Value and Actuation

Christopher R. Perrin, FSA

Presented at
Enterprise Risk Management Symposium
Society of Actuaries

Chicago, IL

April 23–26, 2006

Abstract

The term "enterprise" is used here to refer to the firm as a whole. In order for enterprise risk management (ERM) to support the worthwhile goals of insurance firms, an appropriate conception of "value" is needed. Risk will then be seen, properly, in relation to this conception of value rather than as an inherent property. The purpose of this paper is to outline a framework that measures *decision* and to argue that such a measure is the rightful focus of the enterprise. A central motivating theme is that relatively holistic approaches to valuation should supplant the more traditional methods that set value and assess risk in isolation. The paper develops the principle of maximizing expected utility (PMEU) as a normative decision theory from its philosophical roots in order to motivate its proper usage (i.e., to the enterprise's *ultimate good*). It interprets and applies PMEU in a way that surmounts objections that have been made to it in the form of paradoxes. The theory is then developed mathematically by the direct construction of the general utility function using the concept of a decision lattice. The paper then addresses the misconception that the choice of utility function is hopelessly subjective by establishing rational and objective criteria for its selection. The actuarial community, through the establishment of reliable probability models, is seen as the key intermediary in transforming a subjective view of value and risk into one that can be objectified. Finally, the paper indicates how PMEU can be employed in actuarial activity and insurance company management as the principal measure of a firm's value. Risk mitigation efforts can then be judged in accordance to how they would add to that value.

1. Introduction

The present value of distributable earnings or free cash flow in a stochastic projection model is a natural extension of the notion of "embedded value" (evaluated deterministically). But the result of a stochastic projection model (conditioned on a given decision matrix or strategy) is not a singleton (performance measure), but rather an estimate of the probability distribution of future results (perhaps viewed as a present or future accumulated value). One needs to perform a valuation of *probability distributions* to obtain a "risk-adjusted" performance index that is then directly comparable to alternative distributions (that would be conditioned on alternative decision strategies).

Utility theory has been around a long time; since Bernoulli offered it as the explanation of the St. Petersburg Paradox. Utility theory espouses the principle of maximizing expected utility (PMEU). It is generally brought out as a kind of pedagogically useful fiction in order to introduce and describe certain micro-economic concepts. Its central notion, the utility function, is presented in these treatments in a way that is notably vague. Each market participant is said to possess one, and all the utility functions may have a general shape in common and conform to various constraints, but beyond that, an individual's utility function is presented as inscrutably subjective and unobservable. Indeed the virtue of Option Pricing Theory is that it conveniently allows practitioners to produce a useful result about markets without having to know anything about individuals' utility functions. The individual utility functions lurk in the background, but are unneeded, and so they are put away.

Economists view PMEU as a descriptive theory of individual behavior when confronted with uncertainty regarding the future. Some philosophers have additionally viewed it as a theory of how one ought to act so as to be rational and internally consistent. For them it is used to establish an ethical norm. In this respect, it is referred to as a *normative* theory.

We will concentrate here on the normative aspects of this because, as such, it is much more compelling and illuminating for a theory of value. We shall explore some philosophical ideas and then relate them to actuarial endeavors. This will help develop a normative version of PMEU with a perspective on its proper application. In conjunction with actuarial models (particularly stochastic projection models), we hope to show that PMEU will serve as a formal framework (or accounting basis) for insurance company valuation and performance measurement.[1] The advantage to be

[1] The idea of this use for utility theory is explored in Longley-Cook (1998).

obtained from employing this framework (as opposed to existing frameworks) is the ability to capture the value that is created by the financial intermediation of the insurer.

Some have contended that PMEU is invalid (in either the descriptive or normative form). Their attacks have been presented in the form of various paradoxes. We shall discuss these because they shed light on the theory's proper interpretation. Our thesis is that PMEU, properly employed, is precisely the method best suited for carrying out performance measurement. While directly reflecting risk, it keeps risk in its proper perspective by focusing on the firm's ultimate good (what we will refer to as value). Rather than evaluating how to best reduce risk (for risk-reduction's sake alone) and establishing its cost, PMEU focuses on decisions believed to increase value. It is a notion of value that subsumes the concept of risk.

