last class which we wish to categorize are those that fall into the illiquid category. Typically, these enjoy little secondary market. Capital goods are conventionally applied to this category as we see that normally a high-degree of specialization marks these particular goods, and hence, a large difference may be sustained on the resale market when these are liquidated. The difference between the bid-ask spread on these items may be very high owing this degree of specialization inherent in them.

---

[313] See Davidson (2002: 80-85) for a look at this two-step investment process involving time and liquidity preferences.

Hence, the three categories that we select from when assigning our investment decision based on liquidity preference is really a question of specialization. Money is that homogeneous good above all others that is generally accepted as a medium of exchange. Financial assets – that asset class we define as *merely* liquid – are to a large extent homogeneous, however, there is still some degrees of differentiation between them. They will not be able to be sold as assuredly as money with prior knowledge of what the selling price will bring. Lastly, those more durable capital goods represent a category of illiquid assets whose selling price is largely unknown in advance. They are dependent on having a buyer available with similar requirements for the same capital good at the same time as the seller wishes to dispose of their's.

This two-fold investment decision – time and liquidity preference – dictates to what extent an individual's portfolio holdings will be dominated by differing degrees of liquid assets, as well as the absolute magnitude of these holdings.

1. Time Preference

Time preference represents that irreversible fact that we continually must display a positive preference for consuming in the present over the future. However, the *degree* of this positivity is subject to be very much an individual phenomenon. At this moment it will help us to treat the specifics of that portion of income used for consumption, and leave that portion left for investment to a later date. We will see, then, that that portion contributing towards our present consumption will affect our structure of consumption in a myriad of ways that must be assessed briefly.

Increased Present Want Satisfaction

Increased time preference can come by solely preferring *more* consumptive ends fulfilled in the present. This can be brought about in one of two ways (or a mixture thereof). The first concerns the gross quantity of wants that will be fulfilled. Alternatively, we see that a set, or even declining, number of ends may be satisfied in a more thorough and complete manner.

In the first alternative, increasing time preference manifests through an increased number of wants to be quantitatively satisfied. Hence, on a summer day, instead of satisfying our want for refreshment with one glass of wine, we instead choose to have two, or more perhaps. In this way, we can see that we are consuming a higher amount in the present, at the expense where savings will have to be some degree lower as a result.

In the same way, we can see that a second result of a time preference increase is that a higher quality of want satisfaction will be undertaken today at the expense of that in the future. We have seen that ends may be satisfied in a plethora of ways, with differing degrees of satisfaction resulting. While we search for a refreshment to quench our thirst on the hot Ibiza beach during the summer, we may therefore order a domestic Mahou to deliver this desire, or a higher quality import may be chosen – Heineken or Becks, for example.

Hence, increasing our want satisfaction in the present is a product of two dimensions. The first involves the gross quantity of wants satisfied, while the second dimension concerns the magnitude or quality of these want satisfactions. An increase in either will signal, ceteris paribus, an increase in time preference.

Decreased Duration of Want Satisfaction

We have seen that the expected duration of serviceableness of a consumers' good is, ceteris paribus, directly related to its purchase price. As it will yield valued services for an extended period of time, its discounted marginal value will be higher than the same good with a shorter period of serviceableness. The effect on an increase in time preference it to decrease the duration of serviceableness demanded from consumers' goods.

As a general decrease in the duration of serviceableness of a consumers' good will free-up additional resources (i.e., money) to be used in the present for other want satisfactions. As an example, we can compare two houses which only require an allowance for the maintenance of their roofs. A roof made of slate has an expected duration of 100 years, while a less-expensive roof of asphalt shingles may have an expected duration of only 15 years, but will require much less initial outlay. Ceteris paribus, a house could be built with a roof of asphalt of a much larger size than if one was made of slate. If we assume that size is directly proportional to want satisfaction, we may see that the larger house with an asphalt roof will satisfy a higher degree of wants in the present, albeit for a shorter duration of time.

Decreases in time preference have the opposite effect. They entice us to purchase consumers' goods of longer expected duration of serviceableness. We sacrifice other want satisfactions in the present by paying a higher price for these goods, however, we enjoy their services into the future for a longer period of time.

Increased Durability of Consumable Goods

The durability of a good also becomes effected by the time preference of an individual. We have also already seen that more durable goods will require a higher initial outlay of resources to obtain. The choice is between paying upfront for a good of greater durability, or paying later as an allowance on depreciation. All capital goods have a finite life span.[314] However, this can be lengthened if a continued allowance for maintenance is made.

As more durable consumers' goods will require a larger initial outlay of resources for attainment, we see that decreasing degrees of time preference will have the effect of increasing the

[314] Although, in many cases this life span can be longer than the life of the acting individual. Houses, for example, are able to remain in use by different people over many generations.

durability of consumers' goods, ceteris paribus. By purchasing more durable goods in the present, we are afforded a the good fortune that we will not be required to continually maintain these goods to the extent that is necessary with less durable goods. Increases in time preference will have the opposite effect of decreasing the durability of consumers' goods. Under these scenarios, we wish to have an increased amount of present want satisfaction by paying (renunciating) less for these goods in the present. As this is achieved by utilizing less durable goods, we see a general tendency for these two factors – time preference and durability of consumers' goods – to have an inverse relationship.

Increased Quantity of Consumers' Goods

Lastly, we may look at that category of time preference effects that concerns the total quantity of consumers' goods demanded. This may seem to be the same case we explored previously concerning the increase in present want satisfaction through higher time preference. However, this category is different only in that it involves the purchase of more consumers' goods than one has time available to utilize.

Hence, we see that an individual may own more than one automobile. Two or more may be purchased to satisfy wants during an individual's free time. However, we see that time is a limiting factor as an individual may only drive one of these cars at a time. Hence, the ownership of more than one seems to be problematic for time preference theory – why would someone purchase more want satisfaction in the present than they can possibly consume?

The answer lay in the risk surrounding the consumers' good's availability. An individual could theoretically contract this good on the market at will, however, risk surrounding two factors – its availability and price – may provide impetus for the individual to purchase it in the present, despite not having adequate resources available to utilize it.

Hence, an increase in time preference does not place a limit on the accumulation of consumers' goods. Instead, owing the increasing risk that another individual will make use of the good before they get a chance to, an individual may decide to stockpile additional consumers' goods before they are taken by someone else on the market.

Asset Holding and Time Preference

Besides these four considerations concerning conventional time preference and investment decisions, there is an additional component which receives scant attention. The previous four sections have looked at the goods' side of time preference, specifically, how changes in the time preference of an individual affects their spending decisions on consumers' goods. However, there is also a way in which time preference effects the savings scales of individuals in heretofore unseen ways. Typically, savings are demonstrated as a uni-dimensional factor, comprising a homogeneous quantity available in the present. However, as Garrison (2001: 40) points out, there is also a temporal dimension concerning the availability of these savings in the future:

> People do not just save (S); they save-up-for-something (SUFS). Their abstaining from present consumption serves a purpose; saving implies the intent to consume later. SUFS, our unaesthetic acronym (which we will resist employing repeatedly throughout this volume), stands in contrast to the conventional distinction between "saving," the flow concept (so much per year – from now on?) and "savings," the corresponding stock concept (the accumulation of so many years of saving – to what end?). Saving in capital-based macroeconomics means the accumulation of purchasing power to be exercised sometime in the future.

Hence, savings are not just undertaken in a static moment, never to be used again. Instead, we find that savings are undertaken for a period of time, albeit one that is ill-determined at the time that the individual is saving. It thus becomes clear that savings is not only of concern in the present, but its availability in the future is of importance as well.[315]

An individual's time preference will alter over time. As a result, that ratio between savings and consumption will be altered continuously, or at least cannot be assumed to be constant, and the resultant issues will have to be assessed. Two conditions dominate others in importance.

The first is that liquidity issues surrounding changes in time preference will have to be assessed. Bagus and Howden (2009b) assess some of the problematic results of time preference changes – maturity mismatching. Hence, banks that borrow short-term securities in hopes to make longer-term loans at higher interest rates rely on a future supply of savings to be available at a low interest rate in order to be able to profit from these long-term loans. To the degree that a bank can correctly foresee the

[315] As Davidson (2002: 77) points out, there is a neglect in modern economics of the disparities that may exist between present and future savings. As all contracts are paid for in full at the initial time period, or, a perfect futures market ensures that all settlements are fully offset, there is no room for disconnect between these two components – present and future savings.

future time preference schedules of individuals, this practice may be undertaken successfully. However, if mistakes are made in this process, a liquidity crunch may occur as banks, in this example, are unable to reinvest in new loans to offset the liabilities that they have previously committed to.

The second is that the valuation placed on savings or consumption will be subject to change. Hence, savings may be made available in the future, but the question arises: “At what price?” There is no necessary link between increases in real savings, and changes that result in the price of those savings (i.e., the natural rate of interest). Hülsmann (2008) demonstrates the falsity of this argument by showing that future savings may increase with a raising rate of interest associated with their availability provided that individuals place a higher value on their increased savings.

Conclusion

Time preference changes the structure of two important factors then. The first is the structure of production. By shifting the mix that we demand concerning consumers' goods, we see that time preference dominates ancillary factors in determining the quantity and quality of these goods. Also however, we have seen that the structure of savings available to be invested is also effected by this preference. Typical explanations focus solely on the static, or one-period, trade-off that exists between consumption and savings. However, we see that there is also a temporal element where savings are held for only a limited amount of time. Individuals' time preference is not a unidimensional factor – a trade-off between now and later. It also involves a waiting dimension. The future availability of savings will have drastic implications both for the consumers' goods that will be demanded, and also that availability of funds to be invested, or reinvested, as is required.

2. Liquidity Preference

Once an individual has decided that portion of their income to be dedicated to present consumption, and that to go towards future consumption (thus representing current savings), a secondary decision must be made. This second decision, that concerning an individual's liquidity preference, takes on a binary character. The difference in liquidities of distinct goods represents in reality a two-fold nature. We may assume that no good is fully illiquid; all goods are able to be sold the only question is at what price. Full liquidity, then, involves the certainty that you can sell a given asset and be assured that you will receive your expected price.

Once we realize that there is only one asset that provides this benefit – full liquidity at all times – above all others, we see that the investment decision is with how to divide income between money and other less liquid assets. Money is that asset we hold to combat our fears over uncertainty. As a result, it is a gauge for the amount of *felt* uncertainty that may exist at any given time. As conditions change and make the entrepreneur feel more certain about future events that may obtain, we see that a smaller amount of cash holding will prevail. Conversely, circumstances when uncertainty increase will result in a relatively larger degree of cash holdings on the part of the entrepreneurs.

However, that remaining portion of savings ready to be invested into somewhat illiquid assets will require to be split in some way between illiquid assets, and liquid assets (but not assuredly so). How this split is made is the topic of much contention, as it determines that demand which exists for capital goods (relatively illiquid assets) and financial assets (liquid assets). Previously we have seen that one reason an individual may hold more than one financial asset in an attempt to synthesize a perfect entrepreneur – through the combination of multiple firms. However, why would a trade-off be made between investing by differing degrees of liquidity of assets be of concern, and how should they be allocated?

In book I, chapter IV, we saw that when making a decision, an individual takes a two-step approach. First, they limit the outcomes that they expect to occur to a defined set, thus eliminating the uncertainty they feel in a situation. Second, they assess this bounded situation in terms of risk via a subjective probability distribution. By comparing two resultant focus-points – a focus-gain and a focus-loss – we saw that an individual has no uncertainty in their resultant decision; they expect their projection to obtain. However, there is some degree of risk in the decision-making process. An individual can eliminate the uncertainty they *feel* about a situation (although they cannot eliminate the uncertainty that *exists*) however, they will have to manage a risky element within this *deemed* certain subset. Two facets need to be explored.

The first is that until now we have explored a structure of production phrased in terms of real capital. However, there exists a financial structure of production as well that represents the *claims* to these capital goods. Hence, as value is imputed from the desires of consumers to the real capital that helps them to achieve their wants, we see that the value of financial assets is again derivative of this process. Myrdal (1939) gave us two set of equilibrium conditions that called for an equality between savings and investment and the inter-stage marginal value productivity of capital. Now we can introduce a new equilibrium condition as the profit rates on financial assets must be equal to those on real capital.

Individuals invest their saved funds with the goal of maximizing the expected profit they will receive. One would be led to believe that if returns were identical on illiquid and liquid assets, that the added liquidity would entice individuals to invest solely in this class – primarily financial assets. However, we have seen that liquidity concerning financial assets is always associated with a certain temporal element. It makes no difference what the liquidity is at any specific point in time, only at that point when we wish to sell the asset to redeem it into money. Hence, being always able to sell an asset and have an eager buyer is not as important as being able to sell an asset in the future at a specific date and have *a* buyer available.

Typically buyers of illiquid assets are not concerned with their lack of liquidity. These assets are generally held for the whole useful life and then disposed of – there is little resale market for these assets. As we have seen that an individual undertakes a plan fully expecting that their assessment of the future will be correct, they see no need *ex ante* to concern themselves with liquidity problems. However, if this was the case, why would anyone desire to purchase more liquid financial assets? The two assets – liquid and illiquid – are reliant on each other. Illiquid assets could exist without their financial counterparts, however, the large expense of purchasing these makes it much more likely that they would be shared by several individuals.

Although the commonly cited reasons for owning financial assets are risk reduction or the pooling of resources by distinct individuals, this does not strike the heart of the question as to why financial assets exist. For to answer this question, it proves to be instructive to return to Coase (1937) who asked the same question concerning the firm's existence. In a world of zero transaction costs, there is no reason why all activity could not be contracted on the market without having firm-based organizational structures to internalize these activities. There is also no reason why firms would require financing through external sources, which necessitates the requirement of financial assets to be issued by a firm. In fact, all investment could be financed internally through profits (retained earnings) or loans.

The reason is found in our previous expose of the entrepreneur. We have seen that firms are the result of two classes of entrepreneurs – risk-mitigating and uncertainty-bearing – who are brought together by way of contractual necessity to work in the same institutional structure. Until this point we have presupposed that a firm *can* create a perfect entrepreneur. One implication of this is that a resourceless definition is possible. As the entrepreneurial functions are both combined perfectly, there is no chance of loss occurring – a possibility that necessitates the inclusion of real resources as part of the firm. This assumption must now be shed as we look at how a firm may combine the third entrepreneurial element – resources – under its realm.

At a firm's inception, then, both types of entrepreneurs have to search out the other entrepreneurial class to have the skill set necessary to function competitively. However, if they lack the funding necessary to undertake the venture, they will also be required to seek this requirement out as well. As Rothbard (1997: 247) remarked concerning the entrepreneurial requirement for resources: "I might be keenly alert to a profit opportunity virtually lying at my feet. I may have a sure tip on the stock market. But if I haven't got [*sic*] any money to invest, the profits, perceived opportunity or not, will simply not be made." Issuing financial claims to capital assets is the method used to incorporate this third entrepreneurial requirement – resources – in an attempt to synthesize the perfect entrepreneur.

With the rationale for purchasing financial assets now established, why would an individual purchase, or invest, in a illiquid capital good? The reason is that the additional entrepreneurial roles that one individual can perform increases the pure entrepreneurial profit that will accrue to them. If an individual has it in their power the command of resources sufficient to directly invest in capital goods, they would do so and reap the rewards. The liquidity preference decision at this point is a non sequitur as the entrepreneur expects a purchased asset to be fully liquid *ex ante*. That is to say, at the time of purchase, entrepreneurs will have an expectation that they will be able to sell an asset in the future at a given price. Therefore, they don't forecast an asset to be more liquid than another *ex ante*, although *ex post* the story may be much different.

Money capital is made available to firms through the financial structure that exists – that interrelated structure of financing options that firms have available to them at any given time. Hence, money, or funding more generally, filters through this financial structure into the real structure of production. This make a third class of entrepreneurs that we must now consider – investors. With the provision of the appropriate capital, our two previous entrepreneurial functions – risk-mitigation and uncertainty-bearing – now have the necessary resources to function in an uncertain world. As a return will be earned on these loaned resources, this entrepreneurial class we now call investors will earn an entrepreneurial profit. As this classification of entrepreneurial roles exhausts the analysis, the three returns – wages, salaries, and return on investment – represent the three types of entrepreneurial profit that can exist.

The financial structure of production thus becomes an integral part of the analysis. First we see the trade-off that exists between investing in liquid and illiquid assets is false. *Ex ante* all investment decisions rely on some positive degree of liquidity as the investor thinks their plans will materialize as forecast. Then we see that the choice to invest in that capital we refer to as liquid – financial assets – versus that which is commonly thought of as illiquid – capital goods – is intertwined with the entrepreneurial function. The need for capital through funding implies that a pure entrepreneur entails three functions – risk-mitigation, uncertainty-bearing and resource provision. Suppliers of money capital through the financial market will be rewarded accordingly to the needs of the greater entrepreneurial effort, just as the two other classes of entrepreneurs are. By adding this third component to the entrepreneurial function, we can now complete the entrepreneurial role, and see how this *pure* function is synthesized through distinct individuals on the market.

Money

Money is that asset which we hold to guard against our uncertain future. To hedge against these uncertain contingencies which may arise in the future, money is held. However, we see that value is variable, and as a result, there can be no absolute measure by which we could say that these uncertain contingencies that may arise may ever be met through its use. However, as Hayek (1980: 33) informs us:

> People want, as a medium of exchange, something which reduces as much as possible the uncertainty of future prices. It is inevitable, however, that prices change. They even change unpredictably. The reason for this is that prices are instruments which inform us on events

> about which we have no information and which, by their nature, must be unexpected events. But the uncertainty about future prices can be reduced to a minimum, if the risk of making mistakes in anticipating future prices in one direction is balanced by the risk of making mistakes in the other direction.

Money may be thought of as that asset with the highest degree of liquidity attainable, even though this may not be full-liquidity in the sense we wish to ascribe to the word.[316] The reason why this addition of liquidity is given is that a deferral of decision-making occurs under the use of money. As Shackle (1967: 290) outlines, money is that which allows us to part ways from the more constrictive ways of direct exchange, and hence, delay our decisions for a later date when our knowledge, wants, or means are different.[317]

Money provides us with the highest degree of certainty that we may defer consumption in the present, and place our savings in a means that will be useful in the future. Hence, Garrison's (2001) emphasis on "saving up for something" becomes evident as we place our savings in money in hopes to use it again on consumption goods in the future.

Individuals with high degrees of felt uncertainty will be inclined to hold more of their saved assets in the form of money. This ensures that they will have more adequate means available to cover these uncertain future contingencies as they arise. The corollary is that an individual with full certainty concerning the future would have no need to hold money. As all contingencies would be known in advance, money will be of no use as contractual agreements may be arranged in advance to cover future expenditures. Two conclusions can be drawn from this. First is that any equilibrium setting with a fully certain future will have no use for money. In fact, much like Davidson (2002: 77) makes evident:

> Since Debreau's general equilibrium formulation is the theoretical foundation of all mainstream models, even if orthodox models explicitly have markets for forward delivery, *all payments are made on the spot*... No wonder that the concept of liquidity is given short-shrift

[316] Using this definition of money as the asset with the greatest amount of liquidity aids in removing some of the problematic aspects of defining money's chief characteristic. Osborne (1984), for example, finds ten separate definitions all which try to explain the same phenomenon: money. Ritzmann (1999) provides a history of money as a substitute for confidence. Confidence, in this case, arises as the result of increased certainty regarding events which are otherwise uncertain, such as money aids in achieving.

[317] Of course, this line of thought shares much in common with Hayek's (1941: 409) viewpoint of money as a "loose joint" in the economy.

> in mainstream economic models. No wonder that these models stress a budget line (income) as the sole constraint on spending and ignore the possibility of a more important liquidity constraint on spending.

In any general equilibrium setting, money will play no role. We have already seen in book I, chapter VII how Mises' construction of the evenly rotating economy precluded the existence of a monetary factor. Second, we may see that in the framework for decision-making under uncertainty that we developed in book I, chapter IV, uncertainty plays a prominent role. The inner-range of values an individual considers when making a decision is directly influenced by the degree of *felt* uncertainty concerning a given situation. Hence, as the inner-range will be wider (incorporate more values) if it represents a more uncertain event we may gauge this inner-range through the size of cash-holdings an individual saves. If savings in the form of money are high, felt uncertainty will also be high. Conversely, if cash savings are low, felt uncertainty will also be low. These two points make the link between money and uncertainty more than an exercise in pure theory, but one of real application as well.[318]

However, as Wicksell (1935: 10) has noted: "[F]rom the individual point of view the use of money as a standard of future payments over longer periods is unsatisfactory and incomplete, since the capital saved is not employed in production and thus does not, as a rule, yield any interest." Even though money provides us with these benefits, an individual will not be enticed to hold their whole savings in this form as to do so would forgo the ancillary benefits afforded through less liquid assets.

Liquid Capital

Individuals are never fully uncertain concerning their yet to be discovered future, they are only partially so. As a result, a positive sum of money will always be held, but this will generally not be to the same maximum extent made possible through an individual's savings. A point will be reached where an individual has satiated their search for certainty (or at least has countered sufficiently their felt uncertainty), and they are willing to forgo further increases thereof in exchange for the use of their savings in a more lucrative avenue. This involves sacrificing further increases in certainty through the use of that class of assets which are not fully liquid to the same degree as money. As Rothbard (1978:

[318] Wicksteed (1933: 153) alludes to the fact that money has a unique role concerning uncertainty. Although it may be used to mitigate our uncertainty concerning the future, there is no good besides money which may serve this role directly.

150-51) defines these assets:

> M]any assets are 'liquid,' i.e., can easily be sold for money. Blue-chip stocks, for example, can be easily sold for money, yet no one would include such stocks as part of 'the money supply.' The operative difference, then, is not whether an asset is liquid or not (since stocks are no more part of the money supply than, say, real estate) but whether the asset is redeemable at a fixed rate, at par, in money... The current tendency of some economists to include assets as money purely because of their liquidity must be rejected; after all, in some cases, inventories of retail goods might be as liquid as stocks or bonds, and yet surely no one would list these inventories as part of the money supply. They are *other* goods sold for money on the market.

Hence, Rothbard places the distinction between money and liquid assets as those that can be sold *for* money at some future time.[319] As we are now making a distinction between those non-monetary goods depending on their specific degrees of liquidity, we see there are two specific types.

The first are those which we refer to as "liquid" and define as financial assets. Hence, they are marked by a high probability that they can be liquidated into money without undue loss. Well-developed markets exist trading these financial assets, with the result that buyers and sellers may easily find one another. The claims to real assets, therefore, enjoy a higher degree of transferability than the actual real assets that underlie them.

In distinction, real assets are fundamentally less liquid. They are more specific, and enjoy less developed markets for buyers and sellers to exchange titles to these goods. As a result, if an individual is forced to sell a real asset, there is a higher degree of probability that they will be forced into a disadvantageous position due to the relative illiquidity of these goods.

Davidson (2002: 83-84) notes that there are four attributes of durable assets that are possessed in differing degrees that affect their desirability to be held. 1) the expected rents to be earned from their use (which also includes any dividends or interest) (q), 2) carrying costs of holding the asset for a period (c), 3) the liquidity premium which arises from having to sell an asset during the period (l), 4) the expected spot price appreciation (depreciation) of the asset over the expected holding period (a). Any asset that is fully illiquid will have values of a and l of 0. That is to say, they will be expected to fully depreciate over the holding period, and thus have no resale value at the end. Additionally they

319 Fiduciary media must be included as a type of money, not a liquid asset, as it is never withdrawn from the monetary system (Huerta de Soto 2006: 696fn141; O'Driscoll 1986).

will have no liquidity premium attached to them. That is to say, individuals will not attach any ancillary positive value to these assets owing to the fact that they will be able to be sold into cash quickly and with little probability of undue loss.

Of course, the fully liquid assets – money – will have two significant divergences from these zero values given to illiquid assets. First, the liquidity premium, l, will be equal to unity. Money is that asset which *par definition* is fully liquid at all times. The expected appreciation of an asset over a given holding period will give a positive value to a. As we are unsure of inflationary/deflationary effects concerning the price of money, we will not be able to assign *a priori* a value of 1 to this parameter. However, we will see that this variable will have a greater degree of constancy than any other liquid asset – the value of cash in terms of cash will remain constant compared to other goods.

Hence, financial assets exist that are of greater liquidity than the real assets that they provide claims to. There will exist two distinct structures of assets then. We have already explored the well-developed structure of production, which traces the path of circulating and working capital along the stages that create consumers' goods. However, we see that there is a corresponding structure of financial assets that represent claims to these real goods. At this point, we may extend the two conclusions we made concerning the equilibrium trending conditions that prevail in the real economy.

The first is that the return on financial assets must tend towards a return roughly equivalent to that in the real economy. That is to say, the return on financial assets must be approximately equivalent to that on real assets, given an equilibrium constraint. There is one significant difference that will make these two profit rates differ in non-equilibrium conditions. Note that profitability on real assets – those that comprise the structure of production – depends on decisions made in the past. As assets have to be purchased in the past, the profitability they will exhibit in the present depends on them being arranged in a pattern that is fully conducive to the needs of consumers. Such conditions have previously been explored in book II, chapter V. However, the profit rate on the structure of financial assets will depend on their ability in the future to provide profitable services. This arises as these financial assets represent claims to real assets whose value (but not profitability) depends on those future conditions which were expected to prevail when they were purchased. Hence, the profitability of financial assets will be fundamentally a forward looking estimation, in contrast to the profitability of the real assets which are a more backward looking assessment. Under a steady rate of growth in the economy, these two profit rates would be identical. However, under conditions of a steadily increasing growth rate in the economy, the profit rate on financial assets will outstrip that of their real counterparts. Conversely, a steadily shrinking economy will see the profit rates on real assets higher than those that will prevail on the financial assets, as this second group will depend on the future profits which will stem from a less profitable future.

Second is the consideration that in equilibrium, profit rates between financial assets representing different "stages" of production will tend towards an equilibrium. As a result, a uniform rate of profit (or return on investment) will prevail, or at least have a tendency to prevail, as a trend towards equilibrium occurs. Keep in mind that earlier when we spoke of this equilibrium condition with real assets, the equilibrating tendency was for rates of *pure entrepreneurial profit* to converge. This amounts to a homogenization of accounting profit rates less the appropriate originary interest payment that would be garnered by not partaking in any investment. The same condition applies in this case. However, in this case, the rate of originary interest to be used to remove from the accounting profit rate will differ from the previously stated case. Notice that as the profit on financial assets stems from future profitability, the cash flows have already been discounted at an appropriate interest rate. As this discounting has brought all profit rates back to the same period of time, the interest rate used to determine the pure rate of entrepreneurial profit will be identical for all financial asset holdings. This is a crucially different point when we consider that the pure entrepreneurial profit rate on real assets required the application of an interest rate that was congruent with the time period and duration with which the use of the assets in question (and hence, the use of investable resources that could have been directed elsewhere) to be removed from the ensuing accounting profit rate.

Hence, the structure of financial assets shares many similarities with that of real capital, but many differences as well. Financial assets are forward looking, and their profit rates stem from the expected future profits to be earned. However, in distinction, profit on real assets is fundamentally concerned with the past structure of production – that is to say, how coherent the existing structure of real assets is with the underlying preferences of consumers. However, the same general equilibrating tendencies exist, much like were explained concerning the structure of production in book III, chapter III. Equality between the financial profits between stages in its structure will prevail over time. Likewise a tendency will exist for financial profits to be equilibrated with real profits, although the temporal aspects will differ. By discounting expected future financial profits with the appropriate originary interest rate, we see that the tendency for equilibration between two distinct profits rates – present and future – can be equilibrated to exist at the same moment. However, the question arises: why would anyone invest in a real asset if the possibility of having a greater degree of liquidity available at all times entices them to invest in financial assets? This will be answered in the following section.

The Pure Entrepreneur – Uncertainty, Risk and Resources

Much of the previous considerations of the entrepreneur explained that the pure entrepreneur is a mixture of two qualities – risk-mitigating and uncertainty-bearing. However, we have left the requirement of resources decidedly uncommented on. This has been by nature of the fact that we have assumed a *pure* entrepreneur to be a possibility. Hence, if both entrepreneurial qualities could be met in full, there would be no requirement for resources. As Coase (1937) might surmise, all resources could be contracted on the market, and there would be no need for any real commitments. This arises as the possibility of loses would be eliminated, thus making resources a non sequitur. However, we see that in the real world, entrepreneurs will err and this makes resources a crucial requirement for this role. As such, an additional entrepreneurial role becomes apparent – that of the resource provider.

Entrepreneurs require resources for their plans to materialize. There are two ways that this can be achieved within the setting of a firm. One is through the contracting of loans. A basic requirement of this is that a set duration of repayment is offered.[320] In this case, the entrepreneurial function of resource provision would be supplied for a contracted period of time. After this time has elapsed (or prior if an early payment option is available), the entrepreneurial input of this funding will be repaid, and hence, this entrepreneurial function will be removed from the firm. Second, a firm may issue equity to provide for this funding requirement. The advantage is that this can be done for a *typically* lower cost than a direct loan. The reason for this cost-differential becomes apparent when we look at the implications concerning the entrepreneurial function. For issuing equity involves allowing additional entrepreneurs to partake in the firm's profits for *an unlimited amount of time*. Issuing equity typically involves no specific time limit (although convertible shares may be issued giving an explicit limit to this issuance). As a result, though the issuance of equity often will involve a lower cost than that of contracting a loan, the entrepreneurial profits will have to be shared throughout the whole of the firm's future. Each of these two financing options – contracting a loan and issuing equity – will be explored more thoroughly.[321]

[320] See Huerta de Soto (2006: 3-4) for this argument. Sechrest (2008: 3) has countered that this is not true, citing evidence of mortgages that allow for early repayment schemes as durationless loans. However, Sechrest mistakes loans as existing without durations, with loans that exist with a set duration and an *option* for early repayment. More pertinently, the first does not exist, while the second does.

[321] This addition of a resource providing entrepreneur allows for an addition of a new source of entrepreneurial profit. Kirzner (1973: 54) had already noted the difficulties that exist in trying to disentangle distinct sources of entrepreneurial

Typically, a firm will have a set plan complete with the necessary human capital to perform the operations it wishes. The missing component is the financial capital which will have to be supplied through an external source. If this financial component is contracted through a loan, there are several important points which become important.[322] The first is that the issuer of the loan typically has no direct say in the way that the business is run. Their comments or input are limited to the supply of capital which will be repaid to yield an entrepreneurial profit after a maximum set duration. The terms under which the business will operate are determined in advance by the existing entrepreneurs, and the financing firm will have the decision to accept or reject these terms. Of course the rate of profit on their issued loan will be determined through the interpretation of the production plan, but in general no input will be offered concerning this plan. Entrepreneurs working in the firm may prefer this option as it requires that entrepreneurial profits from the firm's operations be shared with another party for a determined duration of time. Hence, although profits will be *sacrificed* to a third party (the bank) over a period of time, it is known that in the future this will end and all entrepreneurial profits can accrue to entrepreneurs within the firm.

Alternatively, the existing entrepreneurs can raise their necessary financial capital through the issuance of equity to shareholders. There are several points that make this a distinct option to financing via contracted loans. Under this case, equity may be issued to produce the necessary financial capital, however, unlike loans, the duration of this issuance will be forever (i.e., it is a perpetuity). As a result, the share of entrepreneurial profits that will accrue to the providers of these financial resources will also continue (in theory) forever. This becomes unattractive to the existing firm as they will forever be bound to share the profits from a firm's issuance, however it is a very attractive option to individuals who wish to contribute to the entrepreneurial effort, and hence, receive a share of the profits. Additionally, equity *generally* involves voting rights to select management in the company. As a result, this class of entrepreneurs may actively partake in the uncertainty-bearing role of entrepreneur through the use of their equity voting rights. As a result of these differences over financing with loans, an entrepreneur contributing funding via equity will be inclined to *ceteris paribus* accept less in return for their entrepreneurial contribution than their contracted counterparts.

In fact, this ceteris paribus clause concerning remuneration for suppliers of resources can be viewed under the same conditions as we earlier looked at remuneration between uncertainty-bearing

profits, specifically profits belonging to the capitalist-entrepreneurs and pure-entrepreneur. See also Lachmann (1956) for this difficulty.

322 Bonds essentially fall under this category as well. In fact, little distinction can be made between financing through a bond, and that of a contracted loan, expect through the terms of repayment.

and risk-mitigating entrepreneurs. Loan suppliers are paid in terms of an interest contract which is generally fixed at the time of inception and will continue for a predefined duration. In contrast, equity holders will be offered a dividend which roughly equates to the same general concept, however, it may not be a constant throughout the life of the firm. The degree of remuneration that either financing option would receive is relative to the comparative demand and supply conditions prevailing at a given time. At certain times we can see that banks are flush with cash and ready to supply money-capital to needy firms. During these times we may notice that the return to compensate a bank for this loan may be quite low as there is a large supply available. Alternatively, we may see that there are times when there is a large supply of shareholders available to supply equity financing to a firm. At times like this, we may notice that there is little incentive to guarantee these potential shareholders a large share of the entrepreneurial profits that will result.

As a case in point, the period of the late 1990s was marked by the high-tech boom which was defined by expectations of continually increasing share prices coupled with historically low dividend payout ratios for firms. As investors felt that continually increasing share prices would adequately compensate them for their entrepreneurial contributions in the future, a rush to supply firms with monetary-capital gave receiving firms an ample supply which required little need to remunerate them amply through high dividends. As a result, price-dividend ratios fell to historic lows as the ample suply of resource-supply entrepreneurs made this requirement less important (and hence, require less payment of entrepreneurial profit) than the other two entrepreneurial roles.

Hence, we find that the true entrepreneurial role may be divided into three distinct categories. Risk-mitigation and uncertainty-bearing have already been analyzed and the effects their roles play on their respective remunerations have been assessed. With the possibility of loss apparent in a disequilibrium setting, real resources are necessary to complete the entrepreneurs plans. Hence, a third category of entrepreneur – resource-supplier – has been added. The degree to which their services are essential to the entrepreneurial plans will dictate that portion of entrepreneurial profits they will receive for their contributions. In particular, we see that there are two ways in which this resource requirement may be satiated. First, through loans contracted for a specific time duration. Second, through equity issued for an indefinite time horizon. Existing entrepreneurs will, ceteris paribus, prefer to have resources supplied through contracted loans. In this way, the entrepreneurial profits will only be shared with these extra entrepreneurs for a definite amount of time. Alternatively, we find that providers of monetary resources will prefer, again ceteris paribus, to supply these funds via equity with no defined time duration. In this way, they may continue to garner a portion of the entrepreneurial profit for as long as they continue holding claims to the firm's real assets (equity). Hence, the full entrepreneurial role will entail three distinct categories of action – uncertainty-bearing, risk-mitigation and the supply of resources. To the degree that each of these is essential for the functioning of the greater entrepreneurial effort (i.e., the firm) the relative and respective remuneration will reflect this.

Conclusion

The investment decision is two-fold. First, an individual's time preference dictates that portion of their income that will be consumed, and that which will be saved. Out of this saved portion, a secondary decision is made based on liquidity preference. Hence, a portion will be desired to be held in cash – the most liquid of all assets by definition. The remainder will be applied to assets of varying degrees of liquidity as investment.

Money is a gauge for felt uncertainty. Earlier in our model of decision-making under uncertainty, we saw that an inner-range of outcomes exists which an individual deems as being fully possible. That is to say, no degree of surprise will result from one of these outcomes obtaining. The breadth of this inner-range is dictated by the degree of felt uncertainty that an individual experiences at any given time. Hence, under conditions that seem highly uncertain, and individual's inner-range will be quite broad (i.e., encompass many results) compared to that which would result during periods of relatively more (felt) certainty. We have now seen that money is a way we may proxy the subjective degree of felt uncertainty under a potential decision. A higher portion of investable income dedicated to cash signifies that an investor has less certainty about a future situation obtaining. If an individual were to have perfect knowledge about the future state of affairs, they would have no demand to hold investable funds as cash – they could fully contract out this portion of income for the appropriate durations and hence, match their assets with future liabilities.

Typically the remaining portion of investable income (that not held as cash) is said to be invested under the guidance of the trade-off between assets of differing liquidities. However, we see that liquidity cannot be an issue ex ante when the decision-making process is be undertaken. The reason is that the entrepreneur *believes* that there estimation of the future will prevail. For this reason, they only undertake an investment in the first place as they foresee the possibility to gain from it in the future. Liquidity to the individual does not concern any point other than that specific point in time when they expect to sell a specific asset. We see then, that concepts of liquidity based on a general ability to sell an asset at any time in a well-developed market are misguided. To the individual entrepreneur, the only time period that matters it that which occurs when they purchase the asset, and when they expect to sell it.

However, we may see that there is a difference in assets such that in the *general* sense some are more liquid than others. Hence, we have financial assets (claims to goods) which have well developed markets where buyers and sellers converge to deal with each other. The assets that are claimed by these financial assets are the real capital goods which comprise the structure of production. These are generally more specific and hence, enjoy fewer purchasers and sellers in more isolated market. The choice that an individual undertakes is between purchasing real capital goods or claims to them via financial assets.

The entrepreneurial function had previously been defined as only incorporating two classes of entrepreneurs – risk-mitigators and uncertainty-bearers. However, when we enter the real world we see a binding resource constraint makes a third class – resource suppliers – a necessity. As a result, financial assets become inextricably linked with the entrepreneurial role as it provides the means that monetary capital is directed towards firms comprised of the two previously developed entrepreneurs. The return that a resource supplying entrepreneur earns will thus be equatable to the wages that risk-mitigating entrepreneurs earn, and the salaries that uncertainty-bearing entrepreneurs earn. These three categories exhaust the classification of entrepreneurial roles, and their respective remunerations, that exist.

If financial assets allowed for greater liquidity *and* offered the same return as capital goods, there would be no question that all entrepreneurs would purchase financial assets at the expense of capital goods. However, we see that with the addition of entrepreneurs that entrepreneurial profits will have to be divided in multiple ways as well. Hence, if possible, entrepreneurs would prefer to finance the purchase of the real capital assets themselves and hence be the sole recipients of the entrepreneurial profits when they accrue. In reality we see that this may not be possible. The resource requirements of some production processes are so great that external financiers are a necessity. Hence, our third class of entrepreneurs becomes a requirement for a firm to partake in its desired activities.

In general, we have seen that two distinct classes of resource allocators may provide the firm with its financial needs. The first are through contracted loans of set (or maximum at the least) time duration. This is preferred *ceteris paribus* by the existing entrepreneurs as it will limit the amount of entrepreneurial profit that must be sacrificed to an additional entrepreneur. Alternatively, equity financing may be pursued. The main difference this option has over contracted loans is that the duration may be undefined. As a result, the existing entrepreneurs will have to potentially forgo future entrepreneurial profits for an extended period of time. The remuneration for each respective class of financing is through a defined interest rate or a dividend rate, respectively. To the extent that the financing is required and essential to the firm's success relative to the supply that is available on the market, these rates will change to reflect these conditions.

Myrdal's (1939) conditions for monetary equilibrium can be extended to the structure of financial assets we have thus defined. The easiest, or most direct, application of this concept is that the profit rates between stages on the structure of financial will have a tendency to equilibrate. Additionally, we can see that the profit rate between savings (and hence and the structure of production) will tend to equilibrate that of the structure of financial assets, but with one significant twist. For the profit rate on real capital is dependent on those decisions made in the past that must now be reflected in a structure of capital assets consistent with the desires and demands of the consumers. However, the profit rate on the financial assets will depend on a future situation yet to obtain. It becomes clear that these represent two separate temporal periods when they are attempted to be compared. If we wish to compare the future expected profits (via the return on financial assets) with those that actually exist today (via the return on real assets) we must discount those future expected profits back to the present time through an appropriate originary interest rate.

The financial structure of production is an integral part of the analysis. This is so, then, for two reasons. First is that it allows us to see the place where the third entrepreneurial component – provision of resources – becomes apparent in the theory of the firm (itself a component of the more general theory of the entrepreneur). Second, we see that the equilibrium tending conditions that Myrdal first outlined almost 70 years ago are directly applicable to the financial realm, as well as the more commonly considered productive realm. By utilizing both these considerations together, we see that the degree to which these financial assets should be rewarded, that is to say, the degree to which they profit through their contribution, can be pinpointed not only in their present contributions, but their expected future ones as well.

3. Direction of the Real Capital Market by the Financial Capital Market

We have seen that resource funding for production typically enters the productive realm via the financial realm. The two structures that we have previously explored – that of physical production and that of financial assets – are forever separate, but inextricably linked via this temporal process and each other's mutual necessity of one another. Financial resources enter the real productive sector (the structure of production) via the financial structure of production. However, there are limitations on this process. For the only fresh funding that makes its way directly to the productive sector is that which stems from new issues of securities. A secondary and indirect effect stems from the reduction in dividend payments that results from a robust demand for financial securities. However, the base funding that the real sector can utilize can only stem from the issuance of new financial assets.

Given this basic fact, we find that there is ample room for the profit rates to become disjointed between the two realms – at least in the short-run. We will illustrate this via two cases. In the first a real productive structure exists with no external funding available. In the second, the same productive structure exists, but has already utilized a high degree of financing via the financial markets. The disjoint between their respective profit rates will become evident.

In case one, assume that each company in the real economy issues securities for the first time. As investors purchase these financial assets, the funding goes wholly to the real sector. From that point the structure of production may be expanded accordingly. Investors will be rewarded for their supply of resources with a dividend rate for the remainder of the duration that they hold these financial assets for.