An insurance firm is usually viewed as consisting of several distinct, separable parts: marketing, investments, underwriting, administration, systems and other support, etc. They seek their separate, specified ends. Corporate management oversees these parts and their interaction. Its role is to integrate these to form a cohesive whole of the parts. Through its decisions, it endeavors to grow the firm's capacity to provide guarantees that form the stable financial foundation of its customers. The actuary is in a unique position to help in the integration. The actuary's role in this context would be to describe, probabilistically, the future outcomes of alternative decisions through the development of models. The actuary also needs to connect to value by calculating the proper performance measurement; where by "proper," we mean one that tracks what is ultimately valuable to the enterprise.

As for models in general, there is a natural tendency for people to want to reduce complicated matters to something simple. Some sort of model (in the broadest sense of the word) is used in all aspects of decision-making. The essential feature of all models is that they *reduce* a complicated and real situation to something more easily comprehended.

Models are a practical necessity and that explains the tendency. However, in the absence of this necessity, models (or reductions) are undesirable. It would be much preferable to comprehend, directly and objectively, the world in all its (perhaps infinite) complexity. This is because, inevitably, any reduction will leave something (possibly important) unrecognized and it is impossible to know *a priori* what that might be. But, as a practical matter, many decisions must be made and they need to be made economically and in time. So, a considerable degree of reduction is inescapable.

This will all be granted as being true enough. Why is there a need to mention it? It is simply that one should always keep the proper perspective in employing models because it is easy to mistake the model for its intended object. Models are not factual. Unlike propositions asserting facts, models are neither true nor false; rather, they are maps or guides for action (or decision) and are subject to revision whenever a sound reason for such revising is discovered. They assert what we are prepared to believe to the point of acting upon them.

We need to bear in mind that a model is related to its object rather than a substitute for it. We only interact with the "outside" world through intermediary signs. A sign stands for its object as a kind of reduction or model of that object. Models, as we ordinarily think of them, describe some aspect of reality. They serve (for the observer or agent) as a sign for that reality and their functionality resides in that they enable the agent to make decisions, to act. But people tend to take reduction too far, even after the need for that degree of reduction becomes technically obviated. A reduction is needed, but as technology improves, it allows us to stop the reduction at a higher level. We should, at some point in the modeling process, check our habit of making unnecessary reductions. Two related areas of which this seems to be true are a) the notion and measurement of (free) capital and b) the performance measurement of a risk-bearing financial intermediary. These are related because the latter ought to track the former.

We seek a methodology for measuring a risky prospect that connects rationally to value. At this point "value" is a vague term. It will become clearer (it is hoped) as we proceed. Among the possible actions we see as alternatives, we want a way of sorting out the best, the one that adds the most value, whatever "value" turns out to mean. The example we have in mind for such a prospect is the future profit stream generated by the closed or (to the extent it is possible) open blocks of business of a life insurance company in total. We focus on the entire enterprise because that is what is ultimately important.

The sense of the word "value" used should not be merely numerical. While we do seek a numerical measure of it for more or less actuarial purposes, "value" should be taken in its more general sense. Ideally, the measure should be responsive to the firm's overall goal. It should connect with objective, rational and worthwhile goals. Action is value seeking. The measure should contour management decisions in such a way that such objectives are most likely to be met in view of the evidence; that is, it should track our most reliable ways of meeting the objective. Rational decisions require that the value measured be internally consistent with the decision-maker's beliefs and successful decisions require, in addition, that beliefs be formed through a reliable process.

2. Historical Background: Existing Performance Measures and Their Shortcomings

Any performance measure is based on an accounting method. So, the accounting method (particularly how it deals with uncertain, future events) is of key importance to us. Generally Accepted Accounting Principles constitute the accounting standard that is usually embraced. The notion of an objective third party countenancing the results imparts a sense of validity. Statutory accounting is believed to be too focused on solvency and therefore, too conservative. Statutory results are used to establish the side-constraints on management so that risk is controlled.