In distinction, in case two we see that funding has already been provided to the real economy via the financial structure of production, and that there exists a robust financial market. Now we may assume that new investors with available financial resources wish to invest in these financial assets. There are two options which exist. The first is that companies may satisfy this increased demand for financial assets by issuing new securities. In this case, the new funding will work its way directly into the real economy. However, if this new issuance of securities lags the real demand of investment the fresh funding will be directly applied to the existing financial market. With an inelastic supply of financial assets, the only way that this influx of funding can be accommodated is through increasing prices of these financial assets. Notice than that if the issuance of new financial assets directly by the companies defining the real sector of the economy is inelastic that there will be no direct effect on the structure of production stemming from these newly supplied resources. Instead, financial valuations will increase to reflect this new demand for securities outstripping the new supply thereof.

However, there is a secondary effect that will eventually effect the real structure of production. The cost of financing these financial assets is the dividend yield that a company must pay in order to attract buyers to its securities. Hence, as demand increases with the supply being held constant, the dividend offered may fall commensurately. As a result, even in the face of no direct increase in financial resources, a reduction in the cost of the use of resources will result from the dividend rate reduction. In this way, the two entrepreneurial classes that comprise a firm will have to sacrifice *less* entrepreneurial profit to those who provide the resources (in this case, shareholders).

With a reduced remittance of profits via dividends, a firm can change its role in the real structure of production as more resources are available through the use of a higher potential retained profit rate. Hence, while the real productive structure is altered via changes in the financial structure of production, we see that this happens through a direct way (new share issuance) and an indirect way (reductions in dividend rates).

Although as we have previously elaborated, these two structures have an intricate equilibrating tendency, this can be offset in the short-term due to this relationship. A fully elastic supply of equities would ensure that any new demand for financial assets would be fully met by new issuances. If this were the case, dividend ratios would remain constant as well as security prices. However, in reality the supply of equities is not only inelastic, it is also discrete. Firms issue new shares in discrete intervals and at quantities which do not always align with the demand for financial assets. In this case, we see considerable variations in both prices and dividend yields stemming from the demand side to hold these assets being unmatched by the supply side. The structure of production will be altered through the issuance of financial assets, but not always in a direct way, allowing for short-term disequilibria to upset the previous conditions we had listed for a monetary equilibrium.

4. Conclusion

We have shown in this chapter that a complex decision-making process is involved with the investment procedure. Savings are not automatically invested in financial assets, but a two-fold process occurs. The first step is the renunciation of consumption to arrive at a magnitude of savings that can be invested. The secondary choice is that of liquidity preference. From this sum of savings, an individual must divide between a portion to be held as cash, and a portion to be invested in liquid assets. This will depend on the amount of felt uncertainty an individual is experiencing at any one time. As felt uncertainty increases, money holdings increase as a hedge against unforeseen contingencies. Conversely, situations involving less felt uncertainty entice the individual to decrease their cash holdings.

Historically, it has been thought that liquidity preference also manifests through an individual's want to invest between liquid assets (financial assets typically) and illiquid assets (real capital goods). We have shown this to be a false dichotomy. Individuals, at the point in time they finalize their decision-making process, do not factor for differing degrees of liquidity as they believe that their forecast plans will prevail. In this way, we see that liquidity is not the ability to have access to money or a money-equivalent at any given time. Instead, it is a less general case where it is thought that access to money will be available at a *specific point in time*.

However, we see that given this point of view, if illiquid and liquid assets – capital goods and financial claims to these goods – were forecast to have the same rate of return, there would be no incentive to invest in one over the other. We see that the reason that there is a choice made is inextricably linked with the theory of the entrepreneur we developed earlier. We have discussed the two human-qualities that an entrepreneur must have – risk-mitigation (efficiency) and uncertainty-bearing (growth). By contracting both of these qualities within a firm, an attempt is created at making a *pure entrepreneur*. However, we have, until now, left out an important component of this role – that of the provision of resources. This important role is provided by entrepreneurs in the form of shareholders and banks in the form of contracted loans.

In the case of banks being contracted loans, we see that entrepreneurs, or firms in a more specific sense, may contract the use of resources for a determined amount of time. This is quite advantageous for a firm, as it means that the sacrifice of entrepreneurial profits will be dispersed to the provider of financial resources for only a limited amount of time. However, there is the option available to seek out equity purchasers as well. In this scenario, a loan is granted with no duration. An advantage of this is that the firm will have no requirement to pay back this loan. However, the primary disadvantage is that the firm will have to pay-out a portion of its entrepreneurial profits indefinitely into the future. As a result, we find that generally the return on equity is lower that what would be paid on a contracted loan. However, we also see that basic supply and demand conditions affect the magnitude of the dividend being requested by the shareholders. Under conditions when financing is amply available, the portion of entrepreneurial profits which will have to be repaid to the resource providing entrepreneurs will be reduced. An additional benefit can be had for shareholders if they also have a voting right attached to their shares. In this case they will be able to contribute to two entrepreneurial roles – that of uncertainty bearing through the election of company management, and resource provision through the financial funds they make available.

Lastly, we have seen that changes to the real structure of production flow from the financial structure of production in two ways – one direct, and one indirect. As a supply of new equities is made available, financial resources that are used to purchase these new issuances will contribute funding directly to the firms in question. However, if the supply of new equities is sufficiently inelastic, than new financial resources will not directly be introduced to the real sector, but instead will only contribute to continue pushing equity valuations higher and higher. A secondary effect will happen in this case, however, as the increase in equity valuations will reduce the dividend yield on existing shares. In this way, that portion of entrepreneurial profits which are contributed to equity holders in the way of dividend payments may be reduced. As a result, retained earnings may be increased at the firm level, and the structure of real production can be expanded in this way.

This two-fold investment decision – time and liquidity preference – dictates to what extent an individual's portfolio holdings will be dominated by differing degrees of liquid assets, as well as the absolute magnitude of these holdings. The requirement for resources to be supplied to the entrepreneurial firms entices individuals to invest their funds in hopes of earning an entrepreneurial profit in payment. The complex interactions between the supply and demand of financial resources through savers, as well as the elasticity of the supply of new equities will determine in what way and to what extent shifts in profit rates between the structures of real assets and financial assets obtain.

Appendix A. Time Preference and Money

Introduction

One of Menger's greatest contributions to the science is found in writings on the origin of money. He, concurrent to Jevons' (1876) analysis, viewed money arising from the "double coincidence of wants" problem.[323] Two individuals wanting to engage in trade each hold goods that the other may not want. This single issue gives rise to generally accepted commodities that came to be used not only for their direct use-value, but their exchange-value in trade as well. Hence, for Menger, the emergence of money stems from spontaneously finding a commonly accepted medium of exchange. The end goal would be to find a medium that would reduce the bid-ask spread on goods to the maximum realizable amount. For instance, it is true that every good is tradable to some extent. Approaching a butcher to buy a steak with only a hammer at your disposal for exchange may still result in a trade; the question would be at what price. The discount offered for your hammer may be so great so as to dissuade the trade at all, or place you at a considerable bargaining disadvantage.

While recognizing that money emerges as a spontaneous order (1883: 131), Menger would also note that the precious metals have taken on the role of money, but this is only an historical fact.[324] Hence, diametrically opposed to Friedman's conception, for Menger (1892b) we find that, "[m]oney has not been generated by law. In its origin it is a social, not a state institution."[325] Precious metals were adopted due to their particular qualities; however, these have always been particular historical facts, never time-invariant axioms we could use to deduce where future money is to be produced from.

Mises (1912) would add considerably to Menger's insight with his "regression theorem." The

323 Say (1803) would also write of the double coincidence of wants problem. In fact, Say was one of the primary influences on Menger's thoughts on money. See Rothbard (1995: 37) for a further elaboration on this prehistory. Also, see Wu (1939: 126) for the view that Mises was heavily influenced by both Ricardo and Nassau Senior, as well as Menger, while formulating his own theory of money.

324 See also Rothbard (1962: 192) for the actual physical medium of exchange being a historical realm, and outside the realm of pure economic theory as such. We can compare Menger's spontaneous emergence view of money in distinction to his contemporary Walras where he stated, "[l]a monnaie est une affaire d'Etat" (1898, 169). For a closer look at the differences between Menger and Walras' conceptions of money and monetary evolution see Arena and Gloria-Palermo (2008).

325 Interestingly, Menger (1892a; 1909) also viewed the state as being able to "perfect" money. Once money had previously emerged as a market institution, the state, primarily through legal tender laws, could improve upon its acceptance and increase its demand as a medium of exchange. This, however, relied on a pre-existing, market established money. In contrast with some of Menger's followers (and late contemporaries), he held a sympathy towards some state interventions, despite his preoccupation with "cold logic" in his economic analysis (see Hayek 1934: 417). It is indeed unfortunate that a lapse in this logic would occur on this crucial topic.

adoption of money requires two specific individual components: use-value and exchange-value. For general commodities, it makes no difference if they contain exchange-value. However, as Mises (*ibid*: 97) pointed out, for money to have a use-value, it must have exchange-value. To develop this concept one step further, there is no subjective use-value in money, unless it contains an objective exchange-value.

Mises' regression theorem explains why some monies that seem to have no use-value have come into existence. Their demand today is sourced from their demand the prior day, and onward, in a continual regression to the past. At some definite point in the past this regress must end, and at this point, money's exchange-value would be derived from the use-value available at that given moment. The result is that a given money might have no direct use-value in the present, however, for it to have demand in the present, it must have derived its exchange-value from some concrete use-value in the past. The direct implication of this is that no money can originate *ex nihilo*, or more appropriately, no money can emerge from something which has no use-value. The consequences for an thinker seeing the origin of money as state-given are evident.[326]

Later, Mises (1949: 405) clarified his position on exchange-value and money:

> The purchasing power which we explain by referring to the extent of specific demand is not the same purchasing power the height of which determines this specific demand. The problem is to conceive the determination of the purchasing power of the immediate future, of the impending moment. For the solution of this problem we refer to the purchasing power of the immediate past, of the moment just passed. *There are two distinct magnitudes* [emphasis added].

The relationship between demand and purchasing power is thus a complex one. The purchasing power of the immediate past that Mises writes about is conditioned by two things. First is the relationship that exists between goods available at that point in time for purchase and the amount of money outstanding

[326] Hayek (1976) would later advocate a system of competing currencies operating alongside the current government fiat money. In his view, the competing currencies could also be fiat in nature. For a critique of this view point, see Herbener (2002: 6), who argues that all a state can achieve in this is to ratify an existing medium of exchange. This would refute earlier criticisms of the Misesian theory of money by Gilbert (1953: 149) and Patinkin (1956: 71) who argued that fiat money introduced after a monetary collapse could not be explained by Mises' regression theorem. However, we can see that fiat money can never be introduced *ex nihilio*, but must always be offered for exchange with an existing currency (Rothbard 1976).See Tullock (1957) for a discussion of the historical failure of a spontaneously introduced paper money in ancient Persian civilization. He concludes the failure was due to the fact that the Persians had no prior use-value in fiat money, hence, they held no reason to believe it would have an exchange-value.

with which to purchase them. Second is where his regression theorem plays a pivotal role. The demand for money at that particular moment existed due to the finite regression of demand the money had previously experienced. Hence, we can say that demand for a given money was determined in the past by its previous demand and purchasing power thereof.

However, the demand and purchasing power of the "impending moment" is of a slightly different nature. Demand for the future is conditioned by the demand that exists due to the past purchasing power of the money. The purchasing power of the future will be tempered somewhat by the expected purchasing power it will contain at the moment it is expected to be used. Mises' distinction between the two "magnitudes" takes on great significance when viewed in light of this.

Time Preference and the Adoption of Currency

The primary focus that is operative in the adoption of a currency is hence finding a suitable medium of exchange to minimize the problem of the double coincidence of wants. Menger (1871: 208) however, noted that money is given two ancillary roles: (1) measure of value, and (2) store of value.[327] A measure of value (or unit of account) is incidental to money. Indeed, we could not think of the concept of money without also thinking of its inherent measure of value. Menger would downplay the significance of the "store of value" component, stating its existence is of "merely accidental nature" (*ibid*: 208; also 1909: 7). This viewpoint had somewhat detrimental consequences for later economists, as the focus was shifted solely to the static exchange-value money contains.[328]

A store of value, however, is another form of exchange, although with an inter-temporal twist. When an individual requires a store of value, they require a medium of exchange to be used in the future. Indeed, actors must always choose between exchange in the impending moment, and at some future point (Böhm-Bawerk 1889: 260; Rothbard 1962: 767). This finite amount of time will therefore

[327] Menger (1892b) would allude to the temporal element discussed in this section, although fail to expand upon it accordingly. The recognition that, "[t]he interval of time, moreover, within which the disposal of a commodity at the economic price may be reckoned on, is of great significance in an inquiry into its degree of saleableness" points to the fact Menger realized that media of exchange are not only used in the present, but also held for expected future exchanges. Elsewhere Menger (1871) would downplay the significance of this temporal factor, stating that *historically*, commodities have been used for exchange value, and not as stores of value. This historical tendency does not negate the fact that humans must consider this temporal element when choosing their exchange medium. See also Menger (1909) where he views the store of value function as being derived from money's medium of exchange role.

[328] These other specific uses take on increased importance when the value of money is under assault. Röpke (1937: 103-104) notes that in the Weimar hyperinflation, money lost its roles in the following order: first it ceased to be a store of value, second it failed to be a unit of account, and lastly it ceased to be a medium of exchange. However, in the adoption of a currency, we can see different circumstances will dictate the relative importance of each metric. My gratitude to Prof. Philipp Bagus for pointing out this passage.

entail the money having an exchange-value at some future point. Noting that a precondition for money's emergence was a steady (not necessarily stable) value, Wicksell (1935: 23) summed money's role:

> The money he acquires then remains in his hands both as ready money for anticipated future purchases or payments, and as a reserve for unforeseen liabilities. His money thus becomes his means of storing value (though usually only for a shorter period), his potential purchasing power, or future medium of exchange. In other words, it becomes a pledge or guarantee – *de facto* not *de jure* – for the future performance of counter-services to which he is economically entitled by virtue of the services he has performed.

In this way, we can see that an individual's choice of a medium of exchange is tempered by the exchange value of the present, and the expected exchange value of the future. Although it is foreseeable that one commodity may satisfy both wants simultaneously, it is impossible to state this *a priori*. In fact, the future suitability as a medium of exchange will be assessed individually through expectations of the conditions surrounding a given commodity.

An illustration may be helpful. We will consider two commodities. Commodity A exhibits great variability in its exchange-value in the short-term, having wild swings in purchasing power. However, in the long-term the historical precedent, and the future expectation, is that the purchasing power will remain relatively constant. Contrast this with commodity B, whose short-term valuation remains relatively constant. However, the historical precedent and future expectation will be that it will continuously experience a decline in purchasing power, making its future exchange-value greatly reduced compared to its present exchange-value. These two options can be compared in Figure 1:

Comparison of expected purchasing power

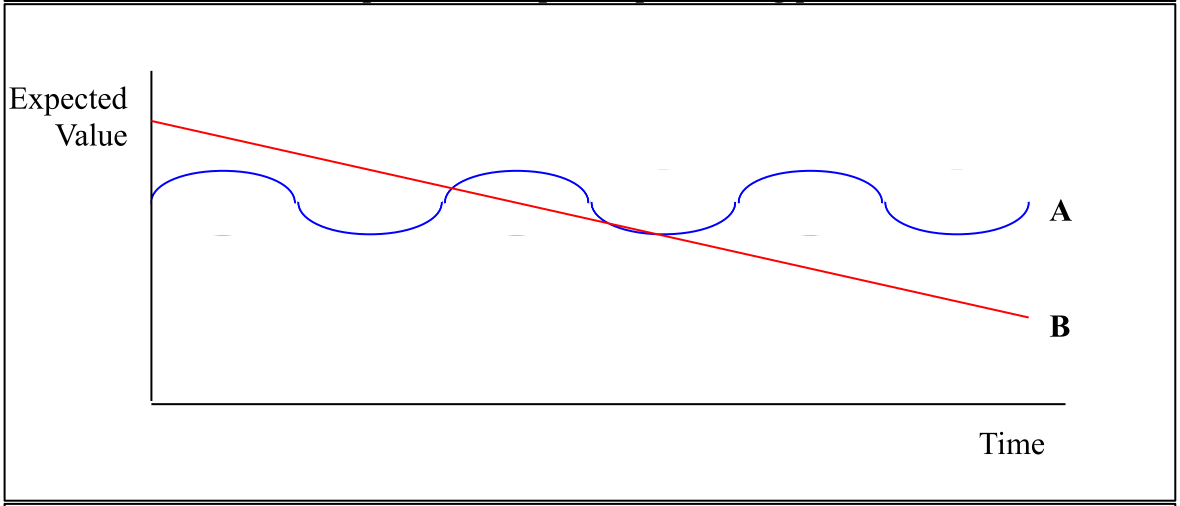

Figure 1

Two suppositions can be drawn at this point. The first is that, for purchases to be made in the near future, the superior medium of exchange would be commodity B. It experiences relatively minor short-term volatility, aiding the exchange process by adding an element of predictability to the future purchasing power inherent in the money. In terms of lowering the bid-ask spread of prices, we can say that a commodity that is relatively non-variable in purchasing power will reduce the short-term risk that changes in purchasing power can add to exchange. In contradistinction, commodity A has experienced, and more importantly, is expected to experience, a relatively higher level of short term volatility. This will not aid in reducing the bid-ask spread in the short-term that is a requisite condition for the choice of a medium of exchange. Given the binary choice between the two, we can state that the commodity exhibiting lower short-term purchasing power volatility (B) will be the preferred medium of exchange for transactions temporally closer to the actors.

Let us now look at a transaction expected to take place at a defined point in the future. This time, we can see that commodity A has historically set an example as having long-term purchasing power stability, despite its short-term fluctuations. Commodity B on the other hand, experiences a continual erosion of its purchasing power. Over any small period of time this may not seem significant, but given a longer time horizon the losses prove quite significant. Hence, for expected transactions of a distant future, commodity A, despite its short-term volatility, will act as a relatively good store of value, thus making its purchasing power stable for the long-term. The result on the bid-ask spread is that it can be expected to be reduced in the future, as contrasted with commodity B's greatly reduced future purchasing power, and hence, larger inter-temporal bid-ask spread.

What commodity is preferred to use as a medium of exchange will depend on the individual actors' time preference scales. Hayek (1980: 33) demonstrated that people wish to reduce the uncertainty of future prices through the use of a stable medium of exchange. The inevitability of future price changes means that although this uncertainty can never be eliminated, it can be mitigated through differing time preferences dictating how far into the uncertain future the money will be used. Two individuals will never share an identical time preference horizon. Some will prefer to exercise their ability to save for the future, and others will prefer to be net-consumers in the present. The result is that some people will prefer having a medium of exchange that exhibits short-term price stability to aid in their present-oriented spending habits. Others may prefer a commodity that acts as a superior store of long-term value, despite the short-term volatility it may experience. In fact, Menger would also identify the trade-off that exists between these two functions, adding that they may exist in a mutually exclusive form concerning their respective attainabilities (Menger 1909: 6). Heterogeneous expectations (stemming from different capital endowments, knowledge, etc) increase incentives for product differentiation in the market for exchange media (Cowen and Kroszner 1994: 35).

To a relatively more present oriented actor, the long-term store of value is a non-sequitur in the choice of money adoption. They may never think, nor act, in the long-term, thus this general criteria may prove to not be a criteria for their individual needs.[329] Likewise the future oriented individual, as they plan on using their medium of exchange in the future, they may not be availed by the relatively short-term price stability afforded by commodity B. The difference however lies in the fact that exchange is a necessary precondition of developed society – it would be difficult to imagine an actor in today's world not partaking in some sort of exchange. Hence it becomes apparent that as much as an individual may be future oriented, a certain degree of exchange must take place in the present.

Mises (1949: 417) would also view money as only existing due to uncertainty as to the future. If there was no inherent uncertainty, there would be no need to hold money. Money is necessarily a creation to deal with the element of change, an element removed in the certain world of his evenly rotating economy. Goods could be directly exchanged against one another as their full use, and hence value, would be known with certainty. Samuelson (1947: 123) would take a similar approach, noting

[329] Keynes may have forwarded that we are all dead in the long run. But we may require a store of wealth until that point, and possibly even longer. Rothbard (1976: 168) viewed Mises' regression theorem as saying that money's marginal utility yesterday is tempered by the expected marginal utility it will have *today*. However, it is not only today that is the subject of inquiry, but the marginal utility at any future date when the money is expected to be used.

that "in a world involving no transaction friction and no uncertainty, there would be no reason for a spread between the yield on any two assets ... in such a world securities themselves would circulate as money and be acceptable in transactions." Samuelson, however, was incorrect in assuming that in a fully certain world, anything would be money. In a fully certain world, nothing would be money – it would serve no role.[330]

Furthermore, differing degrees of price stickiness apparent in different exchange media will give rise to different demands placed on them. Individuals will have different inter-temporal commitments regarding their risks and returns. This heterogeneity will give rise to different actors demanding different exchange media depending on the expected time horizon, or duration, of their liabilities or receipts. As expectations and personal knowledge will serve a large role in determining these demands, we can see a plethora of separate media arising to satisfy these exchange demands.

Time preference will not only dictate *when* people will use their money, *but the actual type of money that will emerge*. As inter-temporal trade-offs dictate when individuals believe they will partake in exchange, the type of commodity held to be used as exchange will differ. Variations in the expected exchange-value of money may cause money to not be adopted, or used further, if they are too severe (Mises 1949: 422). This will also be tempered by the time preferences of the individuals they will be exchanging with. A medium of exchange is only valuable if it can be exchanged to a willing partner. This implies that the two parties to an exchange are equally important in determining what will be exchanged.

A Brief Note on the Separation of Medium of Exchange and Unit of Account Functions

Jevons (1876: 16) first discussed separating two functions of money: the medium of exchange role could be uniquely performed by a money different than the unit of account. Hence, a unit of account could be used for longer-term (i.e., greater than three month) contracts which is of a more steady value. The advantage of this separation of duties would be a stable store of value for exchanges in the future and savings, with a readily accepted exchange medium for present oriented exchanges.

This viewpoint enjoyed a resurgence with the new monetary economists who viewed the central defect of the current monetary system as the instability of the unit of account (see Yeager 1983). Black (1970), Fama (1983), and Hall (1983) have all proposed systems aiming to rectify this issue and lead to

[330] Leontief (1947: 238) would criticize this viewpoint originally by pointing out that demand for money must necessarily be zero in equilibrium, the same formal conclusion Mises had earlier drawn.

greater price stability. In essence, money could, under these forwarded systems, be held in two forms: deposits in banking institutions, or currency. The former would represent a unit of account, and hence, store of value function. This role would be fulfilled by a separate medium than the exchange media used for the second role of currency.[331]

Greenfield and Yeager make a compelling argument that, on the free market, media of account and exchange would be separate roles, performed by unique monies. They (1983: 309) forward that:

> Market clearing forces do not work very well to maintain or restore equilibrium between money's supply and demand because money does not have a single price of its own that can adjust on a market of its own. Instead, the medium of exchange has a fixed price in the unit of account (each dollar on the money market has a price of exactly $1). With no specific price and market to impinge upon, imbalance between money's supply and demand must operate on the dollar's purchasing power, that is, on the whole general price level.

However, as Dowd (1988) has demonstrated, this argument does not apply to money in general, only to non-interest bearing money. On the market, the rate of interest acts as an equilibrating factor on this force.

Doubts have been raised as to if such a separation of monetary roles is possible in practice. Rothbard argues that money cannot be solely viewed as an abstract unit of account, “divorceable from a concrete good” (1981: 4). It cannot represent a *mere* claim, nor a guaranteed fixed price level. Instead, in Rothbard's eyes, money must remain a commodity, and as such, a involve itself as a medium of exchange to have value. Following the original Mengerian line of thought, Shostak (2000: 71) argues that while the roles of unit of account and store of value are important, they are not fundamental to money. Instead, serving as a medium of exchange is what gives rise for these ancillary roles of money to emerge.[332] In fact, as money is saved in order to be used at a future date, separating the two roles only pushes the problem back one step. A unit of account distinct from the exchange medium will be forced

[331] Cowen and Kroszner (1987) and Summer (1990) provide reviews of the history of new monetary economics, and the separation of money's roles.

[332] Niehans (1978: 118) points out that, in any case, money should not be referred to as a *unit of account*, as money is not a unit. Instead, if one were to take this line of reasoning, money should be thought of as a *medium of account*. A unit of account is a specific quantity of a good representing a medium of account. For instance, silver may be considered a medium of account, but a 'ducat', defined as a weight of silver, represents the units of account. We see that a unit of account can never be removed from the medium of account, which in turn evolves from the medium of exchange. White (1986) also provides a view of money with the unit of account being a necessarily ancillary role to the medium of exchange function.

to be converted to a medium of exchange in order to be used at a future date.

We find that it is impossible to separate the roles of money, their values all regress back to money's prime role: that of an exchange medium. We find, however, that these ancillary roles serve to temper the value money has in exchange. Its emergence is not brought about in absence of these considerations, instead it is due to these ancillary roles that a monetary system develops to fill these needs the best.

Money and Value

Mises (1912) stressed that exchange and use-value are identical for money; there could be no other possibility. However, the daunting question that must be answered is where the true value of money derives from. It is easy to see that it *derives* value from the commodity it is composed of, but this is only half the story.

The distinction between commodity money and the commodity contained in money should be made. Although they may be chemically identical, they are different entities to the valuing mind. A commodity proper derives its value from its use-value. It can have varied uses, each with differing ultimate values. It is the economizing actor that values a given commodity among the multitude of choices that may be available for it to be used.

Commodity money, in distinction, will also be valued for its industrial use-value. However, this factor is only one part contributing to its total value. Money's value is always conditioned by expected exchange-value at the time it is to be used (Hazlitt 1978: 75). Exchange-value is determined by many factors, among them the industrial demand that led to the original rise of a given commodity as money.[333] However, we can also see that the degree of acceptability will influence the value of money. Money will be offered through an intermediary, as was previously established. Dowd (1988: 645) points out that private suppliers of money will derive their profits through three main avenues: (1) competitive minting fees, (2) innovation, and (3) *a reputation for probity* to insure its customers that they are not cheated. This profit will depend on the value customers place on their issued money, and hence, the degree to which they are willing to pay to use it.

Commodity money will enjoy a value, or seigniorage, in excess of its commodity counterpart.

[333] Thus exchange-value is strongly influenced by the next most valuable use-value the market may have for the commodity.

This is due to the added use-value it will contain.[334] As providers of money modify a commodity into an exchangeable medium, they add value to it. As more people accept a given money in exchange, its use-value (through exchange-value) grows. As Dowd points out, private producers will work toward increasing innovation in the monetary realm. This could have the effect of increasing consumer confidence in a given commodity money, and hence, increase its value.[335] The conclusion is that two commodities, both made of the same commodity and having the same content, may have different values on the market. Money production on the free-market would operate under the same constraint as any other good – the profit and loss motive (Herbener 2002: 7; Hülsmann 2003: 39).

An example will clarify the matter. Assume Company A and Company B both produce a gold coin of one ounce nominal weight. In both circumstances, the value placed on their gold coins will be more than that of raw gold in a natural state. This is immediately due to the added exchange value that monetary gold would have. A store, for example, would be more apt to accept payment in a readily identifiable gold coin, or known weight, than a random amount of gold ore. This would stem from the ease of identification, as gold ore would necessarily need to be checked for weight and purity before being accepted. Additionally, once it was accepted by the store, the problem would compound in the future of who would accept it from them in exchange. In fact, as McCabe (1989) shows, money reduces the subjective risk each exchange partner has, as each no longer must worry about the future value or acceptability of their received good (i.e., money).

But what of the exchange value between our two companies' coins; what will be the relation there? We can see that it would be a rare occurrence that both companies would issue monetary gold in an identically identifiable manner. For instance, one company may have a reputation for purity, ease of delivery, or any number of ancillary features that would contribute to the monetary gold's exchange value.[336] Further, as competing issuers would derive profit from the value of the money they provide, we can see the emphasis they would place on providing additional value-adding features to their money. Commodity money will always be a subset of a particular commodity (Rothbard 1962: 169). As

[334] Garrison (1992: 62) comes very close to this insight stating, "there is an Austrian preference for gold as money, *not due to its specific qualities*, but also to the historic precedence that has been set over time of people preferring this as their standard" (emphasis added). What Garrison fails to explicitly mention is that some of the specific qualities in gold (or commodity) money will be human added. Menger (1892a) divided money's value into two sources: inner and outer values. In much the same dichotomy we are looking at here, inner money represented that pertaining to other goods, while outer value was relevant to the actual money itself.

[335] Increasing consumer confidence in the purity of money may be the largest source of seigniorage, and has historically been a prominent concern of money issuing groups. See Cowen and Kroszner (1994: 50), Ritter (1995: 134), Garrison (1992: 62) or Howden (2008) for more on the importance of seigniorage creating attributes to money's value.

[336] Economists must realize their limitation in assigning reasons for seigniorage. Market providers will create this value in ways that the economist could never predict; it is a market result that requires a market process for discovery.

such, although it is possible that two competing commodity monies may have identical value, it is more likely the value spread will be time-variant.[337]

Some may argue that Gresham's law would take effect at this point with competing currencies with identical bases. We can see this is decidedly not an issue with our gold issuers. If a gold currency existed of greater value than a competing gold currency, their exchange rate would fluctuate, but this would be separate of the fluctuation of the exchange rate they would share with gold ore. Also, the higher valued gold currency company would never purchase the lower valued currency and homogenize it with their own. The reason is that the cost of doing so would be higher than the cost of buying gold as ore on the market directly. Also, as different individuals will have differing needs for media of exchange, we can see the possibility of competing suppliers operating concurrently and independently on the market. Indeed, in the existence of fluid consumer preferences, more than one type of exchange medium may be the norm (Banerjee and Maskin 1996: 989).

Grubel (1969: 270) argues that the amount of seigniorage that exists will be relative to the degree of competition available. In the existence of perfect competition, seigniorage would disappear completely. Dynamically however, we can see that perfect competition can never exist. Consumer preferences will constantly be shifting. As a result, a product, or currency, demanded will also be ever changing, with continual profit opportunities to be had through seigniorage. We can note however that the organizations enjoying a monopoly privilege through fiat will enjoy the largest seigniorages. It is only through competition of money production that excessive seigniorage can be mitigated (Hülsmann 2003).

Conclusion

The origin of money is primarily viewed as a two-fold issue. One the one hand, actors search for a medium of exchange, on the other, a store of value for the future. The Austrian view-point, following Menger (1892b) and Mises (1912), has focused primarily on the former at the neglect of the latter. Menger's theory on the adoption of money through the search to find a commonly accepted exchange medium to reduce the bid-ask spread of exchange was an instrumental turning-point for monetary

[337] If the reader would like a concrete example, they should ask why Scottish pound notes were exchanged at a discount in England for many years, despite representing the same purchasing power. Or why Scottish, Irish or Welsh pound notes regularly trade at a discount to English notes with street exchangers in many European countries. Additionally, we can see that not all debit card providers charge the same amount for services. Instead of letting this influence consumer prices, however, stores often absorb these costs as a business expense, thus mitigating a flexible rate between using company A's debit machine and that of company B, for the end consumer.

theory. Mises' regression theorem would later almost complete the puzzle. By demonstrating that money must always come into use due to some previous use-value, we can see that the case for the *ex nihilo* emergence of fiat currencies is non-existent.

However, the much neglected store of value component needs to be addressed. Although all exchange takes place in the fleeting temporal present, actors hold expectations of exchange in the future. To the degree that their time-preference dictates when this expectation will materialize, they will search for a store of value to sustain their purchasing-power until that point, as well as reduce the expected future bid-ask spread of exchange. Due to the subjective and personal nature of time-preference trade-offs, this expectation will vary among individuals.

However, many Austrians also assume that these companies will offer a homogeneous commodity money, of equal value. As was shown, due to the personal nature of the time-preference trade-offs inherent in the choice of exchange media, it is likely that no one commodity would be used as money, but several simultaneously. This would provide and enable individuals to choose monies coherent with their personal time-preference scale. It is foreseeable that the same individual may prefer holding multiple monies of differing time-preference satisfying features, thus increasing the need for competition among money producers.

Furthermore, as the necessary precondition of a money is the existence of an underlying medium that is wanted in exchange, we should take care to note this medium is likely not to be unique. In fact, as different regions prefer different commodities, regional tendencies will emerge for preferred media of exchange. Although it is possible that the whole world may function on a common commodity standard, the more likely reality is that several currency-blocks would form, each utilizing a specific commodity for its monetary base.

The true value of a money comes from its exchange value. This is distinct from the underlying base commodity utilized. Additional features such as the issuer's reputation, probity, availability, and ease of use all add to this value metric. Under a system of privately issued, competitively circulating currencies, the value inherent in each may not be identical, as determined by weight or purity. Instead, the value will be dependent on many other factors as well. The fact that these factors will be in a constant state of flux, as demand to hold these currencies, as well as demand to accept them in exchange, are continually revised. Only a system of flexible exchange rates among these competing currencies will ensure that Gresham's law does not take effect, and wreak havoc on the monetary system.

Money's value is not only a feature of concern for the fleeting present, but also must be accounted for in the expected future. People require a medium of exchange *and* a store of value – they intend to use their money now *and* at some future time. To the extent that time preference indicates how far in the future this future use of money will occur, the individual searches for a monetary unit that maximizes their expected purchasing power at that future date. Time preference is an influence not only on when we spend our money, but also what we select to use as our money as well.

VII. ENTREPRENEURSHIP AND DISCOUNTING PROFITS

The starting point for most equity pricing models is the dividend discount model. Hence, under this line of reasoning, a firm's intrinsic value can be estimated to be the present value of all its expected future dividends. This line of reasoning is fully consistent with much of what we have previously developed. Resource-supplying entrepreneurs contribute money to a firm and as a result receive a portion of the entrepreneurial profits. This portion of profit dedicated to the shareholders is what we see today as a dividend.

In fact, two sources of profit exist for these resource-supplying entrepreneurs from financial assets. The first is the dividend payment what will be received. This is primarily determined by the risk-mitigating entrepreneurs of the firm, and their ability to maximize present profitability and hence, earnings. Secondly, the future sale price of the asset is of significant value. This factor will primarily be determined by the uncertainty-bearing entrepreneurs and their ability to steer the firm in the directions that consumers value most highly.

Of great importance is also the expected length of time that an asset will be held for before it is sold. This is affected by two groups of factors. The first are personal to the purchaser, and dependent on factors such as their own life-cycle, age, or propensity to leave assets to individuals different than themselves (i.e., dependents, foundations, etc). Secondly, we will see that there are firm specific factors that affect this holding period. An individual will hold an asset to maximize return, and hence, will sell the asset at the point when this is expected to be achieved. Volatility of an asset will affect the length of time that it will be held for. Highly volatile assets will cause a great amount of uncertainty concerning their future selling price, and hence, cause investors to hold them for shorter periods of time.

By incorporating this temporal aspect of the previously examined concept of time preference, we see that the temporal dimension of saving takes of great importance. Additionally, by incorporating the concept of the entrepreneur, we may begin to identify particular sources of value within financial assets.

1. A Simple Dividend Discount Model

It may prove instructive to give a brief summary of how the common dividend discount model is obtained.

Assume an individual purchasing a stock and expecting to hold it for a one year period. The value they would place on that stock would be no more than the income they expect to receive at the end of that year. Hence, the dividend at the end of year-one, D1, and the future stock price expected to obtain at the end of year-one, P1, will have to be discounted by the appropriate interest rate to create a present value that the investor can compare to their other current alternatives. Hence:

(7.1) $$V_0 = [D_1 + P_1]/[1+k]$$

whereby:

V_0 = present value of stock
D_1 = expected dividend payment at end of year one
P_1 = expected price of equity at end of year one
k = expected originary rate of interest over a one year period

We may then note that if we knew the value of P_1 we would not have to undertake these calculations. The investor could compare this future value to the present equity price and determine if a profit opportunity exists or not. However, instead we see that the value of the equity's price at the end of the year must be estimated. Hence:

(7.2) $$V_1 = [D_2 + P_2]/[1+k]$$

whereby:

V_1 = the intrinsic value of the stock at the end of year one
D_2 = expected dividend payment at end of year two
P_1 = expected price of equity at end of year two

k = expected originary rate of interest over a one year period

If we assume that the stock will be selling at its intrinsic value at the end of year one – that $P_1 = V_1$ – then we can substitute equation (2) into (1) to obtain:

(7.3) $$V_0 = D_1/[1+k] + [D_2 + P_2]/[1+k]^2$$

This equation (3) assumes now that the holding period has changed to two years, after which the equity will be sold. Generalizing this equation for *t* years we obtain:

(7.4) $$V_0 = D_1/[1+k] + [D_2 + P_2]/[1+k]^2 + \ldots + [D_t + P_t]/[1+k]^t$$

Note that this equation does not necessarily exclude capital gains, and only focus on dividends when valuing a security. Each price is assumed to factor in the future dividends to be received, *plus* the capital gains that will be obtained from selling the stock at a future price. The reason then that dividends appear in the formula at the expense of capital gains is that the capital gains are *determined* by the dividends expected to accrue. Thus, as Bodie, Kane and Marcus (2001: 412) summarize the concept, the discounted dividend model "asserts that stock prices are determined ultimately by the cash flows accruing to stockholders, and those are dividends."

However, to make this equation more usable, we may assume that dividends are increasing at a constant rate, *g*. We then see that future dividends can be given by the formula:

(7.5) $$D_t = D_0(1 + g)^t$$

Substituting this equation (5) into (4) we may simplify to:

(7.6) $$V_0 = [D_0(1 + g)]/[k-g]$$

(7.6b) $$V_0 = D_1 / [k- g]$$

This equation (6b), originally developed in Gordon (1959), provides the basis and most fundamental starting point to understanding stock valuation. There are some issues which have not gone unnoticed concerning this simplistic model. First we may see that the growth in the dividend rate must be smaller than the discount rate used to bring future dividends to their present value. Also, the integral role that the dividend plays may be seen to fly in the face of periods of time where dividends are minuscule, or not issued at all.[338]

One significant misgiving comes directly from the assumption that stock prices are determined solely by the future dividend stream expected to be received. Although without an appropriate simplification in some cases, this assumption causes more to be lost in analysis than is gained. On the one hand, for stocks that are expected to continue paying a steady dividend stream (i.e., those established companies known as blue chips) there is more than an ounce of truth in this assumption. However, it overlooks the fact that under conditions when demand to hold financial assets paces beyond the new supply thereof, a decrease in the dividend payout ratio will cause valuations to reflect the expected future selling price, and not necessarily the interim dividends that can be gained. Instead, in these cases where the equilibrium supply-demand balance is upset, there will be much reason to forgo this method of pricing stocks, as we shall see.

Equilibrium and Dividend Discounting

In the previous chapter we had discussed how an equilibrium might obtain in the realm of financial asset pricing. In this case, the equilibrium condition was reliant on the, albeit, unrealistic assumption that the supply elasticity of new financial asset issuance was unity. This would entail that for every monetary unit of new supply demanding to invest in a financial asset, there was an offsetting increase in new supply thereof. This is problematic for several reasons.

First, as was seen, there is a *ceteris paribus* tendency for existing firms to prefer financing through debt owing to the assured limit to the claim of entrepreneurial profit that will result. Loans always entail a *maximum* duration which ensures that they are paid off within a finite amount of time, eliminating the transfer of payments over an indefinite period. In contrast, the issuance of financial assets almost always lacks this maximum time period. As a result, financial assets may be issued with no assurance to the issuer that there will be a cap on the entrepreneurial profits that must be paid to these recipients. As a result, the potential *loss* of profit is unlimited to the firm under this scenario. The

[338] Although, it has been forwarded that if the Modigliani-Miller (1958) hypothesis that dividend irrelevance holds true, dividends may be substituted with earnings. However, the idea that absent taxes, asymmetries and transaction costs (i.e., in a perfectly efficient market) a firm's capital structure has no effect on its valuation places unreasonable restrictions on a model of real stock prices. Gordon (1989) provides almost 50 articles and books that challenge Modigliani-Miller.

result is that, ceteris paribus, preference to given to issue contracted loans in opposition to equity.

Second, issues of equity are undertaken at discrete intervals. Firms as a general rule issue blocks of new securities, underwritten by financial intermediaries to ease the uncertainty that the expected asking price will be met. This provides the supply-side restraint on perfect new security issuance elasticity. On the demand-side, we see that individuals do not always wish to invest an amount of funds perfectly conducive to purchasing the amounts of securities offered. In this way, we see that the supply of fresh funds will not always be perfectly matched to the supply of fresh financial assets.

Last, it would require that the flow of savings *and* the demand to invest a portion of these savings remain constant. To continue our more conventional terminology from the previous chapter, both time-preference and liquidity-preference would have to remain constant. In this way, there would be a constant flow of funds to be invested in the fresh supply of financial assets. As neither of these preferences are seen to be constant, owing to a multitude of factors, we see the problems that develop in relying on an equilibrium situation to obtain reliant on their constancy.