But observing changes to GAAP over the past several years, the manner in which uncertain, future events are treated is quite separable from the rest of the GAAP framework. There seem to be two general (and mutually exclusive) approaches taken within GAAP at present:

1. amortized cost methods that focus on matching costs and revenue incidence, and
2. fair value that has come to mean the use of Option Pricing Theory.

Each identifiable asset and liability is measured under one or the other of these approaches. The second was more recently introduced. The two do not cohere well, however. There seems to be a feeling that the second may ultimately displace the first completely. Proponents for each approach exist, nevertheless. Behind each approach is a philosophy of how the (insurance) world ought to be viewed.

2.1 Amortized Cost Approach

The older, more traditional view is that providing guarantees is a "spread business." In order to be successful, one should take the action that will most likely increase the spread of asset yield over the credited rate on an accumulation product. This embodies a "buy-and-hold" philosophy; once set in place, value and risk are qualities. Defaults and mis-matching cash flows may foil this attempt, so a side constraint (required or dedicated surplus) is established that prevents one from taking too much risk without adequate reward. Required surplus cannot be invested in high-risk assets, only something tame (like short-term, default-risk free bonds). If the required surplus amount properly reflects the riskiness of a particularly risky arrangement, the return on the total investment will suffer relative to alternative arrangements. But the appropriateness of required surplus may be merely a happenstance. Often its level is set by an outside authority and based on incomplete information, and it may lack the refinement needed to respond to differences in the risk levels of particular arrangements. The principles-based method developed so far for risk-based capital is a great accomplishment, and it promises to remedy this situation, but its implementation is being resisted rather strongly at present (i.e., its success will be tempered by any success of the so-called Standard Scenario).

The return on equity (ROE) measure is used to:

1. provide a (common) basis for comparisons across different kinds of business and
2. suggest the best allocation of capital resources.

To the extent that the disability income line of business, for example, has a higher ROE than does the annuity line of business, then corporate decisions will stress growth in disability income product production over that of annuity business. The companion *pricing* method, return on investment (ROI) is viewed as an adjunct to the ROE method. It provides a means of scaling the decision. Set prices for annuities so that their ROI is similar to that of disability income with which annuities compete for capital allocations.

Under the ROE method, the firm is analyzed into profit centers and is overseen by a corporate entity. Assets are divided into three categories: 1) those assets that directly back the liability values; 2) those backing surplus accounts that are dedicated to the liabilities as a kind of cushion (and are proportional to the perceived riskiness of the liabilities); and 3) all other assets which are said to back "free" surplus. Assets of the second and third type constitute the insurers capital base. Each sub-entity has capital assigned to it, and the corporate entity holds whatever excess capital is left over. Return tends to be an earnings figure (as opposed to a general increase in surplus). This is done to isolate "controllable" performance. On a global or company-wide basis, ROE is not much more informative than profit. It is profit divided by equity and since equity is accumulated past profits, this provides no new information. The real objective, it seems, is to relate pricing activity to an overall performance measurement. That is to connect the pricing (performed in a "test-tube" environment) to the effects on the firm's real-world accounting performance. The supposed virtue of this ROI method is its scalability. This feature allows one to focus on the product at hand, on a stand-alone basis. The method is convenient; it uses a minimum of information.

ROE includes free surplus in its denominator and that will encourage the decision-maker to put that free surplus to better use. ROE will encourage the increase in sales to the extent past ROE attributed to a line of business exceeds the ROE observed in the corporate account. It becomes as simple as that. But the extent to which this is true or not true can be measured more reliably and directly, as we will later show. It seems to boil down to pursuing exclusively the strategy believed to produce the highest ROI. Even if consumer demand for the product were high enough to realize that strategy, it would seem to lead to a concentration of risk that detracts from overall value. If consumer demand governs the process toward a more diversified strategy, and if sales are sufficiently large, the result would seem to be a good result for the firm. It seems quite possible however, that using the ROI filter will label superior strategies as non-starters. It builds acceptable strategies but may leave a great many (perhaps the best) out. To prevent this or the concentration of risk situation, an intermediary is needed. Thus ROE and ROI do not measure performance completely.