However, if we assume for a moment that these three conditions are perfectly met, we may at that point see that the use of the previously developed dividend discount model would be fully consistent. The dividend payout ratio growth rate (g) would be zero in this case, as in this case the equilibrium condition would eliminate any net changes in this growth rate. As a result, the perpetuity condition would prevail as follows:

(7.7) $$V_0 = D_1 / k$$

As no future sales would be expected, the current price of a financial asset could be provided by the present value of its future stream of dividends. No future sales would occur as the supply of these financial assets would be perfectly elastic to adjust to new changes in demand. Hence, any demand increases would result in an offsetting increase in the available supply of financial assets.

The removal of any one of these three equilibrium conditions changes the necessary formula drastically. Fortunately, there remains some grain of truth in the above stated formulae despite their functionality being confined to an equilibrium setting. As we will see, with some minor changes, we may more accurately reflect the reality underlying the prices of these assets.

Disequilibrium and Financial Asset Pricing

In disequilibrium, three significant changes must be made to the above stated formulae. First is that the expected future selling price of the asset must be accounted for. As the dividend stream will not be able to account for all of the present value of an asset, we must allow for a disjointing from the dividend stream expected to prevail, and the expected selling price at the time when the asset will be discarded. Second, we see that the assumption that a dividend be offered at all must be dropped. Indeed, under extreme supply/demand imbalances concerning the quantity of financial assets in existence, the yield offered by a company in return for the use of financial resources may be so great that dividends are pushed almost to the point of non-existence. By the same reasoning, the future selling price of a security (P_t) must only abide by the condition of $P_t \geq 0$. In this way, we see that a financial asset may be purchased solely for its expected dividend stream, with no hope of being sold in the future for any non-negative price. Alternatively, a financial asset may be purchased with the hope of selling at a higher future price that is not reliant on any dividend income. Lastly, we see that the discount rate will not be a constant, as is commonly assumed in most pricing models. Instead, there will exist a structure of yield curves at t=0 that will be used by the individual to discount the expected future cash flows back in time to the present.

Provided a dividend stream is constant *and* we assume that all earnings will be distributed via dividends over the life of a firm, the more common dividend discounting model may be used. However, those cases where these two assumptions differ call for a decidedly different formula. One alternative is that a company may never pay-out a dividend, instead reinvesting all new profits to be used to expand their business activities. In this case, the pricing model would have to take into account this lack of steady income, and instead make an estimate of the selling price at a future point in time. Alternatively, we may see that an individual may forecast a firm to supply a valuable good or service, but that due to some future eventuality this value will decline drastically. In this case, we may see an individual holding the financial asset for its interim dividend, and then hopefully a sale of the asset prior to the decline in value that is expected to occur.

The general model that we can apply is a combination of both these cases. On the one hand, we see that future dividend streams must be accounted for. On the other hand, we see that they cannot be the only income accounted for, and that the future sales price must have a more direct estimation, not indirect as through the reinvested dividend stream.

Hence, we may see that the value of a stock today, V_0, will be equivalent to its future stream of dividends coupled with its expected future selling price. Therefore, the value of holding a financial asset for *N* periods is:

(7.8) $$V_0 = D_1/[1+ k] + D_2/[1+k]^2 + \ldots + D_N/[1+k]^N + V_N/[1+k]^N$$

whereby:

V_0 = value of the financial asset today
D_N = expected dividend at time N
k = originary rate of interest
V_N = expected value of financial asset at time N

For simplification, we may take the same assumption that the dividend payout will remain constant (i.e., $g = 0$), we see that:

(7.9) $$D_1 = D_2 = D_N$$

We may summarize equation (7.8) as:

(7.10) $$V_0 = \Sigma\, D_t/[1+k]^t + V_N/[1+k]^N$$

whereby:

$\Sigma\, D_t/[1+k]^t$ = the sum of all expected dividend payments discounted to t=0
$V_N/[1+k]^N$ = the present value of the expected future selling price of the asset at $t=N$

This model explicitly recognizes the future selling price is independent on the expected stream of dividends to be received by ownership of the financial asset. While we have also assumed that there is no growth in the dividend payment, we are willing to accept this point for greater generality. Indeed, for short holding periods (small *N* values) the growth in dividend payments will likely not be large enough to meaningfully affect the results. Additionally, their diminished present value due to discounting will also reduce the significance of any dividend growth. We *have* however made an additional assumption which has gone un-commented, and which will alter the results of this model significantly.

The assumption that the originary interest rate k is a constant implies that a flat yield curve is operative, and expected to remain so for the holding period N. In reality, these values of k will not be constant. Two options exist for an investor with an originary yield curve that is non-flat (i.e., either positively or negatively sloping).

The first is that they may invest their funds for the originary interest rate at that which occurs for the given durations on the present yield curve. That is to say, if a financial asset is bought to be held for two years, the first dividend would be discounted at the one year rate, with the second dividend and future selling price discounted at the two year originary interest rate. However, there is a second option that the investor may chose, depending on their expectations concerning the structure of interest rates.

If the present yield curve is a faithful representation of the future interest rates, there will be no difference between an individual's expectations concerning future rates and the present curve. If the yield curve is significantly steep, that is to say, there is enough positive spread between short-term and long-term rates, the assumption of using the present yield curve to discount future cash flows is warranted. However, if a flat yield curve exists, *and* an individual has the expectation that in the future a significant steepening will occur, than the use of the present yield curve will not be enough to reward the investor.

For example, assume that the one year originary interest yield is 2% and the two year yield is 3%. As long as the individual's expectation is that the one year yield in one year's time will be less than 4%, they will lock-in their investment at the current 2 year rate. Under normal market conditions this type of movement in short-term yields may be unlikely (effectively, 100% in one year). However, if we start from a flat yield curve, the decision is not so easy. For instance, if both one and two year rates are 2%, any expectation that the future one year rate will be higher than 2% would entice the investor to invest their money for one year, and then renew it at the end of the year for another year, rather than invest for 2 years. With a negative yield curve the difference becomes even more extreme, as investors are enticed to invest their funds for short-terms at higher interest rates rather than lock in at shorter rates for longer time periods.

Therefore, based on expectations of future rates, individual investors will base their decision as to what rate is appropriate to discount at upon the more profitable of the two options: 1) locking in a rate today for a longer time period, 2) using a series of short-term rates to maximize returns over longer periods by continually rolling over the investments.

The assumption of a single k to discount all future cash flows implies that the originary interest rate will remain constant indefinitely into the future. However, the rate used to discount the expected cash flows should be the one that reflects the actual time period under consideration. It remains true that at the moment an individual purchases a financial asset, they have the option to sell it at any time they see fit. Using this line of reasoning, it makes sense to discount all cash flows back using the shortest originary interest rate possible. In most literature, the originary interest rate is substituted for a risk-free rate – 30 day T-bills being the most commonly cited proxy for the risk-free interest rate. However, while the *option* to sell these assets at any time exists, and individual will generally have a vague idea of how long they wish to hold on to the asset for at the time when they purchase it. From this point of view, we see that the opportunity cost of the foregone use of their money will have to be compensated by an interest rate commensurate with the length of time their money is expected to be invested.

Therefore, not just one singular k value may be used, but several for each duration corresponding to the time at which the expected dividends and future asset sale will occur. Now we see that collapsing the formula, as was done in equation 7.10 is not a realistic assumption, and that more is lost than is gained through the generalization. As a result, equation 7.8 remains the standard that should be used, with each k dependent on the greater of the two options that an investor is faced with: 1) investing their money for a short-term and reinvesting it (rolling it over) at higher rates in the future, or 2) investing it at the originary interest rate of duration identical to that when the income from the asset sale or dividend will be realized.

2. Risk-Mitigating, Uncertainty-Bearing and Financial Asset Pricing

The entrepreneur plays an integral role in the pricing of financial assets when viewed in light of the equation (7.8) developed previously. We have seen that financial assets are claims to real assets, which are normally owned within the structure of a firm. Firms are attempts to create a pure entrepreneur by combining different types of entrepreneurs. It logically follows that any theory of pricing these firms, and hence, the financial assets claiming them, should scrutinize the actual entrepreneurs that are at the core of the asset to be priced.

The first class of entrepreneurs which we wish to examine are those that we typically refer to as “employees”, but which are really serving the integral role of risk-mitigating within the firm's boundaries. Hence, they work within a closed means-ends framework and serve to increase efficiency to maximize *actual* profits from these available resources. *Profitability* becomes the clear feature of these entrepreneurs. Secondly, we may look at that class of entrepreneurs – uncertainty-bearers – which we normally refer to as upper level managers or of the board of directors. Both of these groups affect profitability in a myriad of ways.

This profitability is reflected in our equation 7.8 in three ways. The first is through the present dividend that is paid out by the firm. Dividends are only able to be paid to the maximum of net profit that a company earns. This profitability, we have seen, is determined by how much risk this class of entrepreneur can mitigate in the present. The second is through the growth in dividends expected to occur over time. As entrepreneurs also undergo a learning process, we see that their relative competitiveness compared to other firms will be reflected through their ability to continue providing a stream of dividends into the future. Last, we see that the future selling price of an asset is dependent on these risk-mitigating entrepreneurs that expect to be employed by the company at the time when it is finally sold. Additionally, this future selling price will reflect the cumulative learning process which has been internalized within the firm and has given it an advantage relative to its competitors.

These three aspects all put a heavy weight on the role that the risk-mitigating entrepreneur serves. However, we see that this weight is placed most heavily on the emphasis given to the present dividend, with declining importance to future dividends, and lastly, the effect they have on the future selling price of the financial asset. This stems from the fact that the *actual* profitability that the entrepreneur creates becomes *less* important than the *potential* profit that the uncertainty-bearing class of entrepreneur makes possible. The uncertain nature of the future also underweights these future profitable occurrences. As they are discounted back to the present, they become gradually less and less important the further temporally they occur (in fact, due to compounding they loss importance at a rate proportional to the mathematical power of their duration).

To begin pricing a financial asset the first place to start is by focusing on the individual entrepreneurial components that comprise the asset that the financial asset claims. By focusing on the specific role of the risk-mitigating entrepreneur we see that a significant portion of financial asset valuation stems from their ability to work efficiently, and hence, maximize profits, in the present.

Current Profit and Dividend

Current profits are almost solely determined by the efficiency at which risk-mitigating entrepreneurs within a company can operate. By working within a fixed means-ends framework, the degree to which this class of entrepreneurs can eliminate inefficiencies (through mitigating all risk) will allow profits to be maximized. Dividends have a maximum value set at the net profit which a company earns in a given year (ignoring retained earnings). As profits are maximized, dividends too may be maximized.

Some may counter that profitability also concerns the demand that customers have for a given product or service. However, this is of no concern to this class of entrepreneurs who are not necessarily able to influence consumer demand for their product except in one way. By allowing for a larger profit margin through increased efficiencies, prices may be reduced without sacrificing profit. This has the effect of stimulating consumer demand. However, by and large, this class of entrepreneurs is only concerned with a *given* demand coupled with the fixed means-ends framework that is available to them.

Learning and Future Dividend Growth

That class of entrepreneurs we have referred to as uncertainty-bearing are the main determinants of the future growth rate of the dividend stream. They are the individuals who create the means-ends framework that the risk-mitigators will have to maximize within. Two aspects of this role are apparent.

The first is that through the decisions made in the past the maximum allowable profit potential is set by the uncertainty-bearers. Past decisions concerning the strategic decisions of the firm will shape the future potentials that can be realized. To the degree that these past decisions have been conducive with the desires of consumers now in effect, profits can be higher than would otherwise be the case.

The second case arises as this process is shifted forward. The decisions made *now* become the basis of future potential earnings. This translates to the growth potential of earnings, and eventually, of dividends. This entrepreneurial class directly determines these future potential growth rates. However, it remains true that these potentials must be *realized* in the future, not only remain possibilities. For this reason it becomes clear that risk-mitigating entrepreneurs must be secured to work for the firm in the future, in order to realize these potentialities.

Risk and Future Asset Valuation

Lastly we must look at the determinant of the future price of the financial asset, at that point in time when we expect to sell it. Its future value will be largely dependent on the resources that it has at its disposal at the given future point in time. Both classes of entrepreneurs will have to be expected to be employed by the firm, with a capital structure that is conducive to the needs of consumers. Much uncertainty surrounds this decision as it will normally be future oriented. However, we may see that problems arise concerning succession of key management positions.

Take an example of a firm with a single very important CEO who is highly valued for their ability to steer the company into profitable directions in the future. If this CEO is expected to step down and retire shortly, there will be significant uncertainty concerning the replacement. Specifically, the replacement's ability to bear and shoulder future uncertainty may not be as sufficient at the existing CEO's. In this case, we can see that the future selling price of the stock would be reduced accordingly as the future ability of the company to move in new valuable directions will be compromised.

Likewise, the quality of risk-mitigating entrepreneurs at the time that the financial asset will be sold will have a significant bearing on the value thereof. For imagine for a moment that at the time when an individual tries to sell their stock, there is a general strike which effectively means that there will be none of this class of entrepreneur working at the future date, and it opens the possibility that they might not be working at a more distant future either. In this case, we see that the negative impact on earnings will significantly effect dividend payments.

Conclusion

Hence, we see that the mix of employees is the core of value determinants that a financial asset has. Of course, capital considerations also factor in, but these are also the result of entrepreneurial foresight. Our two entrepreneurial classes who directly work in a company effect our asset pricing formula in two different ways.

First, the uncertainty-bearing entrepreneurs provide the potential limits for future growth and profitability by steering the firm in the correct direction. By keeping the firm's productive capabilities in-line with that of consumer demands – both in the present and the unknown future – this class of entrepreneurs creates growth potentials that are exploited by the risk-mitigating entrepreneurs. This class then determines the actual profitability that occurs in the present. As dividends rely on earnings, we see that efficient employees are the ultimate source of all dividends within the company.

By viewing financial asset pricing, especially as defined in equation (7.8) we see the integral role that these two classes of entrepreneurs provide.

3. Expected Length of Asset Holding and Pricing

One instrumental aspect of financial asset pricing that has gone un-commented on thus far is the expected time that an individual will hold an asset for. In equation (7.8) we see that there is a crucial distinction between the dividend stream that will be discounted back to a present value, and the future selling price of the asset. Thus, the amount of time that the asset is expected to be held for will considerably affect the value thereof.

Previously we looked at time preference as the trade-off that exists between consuming in the present versus the future. The liquidity preference aspect of the holdings that comprise a portfolio of financial assets has likewise been assessed, but found to be inconsequential concerning the *ex ante* formation of a portfolio. The trade-off that exists when choosing the amount of time to hold an asset for is quite similar to both these concepts, but in many ways quite different. Garrison's (2001) "saving up for something" hints at the concept we wish to use, which is the temporal dimension of saving.

Assets are not bought to be held forever. Instead, they are purchased with a future date in time to be sold back to cash to further consumption. As Böhm-Bawerk (1901) makes clear:

> When Crusoe on his island saves up a store of provisions in order to gain time for the fashioning of better weapons, with which he hopes later to secure a much larger quantity of provisions, these relations are all clearly discernible. It is obvious that Crusoe's saving is no renunciation, but simply a waiting, not a decision not to consume at all, but simply a decision not to consume yet; that furthermore there is no lack of stimulus to the production of capital goods nor of demand for the consumption goods subsequently to be produced by their aid.

Hence, saving is not to be thought of as renunciation (although it is for a given time period). Instead, it is more generally an example of waiting. Resources are put aside to be consumed at a later date – renunciation may happen at the moment, but the consumption desires will be fulfilled at some future date which must be *waited* for.

There exists a temporal dimension to saving which is not to be found in most analysis. We have already demonstrated that general equilibrium models by definition exclude the possibility of a temporal dimension to saving. Likewise, the basic loanable funds model takes a static position of savings and assumes it will continue in this manner until it changes. The duration of savings is a crucial aspect that deserves attention.

Three aspects of the temporal saving aspect become important. The first concerns and individual's life-cycle. Savings are undertaken to consume in the future, which means that at some future point in time *dis*saving will occur. As an individual nears the end of their life-cycle, we see a tendency for the temporal aspect of saving to shorten accordingly. However, there is the case where an individual plans not only according to their *physical* life, but their *economic* life as well. Hence, an individual may have the foresight to save for their children, business partners, dependents, spouses, or others who may benefit from this act. To the degree that these entities extend the *economic* life of the individual, the temporal dimension of saving may be increased.

The next aspect is that which depends on specific conditions concerning the asset to be purchased. Hence, an individual purchases an asset under the plan to maximize their monetary return. To the degree that supply/demand conditions change, general market conditions improve or deteriorate, or other events obtain which will effect the return on an asset accordingly, an individual will alter the length of time this holding period will extend for. This will affect the expected return on an asset.

Lastly, we may look at volatility of asset prices. An asset with a steady value can be held with little risk over an extended period of time. As the price will be expected to remain steady, or move predictably, there will be little added risk in temporally extending the holding period. However, if an asset has a tendency to be quite volatile, a significant cost will be added to an individual by holding the asset for a longer period of time. As more risk will engulf the final value of the asset, individuals will be less inclined to hold these volatile assets for extended durations.

These three aspects all factor to influence the expected holding period an asset will remain in an investor's portfolio for. Savings do not represent a static renunciation of non-consumption over consumption, but have a temporal aspect which becomes clear. By stressing this temporal aspect, we can enrich the analysis of the temporal aspect of saving, and hence, investment.

4. Physical and Economic Lives and Savings

An individual's physical life has a definite limit that affects their saving decisions. Typically, we see that the largest portion of an individual's savings are dedicated towards retirement. As an individual progresses through their life, the time remaining until this point is reached is reduced. Hence, we see a tendency for a generally shortening time horizon concerning the length of time an individual will save their funds for as they near the termination point of their lives.

As an example, we may compare two individuals. Each is saving their money with the same goal in mind – to redeem the proceeds of this saving upon retirement at age 60. The first individual is a 25 year old recent University graduate. They will expect to keep their funds saved for a total of 35 years. However, this same individual's parent of 50 years of age will only expect to keep their savings invested for another 10 years. *Ceteris paribus*, we can see the effect that this will have on the financial assets that are priced by these two distinct individuals.

One particular conundrum that becomes apparent at this point is that investors with longer time-horizons before they wish to redeem their savings exhibit either one of two features relative to shorter time-horizon investors. First, as the duration to redemption increases, uncertainty concerning these future values increases likewise. As a result, longer-duration asset holders may require a higher rate of return to compensate them for this increased uncertainty. Second, we see that the corollary of our first case is that if the expected rates of return are the same for both groups of time-horizons, that those with a longer time-horizon will be less inclined to purchase assets at the same price as the other group as they will not be rewarded for their increased uncertainty accordingly.

First, it may prove instructive to reassess equation (7.8) again:

$$V_0 = D_1/[1+ k] + D_2/[1+k]^2 + \ldots + D_N/[1+k]^N + V_N/[1+k]^N \tag{7.8}$$

If both groups of buyers (long and short time-horizons) are both actively purchasing the same asset, V_0 must be identical as prices will be homogeneous. However, with differing values of N, we see that expectations concerning dividends, D, and the future selling price, V_N, must differ in either their magnitudes, or the confidence in their prediction. As the duration of asset holding increases, the uncertainty premium placed on the expected rate of profit must likewise increase. If future expectations concerning the dividend stream and the future selling price are identical, expected profits must also be equal. However, if future profit expectations are equivalent, individuals with lower time-horizons will not be compensated for their increased holding duration accordingly. A spread in real profit rates will occur if not compensated by a differential in profit rates.

The corollary is that individuals with longer time-horizons may have expectations of greater profit rates (i.e., the dividend growth rate and the growth in expected future selling value) they will bid the present value of these assets higher than would be warranted by the individuals with shorter time-horizons. As a result, short time-horizon investors may not participate in purchasing these investments, as the increased present values conflict with their expectations of future dividend and firm value growth rates. As their expected profit rate will be reduced in this case, they will refrain from purchasing these assets.

We may conclude this by stating that expected profit rates will determine what time-horizon group partakes in purchasing a given asset. If profit rates are dominated by expectations of highly uncertain future earnings growth, present values of assets will be too high to adequately compensate the profit expectations of individuals with shorter time-horizons.

Multiple Savings Goals

It is not only distinct individuals who have separate time horizons, but the same individual will often have multiple horizons concerning distinct future consumption desires they have. An individual will not have a single investment goal, but several of distinct expected dates of realization. An individual will simultaneously save and invest for multiple events.

Hence, in our previous example, a recent University graduate of 25 years of age will have multiple investment desires that will be saved for at any given time. Hence, they may simultaneously be saving to purchase a car, a house, and to retire. As a result, they will require several assets to invest in to satisfy these differing durations.

If all assets shared a homogeneous expected return across all time durations, there would be no difference in diversifying between distinct assets. However, as we see that there are different expectations and that we will have for each asset and as a result, different expectations concerning profit rates across differing time horizons. As a result, there will be a series of investments that we will simultaneously be enticed to invest in, conditional on expected holding rates and resultant profits.

Economic Versus Physical Life

An individual thus invests according to their expected physical life. However, there is also the question of the economic life an individual expects their assets to serve their needs for. As a result, the economic life concerning an individual's investments may well differ from that of the physical life an individual has with which to enjoy these assets.

In fact, we see examples everyday of this occurrence. Foundations are set-up with the purpose of providing funds through their investments for extended periods of time that exceed that of any particular manager within the foundation. Parents may have more funds at their retirement than they expect to use over their remaining lives. As a result, investment decisions factor not their personal remaining lives, but also of those they expect to leave their remaining savings and investments to.

Investment companies provide a particularly interesting example. An investment company as a continuing entity can make investment decisions that far exceed that of the lives of the individual manager or clients. As a result, individuals can choose between different companies that manage investments and choose that which is most consistent with their own personal investment time-horizons.

The conclusion we may draw is that it is not the *objective* physical life of the individual that is of concern to the investor, but the *subjective* economic life of the *use* of the assets which they have at their disposal. As a result, investment decisions are made in light of a subjectively determined time-horizon, which will at times differ widely from the life that an individual has remaining.

Conclusion

As a result of the individual's subjective time-horizon, different assets will be chosen to invest in. Saving is not a decision undertaken with no expectation of *using* the savings in the future. Instead, we see that it involves a period of waiting that implies that savings are consumption renunciated for a certain time period. This time period in question is not necessarily confined by the individual's physical life, but may extend beyond to include that of ancillary individuals. These individuals need not be objectively identifiable, as is the case we find with foundations and trusts that invest for the future concerning abstract individuals.

One significant conclusion we may draw from this analysis is that, as different individuals will have differing expectations concerning future dividends and asset selling prices, while present values remain constant for all, differing profits rates will occur. Individuals with longer time-horizons will require a higher profit rate to compensate for the increased expected uncertainty they will endure over the life of their investments. As a result, individual's with shorter time-horizons will be bid out of the market for some assets, as the relative profit rates will be priced so high that present values will be too high for shorter time-horizon individuals to justify purchasing. As a result, different assets will be priced relative to differing time-horizons. Individuals will choose between these investment alternatives depending on the respective *subjective economic* time-horizons they experience.

5. Profit Maximization and Holding Period

Individual conditions concerning financial assets will influence the holding duration as well. An expected holding duration will not depend solely on the physical aspects of an individual concerning their personal and economic life, but also those factors concerning the individual financial asset. One of the most critical of these asset specific factors will be, as we have previously explored, the quality and changes of the uncertainty-bearing entrepreneurs within the firm.

As this class of entrepreneur effects the future changes that the firm's direction will undergo to serve the needs of consumers better, they will drastically alter the future selling price that a firm will achieve. Individuals will have to forecast any future changes that occur in this entrepreneurial class and factor for it accordingly as they construct their plans about when they will finally sell the asset. Changes forecast in a firm's management will not be viewed in isolation however. Much like Keynes' (1936) beauty pageant, it will not matter at the time that an asset is purchased when an individual expects future changes in management to occur. However, the expectation of when others expect this change to occur will have a drastic affect on the present value of the asset. Hence, the problem becomes one of predicting the expectations of the others and acting accordingly.

Assume a situation where it is widely known in advance that the CEO of a company will be retiring in 10 years time. It is also widely expected that this retirement will adversely affect the future profitability of the company. As everyone knows this fact, there will be an incentive to sell the asset prior to others doing so to maximize the return. However, if everyone shares this same belief and tries to sell before all others, there is a problem created whereby-if everyone tries to sell the asset before others, and endless regress occurs whereby eventually, no one would buy the asset to begin with.

However, we see that the question is not one of an endless regress once we realize that dividend payments made in the interim create an incentive for an individual to continue holding onto the asset. Hence, an individual is faced with a trade-off between the profit they will earn on the dividend stream, and the loss that they may have to take if they wait a longer holding period to sell the asset, during which others might sell in the expectation of a future change in management. The existence of a relatively high dividend will allow individuals the ability to hold onto an asset for a longer duration, while still earning a profit, despite the risk that others may sell the stock and place downward pressure on the price.

As individuals will purchase a financial asset to maximize their monetary return, we have seen that the holding duration is an integral part of the valuing process. However, we see that this is also affected by the factors concerning the individual asset in question – management and supply/demand conditions chief among the concerns. To the extent that an asset pays a dividend stream in the present, an individual will be compensated for holding a stock for a longer period, even in the face of selling in anticipation of negative future changes to the stock price placing downward pressure on it.

6. Volatility and Expected Holding Period

Finally, we may wish to look at how expected volatility affects our holding durations of financial assets. It is important to remember that previously we saw that volatility has no direct influence on asset prices. The reason is that volatility, as defined as the likelihood that an asset will yield the return expected when sold becomes a non sequitur for the investor. Volatility in the interim period is not of significant concern, provided that it does not affect the ultimate expectation that the final selling price will be obtained at the date when it was previously thought to. However, volatility may severely affect the expected holding duration of an asset. Those financial assets that experience what is conventionally called "volatility" may cause an investor to become less certain about the future price of their asset at selling dates more temporally distant from the present. As a result, a shortening of the holding period may result.

We may think back to chapter IV, appendix A, where the two roles of money – medium of exchange and store of value – were assessed. Commodities that exhibited high degrees of long-term volatility, but short-term stability were shown to be very good to use as media of exchange. Conversely, those commodities that were volatile in the short-term, but stable in the long-term were shown to be very good to use as stores of value, but deficient when used as media of exchange.

The reason this distinction arouse was that for a store of value to be of use, there must be a certainty surrounding its value at that future date in time when it will be used. Volatility has a direct effect on changing the duration that an individual will have certainty concerning the future redemption value of an asset.

We may say that, *ceteris paribus*, there is a tendency for highly volatile assets (as we define volatility as uncertainty concerning the future price or value) to be held for shorter holding periods than non-volatile alternatives. As an example, an individual may hold a bond which has a guaranteed value, in nominal terms, at the end of its holding period. This holding period may be significantly long – 30 years or even more in some isolated situations. The reason that durations of this length can be offered is that volatility is minimized through a fixed nominal redemption rate. However, the holding period of stocks, which lack any such fixity in long-term value, is significantly shorter. As stocks exhibit great volatility over long holding period, there is little certainty as to what the future redemption price will be. As a result, purchasers of these assets will only hold them for relatively shorter periods of time than more certain bonds.

Hence, we may make the conclusion that the greater degree of subjectively expected volatility an asset will exhibit, the shorter the period of holding this asset will be. This is not the more commonly considered direct linkage between risk and reward, but an indirect linkage between uncertainty and holding duration. Any asset that is purchased will be held with an expectation of what its future selling price will be. This necessarily arises so that relative expected profit rates can be calculated and compared. Any future uncertainty caused by increased volatility will cause the expected holding period to decrease to a point where it is possible to assess the future selling price.

7. Conclusion

By looking at the two sources of profit holding a financial asset entails, we can begin to identify exact value adding locations. The conventional Gordon dividend discounting model makes several assumptions which make for ease of use, but compromise the usability of the model in the real world. Two of the most egregious errors are that dividends are the sole source of future profits, and that a single risk-free rate can be used for all future cash-flows.

We have broken the conventional model into two separate and distinct parts. In the first we see that a future flow of dividends is valued to a certain point in the future. At this future point, the sale of the asset will occur which will result in the renunciation of any future dividends or cash-flows from it. When these cash flows are discounted back to the present time, we see that a single risk-free rate assumes that the originary interest rate of interest will remain stable over time. Instead, we see that there is an expectation that the originary rate of interest will not only be non-constant for the same duration between different temporal points, but it will be non-constant among different durations at the same temporal point. As a result, there are two options that face the investor when discounting future cash-flows from an investment. The first is the use of an originary interest rate of equal duration to the expected holding period, or time until a dividend is received. However, depending on future interest rate expectations, an investor may chose to discount at a shorter-term rate, which will be rolled over continually if this is a more profitable alternative than contracting a longer-term rate at the initial time period. The greater of these two options will be chosen to discount future cash flows.

Second, by breaking the pricing model into two constituent parts, we can identify the entrepreneurial sources of financial asset value. The first class of entrepreneurs – uncertainty-bearers – affect the values in two ways. First, they provide the means-ends framework that other employees – risk-mitigators – function within to maximize profits. Hence, this first class of entrepreneurs provide the maximum *potential* profit that can be attained within the firm, and hence, also provide the maximum boundary that a dividend could be issued for. Secondly, the future selling price of the firm will be dependent on this class of entrepreneur to guide and direct the firm in the correct future areas that are cohesive with consumer demands. The risk-mitigating class of entrepreneurs affect profitability in the present, and hence, the magnitude of the dividend that may be issued. More efficient entrepreneurs (i.e., those that mitigate more risk) will allow a firm a higher rate of earnings from which to issue dividends.

However, of equal importance is the amount of time that an individual expects to hold an asset for. A longer holding period implies a greater amount of dividends will be earned, but likewise a loss in certainty regarding the payment of these dividends will be achieved. This expected duration of holding time for an asset gives a temporal dimension to the concept of time preference that was explored earlier. We see that not only is there a renunciation of consumption in the present for a definite amount of time as defined by time preference, but there is also the temporal aspect which establishes what this definite amount of time is. This was shown to be determined in two ways. First, the physical life of an individual places constraints on the amount of time with which they may renunciate consumption for, before the physical necessity to partake in consumption before their physical life ends makes necessary the end of the renunciation. However, the possibility opens that individuals save not only for themselves, but also for a myriad of additional individuals or purposes. Hence, foundations function to save and invest not for the lifetime of the managers involved directly, but for an indefinite period of time. The distinction between the physical and economic life becomes apparent.

Two factors affect the holding period that an asset will be invested in for. The first is the characteristics of the particular asset. As monetary profitability will be desired to be maximized, an asset will be sold at that point in time, not necessarily solely concerned with the individual's own time-horizon characteristics, but also those that maximize the expected profits of the asset. Second, we have seen that volatility affects the amount of time an asset will be held for. One necessity of purchasing an asset is having an expectation of its future selling price so that it may be compared with other available options. As assets exhibiting high degrees of volatility complicate the expectational process, the expected holding period will be decreased to compensate for this.

The pricing process for financial assets is a complex process. By looking at not just the objective factors governing the behavior of these assets we can get a much richer analysis. Incorporating personal time-horizon considerations, as well as entrepreneurial theory, we may start to see the more complete and subjective way that these assets are priced.

VIII. CONCLUSIONS AND FUTURE RESEARCH

1. Conclusion

Economics may be one of the youngest of the sciences falling under the umbrella of praxeology, but there are new contenders for this title. Finance is a distinctly separate field, which, although relying on a basis of economic knowledge, proceeds in a way which must be viewed as a distinctly separate field. Economics proper is that field which is concerned with humans acting to achieve their ends in the world concerning their desires. This particularly makes use of the concepts of the goods and means which are used to directly service those ends. However, finance is a distinctly different process. The goal is to analyze this same framework, but from the viewpoint of ancillary means which *derive* their value from these goods. This derivation implies finance, or more correctly – *financial economics* – forever remains a subset of economics. However, the focus of this study is separate; financial assets are one step removed from the consumption/production process.

The present work has built a foundation of a new theory of finance, progressing in two parts. The first volume concentrated on providing an overview of the two most salient theories that provide the foundation for modern financial theory – the efficient market hypothesis and the capital asset pricing model (EMH and CAPM). It is enlightening when we see that although there is discontent with these concepts, no purely theoretical refutation of them has been provided. Instead, we have seen and reviewed a plethora of empirical studies that illustrate problems with these theories, yet they remain the backbone of the financial literature edifice. Hence, a myriad of new theories have been forwarded that try to rectify these problems, however, they remain within the pre-established framework and fail at providing a real, usable alternative.

The second half of the first volume looked at the core theoretical issues that need to be addressed in order to see where the specific faults of EMH and CAPM lay. Faulty conceptions of time, efficiency, interest rates, risk and uncertainty have all been seen to be at the heart of the problems that are empirically surfacing with these two theories. The flaws have arisen primarily due to a faulty methodology that has been neglected since the field of finance was created. With little heed paid to methodological details, there has been no attention given to the micro-foundations of finance, which have been developed in the second volume further. Indeed, this lack of micro-foundations has been heralded by some of the foremost academics in the field. As Ross (1987: 29) views the critical difference between the two fields of finance and economics:

> Finance uses the modeling framework constructed in economics, but, within this scaffolding, finance has taken a different methodological perspective. It is wrong to characterize finance, or financial economics to be formal, as simply another of the specialty areas of economics... While finance is specialized in its focus on the financial markets, the differences between economics and finance only begin there. The principle distinction is one of methodology rather than focus.

Hence, finance has come to be seen as wholly distinct of economics, and not in need of its considerable insights. This disconnect has directly resulted in the problems now surfacing.

However, we have shown, the single most egregious omission in all work building off the EMH and CAPM is the lack of entrepreneurial foresight. With this key function missing – entrepreneurship – the body of work has been wholly separated not only from the field of economics, but the larger body of work concerning human action. With no agent functioning to discover and utilize market prices, there is no basis for these prices to be created in a way consistent with living, rational beings. It is not even a question of *homo agens* being replaced with *homo economicus*; even the latter has been assumed away completely.

The second volume of this work has tried to create a new foundation for finance centered around this previously missing entrepreneur. Some of the readers who have followed this work to the end may question the necessity of the first part of this work – *Action and Its World*. This first part may be the most difficult to read, but also provides the most important insights, the completion of which the following parts could not have been achieved.

Risk and uncertainty have been given the bulk of attention in this first part. The reason is that risk and uncertainty serve to create nearly all the problems which confront the economist. However, we have at the same time demonstrated the source of true uncertainty to be not ontological in nature, as many Keynesians might assume, but epistemological.[339] It is the limits on our reason that result from the incomplete and necessarily imperfect knowledge we have as humans which causes us to act in a fundamentally risky and uncertain world. The only way we can solve this problem of epistemological

[339] See Davidson (1993) for this distinction. Interesting, Davidson (2003-04: 264) correctly characterizes the Austrian method to dealing with this uncertainty as the market process, despite failing to realize the true source of uncertainty itself.

uncertainty is through the market process. Huerta de Soto (1992) provides a particularly deep look into the entrepreneurial function spreading knowledge and becoming the *deus ex machina* that accomplishes this task. In fact, the particular dues ex machina – the entrepreneur – has become the focal point of this second volume.

As the entrepreneur has been developed, we have seen that a three-fold role is served. None of these three particular roles could ever be fully had by a single individual, for if such a possibility occurred, this individual would be found to live in a state of *in*action. The first role is the bearing of future uncertainty. This is the typical Misesian (1957: 320) entrepreneur, one who must look to the future with the eyes of an historian. Second is a risk-mitigating agent. As risk is reduced efficiencies are increased and production and want satisfaction can be expanded accordingly. Lastly, we see that provided that these two functions are not perfectly endowed in an individual, resources will be required to be supplied. This is the third function of the entrepreneur, and one that is of particular interest to this work as it is where the role of financial assets becomes so important. Due to the non-contractual nature of human accomplishments, we see that individuals serving these roles in differing capacities will unite under the common umbrella of a firm to increase the benefits of each others' abilities. This is the firm as we commonly know it. Previous theories have assessed specific *functions* of the entrepreneur, while side-stepping the actual *reason* that they provide these functions. It is for this reason that I would refer to entrepreneurs as viewed under prior theories as "rebels without a cause." They are surely swimming against the stream, but no formal heed is given as to why they are doing so. Once we realize the ultimate cause is to increase consumer want satisfaction, the trichotomy of entrepreneurial functions become not only clear, but necessary.

Hence we see that any theory aiming to price a financial asset must necessarily include these entrepreneurial roles as *they are the financial asset*. Firms, claims on their assets, and the derivative values do not exist in absentia the actors that comprise these institutions. Instead, we see the human element is at the heart of not only the firm, but the value derived thereof.

Additional considerations have been given to those produced means of production that entrepreneurs use to create consumers' goods – capital goods. In particular, Austrian capital theory has been expanded in two significant ways. The first is by explicitly paying attention to *fixed capital* in the structure of production. Previously only circulating capital has been given attention, which has generally ignored concerns such as depreciation. Additionally, the use of Hayekian triangles to illustrate the productive structure has given rise to a belief that production is action which occurs in "stages." One particularly terrible result has been that consumption has been ignored as no distinct "stages" of consumption are identifiable. However, once we view production for what it really is – a series of actions joined in the common goal of producing a consumers' good to increase consumer want satisfaction – we see a definite consumptive counterpart exists. As goods are furthered along the structure of production, each individual action increases the value that a unit of circulating capital will have concerning consumption. Likewise, as a consumers' good proceeds along the structure of consumption, each individual action (or use) will remove value from these goods until they are valueless.

This insight opens of up new conclusions from Myrdal's (1939) monetary equilibrium. For we can now see graphically that savings must equal investment in the long-run. Austrians have generally neglected consumption previously, with the result that its full importance in capital theory has been ignored.

Using the structure of production as a micro-foundation for finance has many benefits. In this way, we can rectify the issues that Shostak (1997: 30) raises whereby much confusion in the finance arena stems from the disjointing of the real and financial sectors. As Lachmann (1956: 86) correctly demonstrated, the capital markets are not markets for physical goods, but the titles to them. As such, we need a theory of finance that recognizes this distinct nature of the two sectors, but also the need to base the one on the other. This linkage can be provided through two interacting structures – production and financial. Firm's are easily able to be diagrammatically shown on the structure as the joining of several distinct actions along the productive structure. This allows us to illustrate concepts such as vertical and horizontal integration. As firms will have more or less optimal degrees of these integrations, value will be added or removed accordingly by investors. Likewise, we see the equilibrating tendencies that exist as firms along the productive structure will have, in the long-run, equivalent profit rates. The entrepreneurial influence assures this as these individuals move from sector to sector, and firm to firm, in search of higher monetary profits.

Lastly, we have culminated all these insights by incorporating them into the common dividend discount model for pricing financial assets. There are three significant differences with our new portrayal of this concept. The first is that dividends and the future selling price of an asset are distinct factors. The price of an asset cannot be collapsed into its dividend stream without assuming away the most important element of the firm – the entrepreneur. Instead we see that individuals price these assets with these two distinct factors – dividends and the proceeds from the future sale of the asset – in mind.

Changes in the structure of entrepreneurship within a firm will affect both these factors accordingly. It is for this reason that we cannot view the entrepreneur as a unique agent with a solitary goal, but a heterogeneous mix of skills which create and alter a structure of entrepreneurship to add or negate value from financial assets. Second, we see that a temporal dimension of time preference becomes importance. The typical loanable funds model used to illustrate interest rate formation under time preference theory leaves no room for a temporal dimension to savings. That is to say, savings represent not only the renunciation of consumption in the present, but also for an expected period of time. Hence, individuals who invest in financial assets do so not fully unsure of the temporal element this will be done for, but ever mindful (even if in only a vague way) of the time an asset will be held for before it is redeemed for cash. The affects on financial assets have been addressed accordingly.

Lastly, the particular interest rate used to discount future cash flows must be reassessed. We have seen that a single unique risk-free interest rate is commonly used to discount future cash flows to the present. However, the appropriate interest rate will be more complex than this simplistic approach. Hence, entrepreneurs are first faced with a two-fold choice surrounding the discounting of cash flows. Either they can use the appropriate rate of interest of equal duration as the investment, or the sum of a series of rolled-over short-term rates. Second, we see that the interest rate will not be assumed to remain constant. Instead, it must be based on the originary interest rate which is subject to continual change. As a result, the discounted dividend model cannot be collapsed using a single interest rate, but must account for the choice of options facing entrepreneurs (locking in a long-term rate or continually rolling over a short-term rate) and the dynamic change that will engulf future expectations concerning the originary rate of interest.

These are hence the main conclusions that the present study yields. However, it should be noted that to reach these conclusions, many changes in the existing edifice required modification. Not only have we rewritten many mainstream finance theories, but the prevailing Austrian doctrine has been reassessed, with changes made where necessary. At the end of this chapter we will summarize the individual conclusions that have been developed herein, and that will be further expanded upon in future articles. Indeed, these articles will fill my own research agenda for some time to come

It is of passing interest that we may address a point raised by Cochran and Glahe (1999: 29):

> Keynes started from a model similar to Wicksell's and turned toward the financial sector to look for the cause of economic problems. Hayek and von Mises, in their Austrian business

> cycle Theory (ABC), took the Wicksellian change in the money rate relative to the natural rate as the cause of the cycle. The theory turns toward the real sector to analyze all the effects from this monetary cause. Hayek disaggregates the real sector; Keynes disaggregates the financial sector.