2.2 Methods Used by Active Traders—Market Value Based Methods

The more recent view is seen by its proponents to better reflect the inherent risk that resides in each asset and liability. The value of the firm is its hypothetical liquidation price. "Fair value" of an insurer is defined as the market value of assets less the market value of liabilities. Market values for most of the assets held by an insurer are generally available and observable. One attractive aspect of market value (as a fair value candidate or a performance measure) is its apparent objectivity. But not all assets (and very few of the insurer's liabilities) have an observable market value. So a very comprehensive theory that is based on a collective coherence principle is used to complete the market, and so unobservable values are at least consistent in a constructive way with the observable. Option pricing models effectively reduce the fixed income securities market to a few factors (most notably the yield curve, default risk category), and reduce the equity markets to these along with some others (like equity markets' volatility rates and their correlation). The liabilities (and so the firm's value) are likewise reducible to market factors. They do this by reducing the market participants' various (subjective) perspectives to that of a reified "market."

Key to the option-pricing methodology is an interest rate model in which all risk-free arbitrage opportunities have been eliminated. This means that no one can, through simultaneous trades, create cash without being exposed to financial risk or obligation in the future. This is a kind of coherence principle for financial markets. (Compare this to the Dutch Book Principle discussed below.) But this applies to cash. No cash can be generated from trades. Cash today represents certain value. It has a factually concrete quality that is lacking (to some degree) in promises to pay. Any worthy and rational valuation framework will treat cold facts like today's cash the same way.

The condition that prohibits the presence of risk-free arbitrage reflects merely a "here-and-now" coherence; it is valid as far as it goes but insufficient for true value assessment in a broader sense. The principle explains why one will pay neither more nor less than the modeled price for a given security provided they possess market knowledge (it gives a rather convincing explanation in terms of a rational process), but not why a given decision-maker should buy a given security (or sell one to another) in the first place. It ignores how a particular asset (or liability) relates to the whole enterprise; it ignores the asset's best use within the context of the firm.

The market is actually a phenomenon, and not a thing. When viewed as a result of reification, is the market coherent over time? One should not expect this if one considers that the market consists of individual decision-makers that contribute to the market's set of (partial) beliefs regarding future events not to mention different attitudes towards risk-taking. The set of participants (even if we assume that they are inter-temporally rational) will change over relatively short intervals of time. New market participants with more optimistic predictions will buy out positions of those with marginally negative beliefs. This transacting will occur until there is no one left willing to buy and the market will settle provided the views of the actual and potential participants have not changed. Therefore, it seems that the market's view (if it can be said to possess a view) will generally be shifting somewhat erratically over time. Any temporal coherence of the market would seem to be coincidental. If you accept the idea that valuation ought to be normatively grounded, then you might think the extension of Option Pricing Theory to a framework for assessing value suffers from the logical fallacy of reification.

Nevertheless, OPT is a marvelous boon. It is a brilliant descriptive theory, conveniently reducing prices to something that can be modeled and used. This makes it extremely valuable. However, within the broader context of value, it takes the reduction farther than is wanted. These very reliable models of market price facts should be incorporated within the decision-making framework for their descriptive power (i.e., we should believe the prices OPT produces in case we need or want to buy or sell in some hypothetical circumstance); but it is going too far to suggest that such models *establish* the framework in which prescriptive choices are judged.

2.3 Unresolved Debate

There is an unresolved debate regarding which of two types of approach to valuation to use. The two sides to the debate are the "buy-and-hold investors" and the "traders." The two sides disagree in a profound way. Could a firm be in one camp at certain times or situations and the other at other times or situations and still be coherent, sometimes a "buy-and-holder" and at others a "trader"? Such a view would be coherent if some coherent process mediates the decision of when to switch from one view to the other. It is doubted that GAAP, in its switching between FAS133 and non-FAS133, constitutes such a process.