The present work has been heavily influenced by Hayek, especially his *Pure Theory of Capital*. It is only by providing the correct micro-foundations of capital and production that we may move to the further challenges posed by financial assets. Hayek had originally intended his *Pure Theory* to be a two-part work, however, the second piece was never completed. In lament of this fact, Hayek (1994: 84-85) had the following to say regarding his work on capital theory:

> I rather hoped that what I'd done in capital theory would be continued by others. This was a new opening which was fascinating.... [However], [n]o one has done what I hoped would be done by others.

I don't believe this present work has filled those giant shoes, but I do think it is a leap in the right direction.

2. Summary of New Contributions

1. A thorough theoretical critique of the efficient market hypothesis (EMH) has been provided. First, it has been shown that an erroneous view of time has been used. By focusing on Newtonian (i.e., linear) time, the real effects of change through time are neglected. Second, EMH neglects uncertainty from occurring in the future, allowing for only risky deviations from the present expectations. Third, a neglect for the entrepreneur's role leaves EMH unable to account for how information affecting prices is spread through the market. Fourth, EMH has been, from inception, a result searching for a theory. Unfortunately, this lack of a theory has had detrimental effects as it has left the hypothesis on a shaky foundation. Fifth, information and knowledge are confused, as it is assumed that they are equivalent. No allowance is made for the personal interpretation of information as knowledge, an occurrence which allows for undiscovered opportunities to abound (which EMH rules out). Sixth, EMH rests on a concept of static efficiency, an assumption which focuses on correctly pricing the information *as given*, but neglects the future possibilities that may occur. Finally, EMH assumes that prices are independent and random (i.e., that they follow a random walk). This erroneous assumption implies that the individuals who create these prices operate in a random, purposeless way as well, and assumption which is not unobjectionable to discard.

2. A thorough critique of the capital asset pricing model has been provided. First, time is used as a homogeneously valued factor which is used as a comparable standard ex ante, a view point which is shown to have detrimental effects. Second, CAPM treats uncertainty as a merely risky element, with future risk-return profiles (and their associated probabilities) known in advance. Third, CAPM treats individuals as price-takers, unable to account for the entrepreneurial element that searches for price disparities and acts on them accordingly. Fourth, the methodological approach has been almost purely via empirical testing which has left the model with many logical inconsistencies – such as those being expounded herein. Fifth, CAPM posits choice as occurring and being arbitraged between only two variables – risk and return. Even new multi-beta pricing models fail to account for the plethora of price determinants which are subjectively perceived by the individual, and acted on accordingly. Sixth, the use of the risk-free asset is found lacking. Ex ante, there will always be an unknown amount of uncertainty concerning an asset's future purchasing power, thus complicating Tobin's (1958)

assumption that a risk-free bond can be substituted for interest-free cash with no negative effects. Finally, risk as the standard deviation of return is found lacking justification. Investors are not concerned with intermediate price movements (to a point), but instead are concerned with the selling price they expect to attain at a future point in time when the asset is sold. To the extent that interim volatility does not disrupt the present expectation that this future price will obtain, short-term volatility will not affect the pricing process.

3. The privatization of time through interest rates is the means we use to make inter-temporal co-ordination possible. Just as money prices allow for static co-ordination, interest rates become the necessary factor to create inter-temporal co-ordination. As action always involves a good and a temporal element, time becomes embodied in the use of specific objects. Through interest rates, we assign a value to our time, and hence, are allowed to trade it to others for ends we value more highly.

4. Probability and time are inextricably linked. If there is a given probability of an event occurring, it must be constrained by a temporal element. If we assume that an action could be continued for an unlimited duration, an positive probability becomes a certainty. Hence, probability as used as a risk-measure in finance theory is conditioned by the temporal element that the risk is expected to be experienced over. By way of example, to say that a company has a 5% chance of bankruptcy is meaningless without a temporal constraint. If we held stock in this fictional company forever, we would be assured of bankruptcy at some point in time owing the positive probability of this occurrence. However, if we know that the risk of the company going bankrupt is 5%, and we expect to hold the stock for only a finite amount of time, we are no longer certain that bankruptcy will obtain during our holding period.

5. Risk as defined by a variation of return is a fundamentally flawed conception. Individual investors are concerned with the expectation that the future price of their asset will remain consistent with the expected profit rate when they purchased the assets. Variation in the interim holding period is of no significant concern, provided that it does not negatively affect the expectation that their desired profit rate will obtain.

6. Logic is an absolute base which the mind applies within each individual's unique knowledge set to

create decisions. Logic is given to us as humans, and is the defining characteristic that we exhibit over animals. It represents absolute, immutable knowledge. As a result, we base decisions on two distinct knowledge sets. An a priori and complete logic set, which is utilized based on a limited and imperfect learned knowledge set, act as the two knowledge foundations that decisions are based upon.

7. The entrepreneurial function rests on the two fundamental concepts of risk and uncertainty. Misesian entrepreneurs bear the uncertainty of the future by foreseeing this unknown eventuality, prior to anyone else. However, entrepreneurs that mitigate risk may also move the market forward. While the previously mentioned entrepreneurial function delivers want satisfaction to consumers that they themselves do not know about, the second delivers established wants in a more optimal way. Hence, risk represents a set of known outcomes, which we know the probabilities of occurrence for the total set, but not for the individual components. A risk-bearing entrepreneur may reduce the risk of loss in a production process, and hence correspondingly reduce the price at which the produced good may be sold for. Once we see the end that an entrepreneur strives to achieve – the satisfaction of consumers' wants – we may see that this may be achieved in two ways: 1) uncertainty bearing, and 2) risk mitigation.

8. Uncertainty represents an absolute element – we cannot speak of differing degrees of uncertainty. When an individual says that a decision is being made under conditions of more uncertainty, they are always referring to the degree of *felt* uncertainty. Hence, when we look at decision-making under uncertainty, it is only the felt uncertainty that is relevant. Additionally, uncertainty cannot be reduced, however, the degree to which it is felt to be so can differ substantially, altering the confidence we have during certain circumstances.

9. The entrepreneur as previously defined assumes that the perfect entrepreneur is achievable. One consequence is a neglect for the necessity of resource ownership. Once we place the entrepreneur within the disequilibrium setting which necessitates their existence, we see that ownership of resources becomes essential. Hence, a new source of entrepreneurial profit is accrued to those resource providers – those individuals who supply the necessary resources to allow other entrepreneurial classes to function.

10. Originary interest as being caused by time preference has faced criticisms by some who view it as only an historical tendency. We forward a new defense based on a better understanding of *real* time. Hence, time preference becomes a necessity for individuals when faced with the realization that time forever is a declining balance, and unable to be extended. The wants which we continually want fulfilled *at or during a specific time* will require us to have a preference for their achievement earlier rather than later in order that we may economize on the scarcity of time. Hence, time preference is not a universal of human action. Instead, we see it is a universal for those *aware* of time's true, limited nature.

11. A trichotomy of action exists – consumption, production, and exchange. This final category is forwarded to contrast against some of the new literature (especially Barnett and Block 2007) who forward that action is binary – consumption and production. Instead, the three categories are shown to be categorically different, with further implications as follows below.

12. Productive are those actions which create greater want satisfaction upon completion. As production is always separated from consumption by a temporal element, we are forever shrouded in uncertainty as to whether the production will lead to greater eventual want satisfaction. As we can only know ex post if production was successful, ex ante actions must be divided between exchange and consumption, while ex post the conventional trichotomy becomes applicable.

13. Following a trichotomy of actions, we find that a trichotomy of goods must likewise exist – consumers' goods, capital goods, and exchange goods. This counters current literature suggesting that goods' nature is necessarily limited to a binary distinction (i.e., consumer versus producer).

14. Production goods will only be known to have been productive ex post. Prior to the fruits of their use being made evident, they will have no value as capital goods, except to the extent that the market estimates that they will eventually yield valuable consumers' goods. Hence, ex ante the price of a capital good will be the market's best estimate of the productive capabilities that the capital good will produce over its useful life.

15. The structure of production has been augmented with a corresponding structure of consumption.

This should interest Austrian economists for two prime reasons. First is the emphasis that consumption value has on the productive goods used in their construction via the imputation of value. Second is the fact that all production is undertaken to physically supply individuals with consumption goods. Elements of the structure of consumption are assessed.

16. Hayek's use of the “production-time” as an axis on his triangle has had unfortunate consequences. Chief among these has been the influence this has had on present-day Austrian macroeconomics, particularly as portrayed in Garrison (2001). As production is shown to be taking place in production time, the x-axis of the Hayekian triangle has been simplified to be phrased in units of “stages of production.” This generalizes the issue, and ignores the true essence of production – a series of individual actions adding value to a unit of circulating capital as it progresses further to that point at which it becomes a consumers' good. By addressing this issue, and modifying the triangle to be labeled in terms of production time as an almost infinite series of value-adding actions, we place the structure of production on much more solid footing, and salvage the use of the Hayekian triangle.

17. Hayek's use of triangles to portray the structure of production were made at the expense of losing focus of the consumptive aspects of the business cycle. This arises as a Hayekian triangle allows for structural shifts and changes in production, but it remains point-output, implying that consumption never undergoes alterations. This helps to explain one fundamental difference between Hayek's rendition of the business cycle, which focused exclusively on malinvestment, and Mises', which allowed for both malinvestment *and* overconsumption.

18. The structure of consumption is created to offset the structure of production. This is shown to be comprised of three types of consumers' goods – services, non-durables, and durables. These are integrated giving three considerations to the shape of the structure. First, the value component that will be provided by these goods determines the height of the triangle. Second, the length of time that these goods are expected to provide services for will approximate the length of the x-axis of the consumption triangle. Lastly, the degree of depreciation on durable consumers' goods will require a continual reinvestment to maintain the value provided by these goods. These factors provide trade-offs as more value generally entails a more costly good, while the same holds true for a longer-lived versus a shorter-lived consumers' good. Likewise, more durable goods as a general rule are also more costly to

produce vis-à-vis less durable goods. The degree of time preference an individual exhibits will determine the trade-off between the attainment of these factors, and the renunciation that must occur in exchange for them.

19. Five factors stemming from the structure of consumption signal to producers the types of circulating capital goods to produce to maximize profitability. First, the exact demand that consumers wish satiated must be gauged. Second, the degree to which consumers wish this demand met must be produced (i.e., want satisfaction). Third, the period of serviceableness of a consumers' goods must be forecast correctly. Fourth represents the degree of depreciation that consumers will tolerate concerning the maintained value of a good. Lastly, the quantity demanded by consumers will dictate the quantity of circulating capital that will be produced. These five factors taken together will determine whether an equilibrium situation can be reached as producers strive to meet consumers' demands.

20. Alterations in the market rate of interest inconsistent with the true originary interest rate have a negligible effect on the existing stock of fixed capital. However, two important alterations are made as a result. Assuming a reduction in the market rate of interest, we may find that: 1) the production of circulating capital will be reduced as the expectation is that a decreased demand for consumers' goods will be forthcoming. A shift to fixed capital will result as it is forecast that goods of higher quality and durability will be desired, and 2) an increase in new projects will be undertaken to fully take advantage of the greater profit opportunities now available owing a lower “hurdle” rate (i.e., interest rate). As the shift towards fixed capital projects is undertaken, previously unprofitable projects in this field will now be started. This second effect will be more pronounced the further away from consumption the capital project temporally exists.

21. Interest rate reductions cause a primary, and secondary effect into the production of fixed capital relative to circulating capital. The primary effect is that which is normally explained through Austrian business cycle theory. A secondary effect stems from the lower interest payments enticing entrepreneurs to demand lower durability capital, as they finance purchases through short-term lease agreements becomes a relatively more attractive option compared to outright purchases. As a result, this lower durability capital must have a higher replacement rate to compensate for the increased depreciation of lower durability capital, a higher portion of fixed capital must be produced to

compensate for this secondary effect.

22. Lengthening the structure of production is an issue commonly viewed as a temporal lengthening of the production process. However, it has been shown that it is really a shift to increase the degree of fixed capital relative to circulating capital produced. This can occur in two ways. First, an absolute shift from circulating to fixed capital may occur, thus increasing the degree to which the production process is capitalistic. Second, fixed capital of greater durability may be produced, resulting in a lower demand for fixed capital production to replace the depreciated capital stock.

23. The pure theory of the entrepreneur was shown to be focused on two abilities. First was the role played by bearing the uncertainty of the future – to discover opportunities not known to exist. Second was the ability to mitigate risk, and hence undertake plans with a high level of efficiency which need not rely on any degree of uncertainty. These actions serve to increase consumer want satisfaction. No one individual could possess these qualities fully, and hence, the possibility of a pure entrepreneur existing is an impossibility. Firms are shown to be the natural outgrowth of risk mitigators and uncertainty bearing entrepreneurs as they try to synthesize the pure entrepreneur through contracting their services together within the boundaries of the firm.

24. Each entrepreneur within the firm works at either promoting a growth, or profitability, center. Risk-mitigating entrepreneurs are those primarily determining the profitability that a firm will have at any given moment through the level of efficiency they operate with. Uncertainty bearing entrepreneurs will be the drivers of growth, as they move the firm into previously uncharted territories. Continued uncertainties are essential for profits to be maintained, however, two caveats must be stressed. First is that the elimination of uncertainty will, at least in the short-run, allow for greater profits as risk-mitigating entrepreneurs may maximize their static efficiency within a closed-ended system. Second is that the continual discovery of new uncertainties drives the firm forward, and is the engine of growth for the future.

25. Employee remuneration is inextricably linked with the entrepreneurial role – whether risk mitigating or uncertainty bearing – within the firm. As a general tendency, we find that uncertainty bearing entrepreneurs (managers) typically earn higher wages than their risk mitigating counterparts

(i.e., lower level employees). However, this is only a general tendency. Higher level managers see the benefits of their future oriented actions accrue over a significant time period, hence, the profitability (or marginal revenue product) that they contribute to the firm can be quite high. In comparison, the profitability that a lower level employee provides is mainly in the present, and does not accrue over a significant time period. As a result, they are paid accordingly. Exceptions abound, for instance, Internet start-up companies may see lower level computer programers earn higher wages than the managers who provide the ability to see the future profit opportunities. However, if profitability cannot be maintained (or at least a valuable product produced) the firm will never live to see the manager's future potential profit opportunities.

26. The typical loanable funds model has assumed that durations of savings are equivalent. Using Garrison (2001) and Böhm-Bawerk's (1901) concepts of "saving up for something", we find that the expected duration, and hence availability, of savings plays an integral role in the analysis. Hence, the loanable funds model is not a purely two-dimensional trade-off, but will also incorporate a temporal component that includes the expected duration that savings will be available for.

27. Monetary equilibrium along the structure of production requires two sub-equilibria. First is that the rates of savings and investment will tend to equilibrate. Second we find that entrepreneurial profit rates must trend towards equilibration across industries of the structure of production, and also among firms within the same industry. As entrepreneurial profit is that earned in excess of the rate of interest of identical duration, we see the possibility of a firm suffering an accounting profit, despite earning entrepreneurial losses. Exogenously determined interest expenses that are inconsistent with the true endogenously determined interest rate will disrupt the entrepreneur's ability to gauge the appropriate rate of entrepreneurial profit, and may cause malinvestments as a result.

28. Firms will vertically integrate along the structure of production to the extent that the costs of using the command based organization of inter-firm resource allocation is less than that of using the market pricing system. Hence, firms may not be at an "optimal" level of vertical integration and have their value effected accordingly. Firms that are *too* vertically integrated will see a decrease in the value of their components compared to what they would be worth separately on the market. Likewise, firms can increase their value if they are *under* integrated by including additional productive steps along the

structure of production.

29. Firms as syntheses of pure entrepreneurs are not the end of the process aiming at creating this function. Individuals may also attempt to create pure entrepreneurs by holding portfolios of investments in distinct firms, thus accruing the benefits that each individual firm has. Limitations on these profits will be found in the existence of search costs still needed to create the portfolio. Also, as the distinct firms will still be dependent on the price mechanism, ancillary costs will not be fully avoidable (i.e., marketing, advertising, etc). Individuals can influence firms that they hold as a portfolio to integrate with one another through their rights as shareholders at annual meetings designed to elect directors who are essentially the uncertainty bearing entrepreneurs at the firm, driving it to new, unforeseen perspectives.

30. Adding a *z*-axis to the structure of production allows us to view production not solely in value terms, but physical terms as well. Although the significance of the value creating aspect of production cannot be over-emphasized, the addition of quantitative considerations also cannot be overstated. Previously, we showed that entrepreneurs must estimate four qualitative factors concerning goods' remanded by consumers, plus the quantitative consideration of how much of a good will be demanded. By adding the physical production axis to the structure, we may see how this quantitative component differs at different stages, as well as measure excess capacity that firms have at each given stage.

31. Horizontal integration may be defined as that percentage of a good's market an individual has control over. Ranging on a scale from 0 – 100 percent, there are gains and losses to be made by integrating further. Gains may be made by capturing a greater share of a profitable market. However, the risk is continually run that a firm will see itself shouldering a large loss if the good in question becomes unprofitable through unforeseen shifts in consumer demand.

32. The investment decision facing investors is a two-fold process. First, the time preference trade-off dictates how much savings versus consumption will be undertaken in the present. Next, a liquidity preference trade-off will dictate which of three asset classes this savings is channeled into – money, liquid assets (i.e., financial assets), and illiquid assets (i.e., capital goods).

33. The pure entrepreneur previously defined lacked any resource requirement. This followed from the assumption that a pure entrepreneur could be created as a firm which never suffered loses, an occurrence which would negate the need for resources. Failing this possibility, and once we accept the possibility of entrepreneurial losses, we must allow for resource dependence. Hence, a third class of entrepreneurs is necessary – resource providers. These individuals are the equity and bond holders in a firm who provide the necessary resources for the other two entrepreneurial classes – risk mitigators and uncertainty bearers – to function.

34. The model of decision-making under uncertainty developed in book I, chapter IV had an inner-range of event values which represented those that an individual would fully expect to obtain. One previous conclusion was that as felt uncertainty increases, the inner range will likewise increase as an individual can no longer be certain that a limited number of events will obtain, and instead will have to widen their expectations accordingly. This can be augmented and developed using cash holdings as a proxy for felt uncertainty. Hence, larger amounts of cash holdings translate to larger degrees of felt uncertainty.

35. In addition to the previously mentioned equilibrating tendencies concerning profits, with the addition of a resource provider as entrepreneurial class, a new condition must be considered. The profit rates on the three general entrepreneurial classes – risk mitigating, uncertainty bearing and resource providing – must tend towards equilibration in the long-run. Hence, excess returns in one of these three roles signal to other entrepreneurs that more resources are required accordingly. Additionally, the difference between equity and debt financing by the resource providing entrepreneur has been discussed. The first two entrepreneurial classes will generally prefer to finance through debt, which entails a set limit to the duration that they must share entrepreneurial profits with the third entrepreneurial class. As a result, the yearly return on debt will generally be higher than equity, although owing the perpetual duration of equity, the aggregate profit may eventually exceed that earned on debt.

36. Fresh funding in the real sector will filter through from the financial sector. If we assume that 100 percent of funding is through equity issuance, a price equilibration could only be obtained through a continual equilibration of share issuance (supply) and demand for purchases. Share issuances are

inelastic and discrete, which implies that at any given moment, supply and demand conditions may be far from equilibration. If fresh share issuance is unable to keep pace with increased (decreased) demand for share purchases, increases (decreases) in nominal values, as well as price/earnings ratios will result.

37. Time preference dictates not only when we spend our money (i.e., future versus present) but also what type of money we prefer to hold. Money's two primary functions – medium of exchange and store of value – are two sides of the same coin; a store of value is nothing more than an intertemporal medium of exchange. As a result, the expected volatility that a money will have in the future will be compared to the time preference of an individual, and hence, when they expect to spend their savings, to determine what degree of volatility (and hence what good money is made of), they deem acceptable in a money.

38. The dividend discount model of pricing stocks relies on the expected future selling price being collapsed into the dividend stream accruing over the stock's holding period. Although this may be a viable assumption in equilibrium, in disequilibrium we must focus on the dividend stream, as well as the expected future selling price. Stocks are seldom bought to be held forever. Instead, an individual will have a rough estimate of the time period they will endure before selling the asset. Selling may occur in the expectation that a value diminishing event will negatively effect the stock's price, and hence, provide an incentive for the investor to exit prior to this event obtaining. As a result, the conventional dividend discount model has been reformulated to more concretely account for this future selling price.

39. The discount rate used in dividend discount models is assumed to be based on a flat yield curve. Instead, we see that a trade-off will exist, which results in giving two options that an investor will discount future cash flows with. As an investor will be interested only in that income earned in excess of what they could have earned in originary interest, the two options available are: 1) locking in a rate today for a longer time period, 2) using a series of short-term rates to maximize returns over longer periods by continually rolling over the investments. The option which the investor deems most profitable will provide the applicable interest rate to discount future income streams with.

40. Firm profitability is an integral part of our previously derived asset pricing formula. Profitability is

evident in this formula in three specific components. First, dividends may only be paid out to the extent that firms are profitable. This, in turn, depends on the degree of risk that a risk-mitigating entrepreneur can eliminate. Second, the growth rate in dividends over time will be dictated by the growth in profits. Uncertainty bearing entrepreneurs are those who fundamentally drive the firm into new territories and markets, thus exposing it to greater opportunities for growth. Last, the future selling price of a firm will be dependent on the future markets that the current uncertainty bearing entrepreneurs have driven the firm to, as well as the level of static profitability that the future risk mitigating entrepreneurs can achieve. Hence, this future selling price will be the result of a cumulative learning process within the firm giving it an advantage vis-à-vis its competitors.

41. The expected length of time an asset will be held for significantly influences its present value. Three aspects of this duration are important. First, an individual's unique life-cycle determines to a large extent the duration they may hold an asset for. Older individuals have a comparatively shorter time horizon compared to their younger counterparts. Additionally, it is not the physical life that is important, but the economic life of an individual. As individuals may plan to save beyond their personal lives, and instead save and invest for dependents, the expected holding period is adjusted accordingly. Second, the duration will depend on the specific asset to be purchased. Individuals will hold an asset in order to maximize the expected profit. Depending on firm-specific factors, or general market tendencies, the duration will be lengthened or shortened accordingly. Last, highly volatile asset prices provide a disincentive for an individual to hold them for extended periods of time as the investor is engulfed in a greater degree of felt uncertainty regarding the future selling price. As a result, investors will shorten their expected holding periods to an expected duration consistent with that which they deem appropriate for the level of felt uncertainty.

REFERENCES

Alchian, Armen A., and Harold Demsetz. 1972. Production, Information Costs, and Economic Organization. *The American Economic Review* 62: 777-795.

Alchian, Armen A., and Susan Woodward. 1988. The Firm is Dead; Long Live the Firm. *Journal of Economic Literature* 26: 65-79.

Aoki, M. 1990. The Participatory Generation of Information Rents and the Theory of the Firm. In (*eds*.) M. Aoki, B. Gustaffson, Oliver E. Williamson, *The Firm as a Nexus of Treaties*. Los Angeles, CA: Sage Publications.

Arena, Richard and Sandye Gloria-Palermo. 2008. Menger and Walras on Money: A Comparative View. *History of Political Economy* 40(2): 317-343.

Bagus, Philipp. 2007. Asset Prices – An Austrian Perspective. *Procesos de Mercado: Revista Europea de Economía Política* 4(2): 57-93.

Bagus, Philipp. 2008. Monetary Policy as Bad Medicine: The Volatile Relationship Between Business Cycles and Asset Prices. *The Review of Austrian Economics* 21(4): 283-300.

Bagus, Philipp, and David Howden. 2009a. *Iceland's Banking Crisis: the Meltdown of an Interventionist Financial System*. Ludwig von Mises Institute, Daily Article, June 9th, http://mises.org/story/3499

Bagus, Philipp, and David Howden. 2009b. The Legitimacy of Loan Maturity Mismatching: A Risky, but not Fraudulent, Undertaking. *The Journal of Business Ethics*.

Banerjee, Abhjit V., and Eric S. Maskin. 1996. A Walrasian Theory of Money and Barter. *The Quarterly Journal of Economics* 111(4): 995-1005.

Baumol, William J. 1959. *Business Behavior, Value, and Growth*. New York: Macmillan.

Baumol, William J. 1962. On the Theory of Expansion of the Firm. *American Economic Review* 52(5): 1078-1087.

Berle, Adolph A., and Gardiner C. Means. 1932. *The Modern Corporation and Private Property*. New York: Macmillan.

Black, Fischer. 1970. Banking and Interest Rates in a World Without Money: The Effects of Uncontrolled Banking. *Journal of Banking Research* 1: 8-20.

Bodie, Zvi, Alex Kane, and Alan J. Marcus. 2001. *Essentials of Investments*, *4th ed*. New York: McGraw-Hill Irwin.

Böhm-Bawerk, Eugen von. [1889] 1959. *Capital and Interest. Volume II: Positive Theory of Capital*. South Holland, IL: Libertarian Press.

Böhm-Bawerk, Eugen von. 1901. The Function of Saving. The *ANNALS of American Academy of Political and Social Science* 17: 58-70.

Boudreaux, Donald, and Randall G. Holcombe. 1989. The Coasian and Knightian Theories of the Firm. *Managerial and Decision Economics* 10: 147-154.

Casson, Mark C. 1982. *The Entrepreneur: An Economic Theory*. Oxford: Martin Robertson.

Casson, Mark C. 1997. *Information and Organization: A New Perspective on the Theory of the Firm*. Oxford: Clarendon Press

Cheung, Stephen N. S. 1983. The Contractual Nature of the Firm. *The Journal of Law and Economics* 26: 1-22.

Coase, Ronald H. 1937. The Nature of the Firm. *Economica*, *N.S.* 4(16): 386-405.

Cochran, John P., and Fred R. Glahe. 1999. *The Hayek-Keynes Debate: Lessons for Current Business Cycle Research*. Lewiston, NY: The Edwin Mellen Press.

Cowen, Tyler, and Randall Kroszner. 1987. The Development of New Monetary Economics. *Journal of Political Economy* 95 (3): 567-590.

Cowen, Tyler, and Randall Kroszner. 1994. *Explorations in the New Monetary Economics*. Oxford: Blackwell.

Davidson, Paul. 1993. Austrian and Post Keynesians on Economic Reality: A Response to Critics. *Critical Review* 7(2-3): 38-55.

Davidson, Paul. 2002. *Financial Markets, Money and the Real World*. Cheltenham, UK: Edward Elgar.

Davidson, Paul. 2003-04. Setting the Record Straight on *A History of Post Keynesian Economics*. *Journal of Post Keynesian Economics* 26(2): 245-272.

Demsetz, Harold. 1995. *The Economics of the Business Firm*. Cambridge: Cambridge University Press.

Dowd, Kevin. 1988. Automatic Stabilizing Mechanisms Under Free Banking. *Cato Journal* 7 (3): 643-659.

Fama, Eugene. 1980. Agency Problems and the Theory of the Firm. *Journal of Political Economy* 88: 288-307.

Fama, Eugene. 1983. Financial Intermediation and Price Level Control. *Journal of Monetary Economics* 12: 7-28.

Friedman, Milton. 1953. The Methodology of Positive Economics, in *Essays in Positive Economics*. Chicago: University of Chicago Press.

Foss, Nicolai J. 1996. Capabilities and the Theory of the Firm. *Revue d'Economie Industrielle* 77: 7-28.

Foss, Nicolai J., and Peter G. Klein. 2002. *Entrepreneurship and the Firm: Austrian Perspectives on Economic Organization*. Cheltenham, UK: Edward Elgar.

Foss, Nicolai J., and Peter G. Klein. 2006. The Emergence of the Modern Theory of the Firm. Center for Strategic Management and Globalization, Copenhagen Business School, WP 1/2006.

Garrison, Roger W. 1978. Austrian Macroeconomics: A Diagrammatical Exposition. In (*ed*.) Louis M. Spadaro, *New Directions in Austrian Economics*. Kansas City: Sheed Andrews and McMeel Inc.

Garrison, Roger W. 1992. *The Gold Standard: Perspectives in the Austrian School*, (*ed*.) L. H. Rockwell Jr.. Lexington, MA: D. C. Heath.

Garrison, Roger W. 2001. *Time and Money: The Macroeconomics of Capital Structure*. London: Routledge.

Garrouste, Pierre. 2002. Knowledge: a challenge for the Austrian theory of the firm. In (*eds*.) Nicolai J. Foss, and Peter G. Klein, *Entrepreneurship and the Firm: Austrian Perspectives on Economic Organization*. Cheltenham, UK: Edward Elgar.

Gilbert, J. C. 1953. The Demand for Money: The Development of an Economic Concept. *Journal of Political Economy* 61(2), 144-169.

Gordon, Myron J. 1959. Dividends, Earnings and Stock Prices. *Review of Economics and Statistics* 41(2): 99-105.

Gordon, Myron J. 1989. Corporate Finance Under the MM Theorems. *Financial Management* 18(2): 18-28.

Greenfield, R. L., and Leland B. Yeager. 1983. A *Laisser Faire* Approach to Monetary Stability. *Journal of Money, Credit and Banking* 15: 302-315.

Grubel, Herbert G. 1969. The Distribution of Seigniorage from International Liquidity Concerns. In (*eds*.) Robert A. Mundell & Alexander K. Swoboda, *Monetary Problems of the International Economy*. Chicago: The University of Chicago Press.

Hall, Robert. 1983. Optimal Fiduciary Media Systems. *Journal of Monetary Economics* 12: 33-50.

Hall, R., and C. Hitch. 1939. Price Theory and Business Behaviour. *Oxford Economic Papers* 2:12-45.

Hart, Oliver. 1995. *Firms, Contracts, and Financial Structure*. Oxford: Clarendon Press.

Hayek, F. A. 1933. *Monetary Theory and the Trade Cycle*, (*trans*.) N. Kaldor, and H. M. Croome. New York: Sentry Press.

Hayek, F. A. 1934. Carl Menger. *Economica, New Series* 1(4): 393-420.

Hayek, F. A. [1935] 1967. *Prices and Production, 2nd ed*. New York: Augustus M. Kelly.

Hayek, F. A. 1941. *The Pure Theory of Capital*. London: Routledge.

Hayek, F. A. 1976. *The Denationalisation of Money*. London: Institute of Economic Affairs.

Hayek, F. A. [1980] 1984. The Theory of Currency Competition. In (*ed*.) Pascal Salin, *Currency Competition and Monetary Union*. Boston: Martinus Nijhoff Publishers.

Hayek, F. A. 1994. *Hayek on Hayek: An Autobiographical Dialogue*, (*eds*.) Stephen Kresge and Leif Wenar. London: Routledge.

Hazlitt, Henry. 1978. *The Inflation Crisis, and How to Resolve It*. New Rochelle, NY: Arlington House.

Herbener, Jeffrey M. 2002. After the Age of Inflation: Austrian Proposals for Monetary Reform. *The Quarterly Journal of Austrian Economics* 5(4): 5-19.

Holström, Bengt R., and Jean Tirole. 1989. The Theory of the Firm, in (*eds*.) Richard Schmalansee and Robert D. Willig, *Handbook of Industrial Organization, Vol. I*. Amsterdam: North-Holland.

Howden, David. 2008. Stability of Gold and its Selected Consequences: A Comment. *Procesos de Mercado: Revista Europea de Economía Política* 5(1): 159-175.

Huerta de Soto, Jesús. 1992. Socialismo, Cálculo Económico, y Función Empresarial. Madrid: Unión Editorial.

Huerta de Soto, Jesús. 2004. La Teoría de la Eficiencia Dinámica. *Procesos de Mercado: Revista Europea de Economía Política* 1(1): 11-71.

Huerta de Soto, Jesús. 2006. Money, Bank Credit and Economic Cycles. Auburn, AL: Ludwig von Mises Institute.

Huerta de Soto, Jesús. 2008. *The Austrian School: Market Order and Entrepreneurial Creativity.* Cheltenham, UK: Edward Elgar.

Hülsmann, Jörg Guido. 1998. Toward a General Theory of Error Cycles. *The Quarterly Journal of Austrian Economics* 1(4): 1-23.

Hülsmann, Jörg Guido. 2003. Optimal Monetary Policy. *The Quarterly Journal of Austrian Economics* 6(4): 37-60.

Hülsmann, Jörg Guido. 2008. Time Preference and Investment Expenditure. *Procesos de Mercado: Revista Europea de Economía Política* 5(2): 13-33.

Jensen, Michael C., and William Meckling. 1976. The Theory of the Firm: Managerial behavior, Agency Costs and Ownership Structure. *Journal of Financial Economics* 3: 305-360.

Jevons, William Stanley. 1876. *Money and the Mechanism of Exchange*. New York: D. Appleton and Company.

Keynes, John M. 1936. *The General Theory of Employment, Interest, and Money*. New York: Harcourt Brace.

Kirzner, Israel M. 1973. *Competition and Entrepreneurship*. Chicago:University of Chicago Press.

Knight, Frank H. 1921. *Risk, Uncertainty, and Profit*. Boston: Houghton Mifflin.

Lachmann, Ludwig M. [1956] 1978. *Capital and Its Structure*. Kansas City: Sheed Andrews and McMeel.

Langlois, Richard N. 2002. Modularity in technology and organization. In (*eds*.) Nicolai J. Foss, and Peter G. Klein, *Entrepreneurship and the Firm: Austrian Perspectives on Economic Organization*. Cheltenham, UK: Edward Elgar.

Langlois, Richard N. 2005. The Entrepreneurial theory of the Firm and the Theory of the Entrepreneurial Firm. University of Connecticut, Department of Economics Working Paper Series #2005-27R.

Langlois, Richard N., and Metin M. Cosgel. 1993. Frank Knight on Risk, Uncertainty, and the Firm: A New Interpretation. *Economic Inquiry*, 31: 456-465.

Leontief, W. 1947. Postulates: Keynes' General Theory and the Classicists. In (*ed*.) S. Harris, *The New Economics*. New York: Knopf.

Loasby, Brian J. 2002 Explaining Firms. In (*eds*.) Nicolai J. Foss, and Peter G. Klein, *Entrepreneurship and the Firm: Austrian Perspectives on Economic Organization*. Cheltenham, UK: Edward Elgar.

McCabe, Kevin A. 1989. Fiat Money as a Store of Value in an Experimental Market. *Journal of Economic Behavior & Organization* 12(2): 215-231.

Menger, Carl. [1871] 2007. *Principles of Economics*, (*trans*.) J. Dingwall & B. F. Hoselitz. Auburn, AL: Ludwig von Mises Institute.

Menger, Carl. [1883] 1963. *Problems of Economics and Sociology*, (*trans*.) J. Dingwall & B. F. Hoselitz. Urbana, IL: University of Illinois Press.

Menger, Carl. [1892a] 2005. Money as a Measure of Value, (*trans*.) Gilles Campagnolo. *History of Political Economy* 37(2): 245-261.

Menger, Carl. 1892b. On the Origins of Money. *The Economic Journal* 2: 239-255.

Menger, Carl. [1909] 1977. *Money*, (*trans*.) A. H. Zlabinger. As Prepared for the “Carl Menger

Symposium” at Claremont Men's College. May 26-28.

Mises, Ludwig von. [1934] 1971. *The Theory of Money and Credit*, (*trans*.) H. E. Batson. New York: The Foundation for Economic Education.

Mises, Ludwig von. [1949] 1998. *Human Action*. Auburn, AL: Ludwig von Mises Institute.

Mises, Ludwig von. [1957] 1985. *Theory and History*. Auburn, AL: Ludwig von Mises Institute.

Modigliani, Franco, and Merton Miller. 1958. The Cost of Capital, Corporation Finance and the Theory of Investment. *American Economic Review* 48(3): 261-297.

Myrdal, Gunnar. 1939. *Monetary Equilibrium*. London: William Hodge and Co. Ltd.

Nelson, R. R., and S. G. Winter. 1982. *An Evolutionary Theory of Economic Change*. Cambridge, MA: Harvard University Press.

Niehans, Jürg. 1978. *The Theory of Money*. Baltimore: John Hopkins University Press.

O'Driscoll, Gerald P. 1986. Money: Menger's Evolutionary Theory. *History of Political Economy* 4(18): 601-616.

Osborne, Dale K. 1984. Ten Approaches to the Definition of Money. *Federal Reserve Bank of Dallas Economic Review* (March): 1-23.

Patinkin, Don. 1956. *Money, Interest, and Prices*. Evanston, IL: Row, Peterson & Co.

Penrose, Edith. [1959] 1995. *The Theory of the Growth of the Firm, 3rd ed*. Oxford: Oxford University Press.

Richardson, G. B. 1972. The Organization of Industry. *Economic Journal* 82: 883-896.

Ritter, Joseph A. 1995. The Transition from Barter to Fiat Money. *The American Economic Review* 85(1): 134-149.

Ritzmann, Franz. 1999. Money, a Substitute for Confidence? Vaughan to Keynes and Beyond. *American Journal of Economics and Sociology* 58(2): 167-192.

Röpke, Wilhelm. [1937] 1963. *Economics of the Free Society,* (*trans*.) P. M. Boarman. Chicago: Henry Regnery Company.

Ross, Stephen A. 1987. The Interrelations of Finance and Economics: Theoretical Perspectives. *The American Economic Review* 77(2), Papers and Proceedings of the Ninety-Ninth Annual Meeting of the American Economic Association, (May): 29-34.

Rothbard, Murray N. [1962] 2006. *Man, Economy, and State*. Auburn, AL: Ludwig von Mises Institute.

Rothbard, Murray N. [1963] 1975. *America's Great Depression*. Kansas City: Sheed and Ward.

Rothbard, Murray N. 1976. The Austrian Theory of Money. In (*ed*.) E. G. Dolan, *The Foundations of Modern Austrian Economics*. Kansas City: Sheed and Ward Inc.

Rothbard, Murray N. 1978. Austrian Definitions of the Supply of Money, in (ed.) Louis M. Spadaro, *New Directions in Austrian Economics*. Kansas City: Sheed Andrews and McMeel.

Rothbard, Murray N. 1981. *What Has Government Done to Our Money?* Novato, CA: Libertarians Publishers.

Rothbard, Murray N. [1995] 2006. *An Austrian Perspective on the History of Economic Thought, Volume II: Classical Economics*. Auburn, AL: Ludwig von Mises Institute.

Rothbard, Murray N. 1997. *The Logic of Action Two*. Cheltenham, UK: Edward Elgar.

Salerno, Joseph T. 1994. Ludwig von Mises's Monetary Theory in Light of Modern Monetary Thought. *The Review of Austrian Economics* 8(1): 71-115.

Samuelson, Paul A. 1947. *Foundations of Economic Analysis*. Cambridge: Harvard University Press.

Sautet, Frédéric E. 2000. An Entrepreneurial Theory of the Firm. London: Routledge.

Sechrest, Larry J. [1993] 2008. *Free Banking: Theory, History, and a Laissez-Faire Model*. Auburn, AL: Ludwig von Mises Institute.

Shackle, G. L. S. [1949] 1955. Expectations in Economics': Some Critics Answered. Reprinted in *Uncertainty in Economics*. Cambridge: The University Press.

Shackle, G. L. S. 1952. *Expectation in Economics*. Cambridge: The University Press.

Shackle, G. L. S. 1967. *The Years of High Theory: Invention and Tradition in Economic Thought 1926-1939*. Cambridge: Cambridge University Press.

Shostak, Frank. 1997. In Defense of Fundamental Analysis: A Critique of the Efficient Market Hypothesis. *The Review of Austrian Economics* 10(2): 27-45.

Shostak, Frank. 2000. The Mystery of the Money Supply Definition. *The Quarterly Journal of Austrian Economics* 3(4): 69-76.

Simon, Herbert A. 1951. A Formal Theory of the Employment Relationship. *Econometrica* 19: 293-305.

Simon, Herbert A. 1991. Organizations and Markets. *Journal of Economic Perspectives* 5: 25-44.

Summer, Scott. 1990. The Forerunners of 'New Monetary Economics' Proposals to Stabilize the Unit of Account. *Journal of Money, Credit, and Banking* 22(1): 109-118.

Tobin, J. 1958. Liquidity Preference as Behavior Towards Risk. *The Review of Economic Studies* 25(2): 65-86.

Tullock, Gordon. 1957. Paper Money – A Cycle in Cathay. *Economic History Review, Second Series* 9(3): 393-407.

Walras, Leon. [1898] 1992. Etudes d'économie politique appliquée. In *Oevres économiques complétes* 10. Paris: Economica.

White, Lawrence H. [1986] 1989. A Subjectivist Perspective on the Definition and Identification of Money. Reprinted in (*ed*.) Lawrence H. White, *Competition and Currency: Essays on Free Banking and Money*. New York: New York University Press.

Wicksell, Knut. [1935] 1967. *Lectures on Political Economy. Volume II: Money,* (*ed*.) Lionel Robbins. Fairfield, NJ: Augustus M. Kelley.

Wicksteed, Philip H. 1933. *The Common Sense of Political Economy and Selected Papers and Reviews on Economic Theory. Volume I*. London: Routledge & Kegan Paul.

Williamson, Oliver E. 1975. *Markets and Hierar*chies: Analysis and Antitrust Implications. New York: The Free Press.

Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism*. New York: Free Press.

Willson, Matthew C. 2009. Creativity, Probability and Uncertainty. *Journal of Economic Methodology* 16(1): 45-56.

Wu, Chi-Yuen. 1939. *An Outline of International Price Theories*. London: George Routledge and Sons.

Yeager, Leland B. 1983. Stable Money and Free-market Currencies. *Cato Journal* 3: 305-326.

INDEX OF AUTHORS