OPT has conceptual appeal to those who employ an active trading strategy in making investment decisions. The use of market value as value provides objective results. Option pricing models allow the formation of portfolio benchmarks that "replicate" the liabilities. The benchmark risk-adjusts the total return of the actual asset portfolio. The more traditional basis for performance measurement (amortized cost) is incomplete and inadequate in terms of addressing risk; specifically, surplus as it is measured ignores a change in the riskiness (captured by market value) of a given insurance contract induced by a change in the financial markets.

But those who adopt a "buy-and-hold" strategy favor a measure based on amortized cost. For them, the sometimes wildly fluctuating results of total return seem much too arbitrary and capricious; even meaningless. The ability to wait out financial market disruptions is of considerable value. To the extent one does not envision selling an asset (or at least expects to not sell it with a high probability), the day-to-day market value of that asset is irrelevant. For them the risk-adjusting via benchmark portfolios focuses on the wrong risk.

Both criticisms have merit, but at least one must be mistaken (and we've already pointed out shortcomings in each). Frank Plumpton Ramsey (in a different context) wrote:

> In such cases, it is a heuristic maxim that the truth lies not in one of the two disputed views but in some third possibility which has not yet been thought of, which we can only discover by rejecting something assumed as obvious by both the disputants.[2]

The above bases for measuring value share a couple of properties that are taken for granted. For both, value is linearly combinable. By examining each part of what is being valued, one can merely add up the value of the pieces to obtain the total value. Indeed, for the option-pricing proponents, an insurance contract is merely a collection of options and of no value in and of itself beyond the sum of the objectively discernible option values. For the static measures based on amortized cost, like ROI, the values of the various product lines depend only on features of the liabilities and the assets backing them and are independent of the other product liabilities (and their attending assets) that may be present. In day-to-day decision-making about some part of the company, the view common to both is that one only needs to consider information

[2] See Ramsey's "Universals" (1925).

regarding that part. The valuation can be applied in isolation from the whole and the result of maximizing value locally is the maximization of overall value. These views are overly reductionist. That is their common ground and also their principal flaw.

Another idea common to both is that value is independent of an observer. For the buy-and-hold camp, (amortized) value is a quality or property of a given asset or liability that resides within it. For those who espouse the option pricing view, changes in value appear to be caused by changes in the external factors to which value is reducible. In either case, the observer is unrelated to the object of valuation.

Amortized cost valuation is (logically speaking) monadic; it is representable as a one-place relation. Value is an embedded quality independent of the decision-maker. OPT describes a dyadic (two-place relation) view in which the "market" *causes* value. Missing from both is the information regarding what *use* the assets or liabilities have *for* the observer/decision-maker. In either case, the observer/decision-maker necessarily is present but apparently not on the same level as the objects of valuation. Rather, the observer looks down from some higher meta-level. But as decision-makers, we are "in-the-fray," so to speak, and our actions affect value. We deal as best we can with our limited perspectives. Leaping to a meta-level (above the fray) is not really an option available to us.

Let us express a view that denies the assumptions held in common by the two camps. Value is not an inherent or essential property of the object of valuation. Also, value is not decomposable into natural parts that sum up. Rather, value is a triadic relation among a) the entity (i.e., world) valued, b) an action taken (transforming the world), and c) the evaluator (or observer) who decides to take an action. The observer has its own perspective on the world. The subject of the valuation is a totality, and it contains the observer and his or her perspective. Value subsists only in the presence of the observer, the world, and the attempt to improve the world.

Risk is not a side constraint. It is rather a dimension of value whenever the object of valuation is limited. Risk is not variability, *per se*, but rather variability *from* that which is judged to be desirable by the observer *through valuation*. Risk is not an inherent or essential part of the object of valuation; rather it involves the observer directly and so is a relation as well. The concept of risk is subsumed by that of value.

2.4 Rational Goals

We will hold as a basic tenant that risk-avoidance should be seen as part of the firm's overall objective, especially if the firm is an insurer. A conservative attitude toward risk is fundamental to the insurance business. The ability to allay others' risk is the real product or service that an insurer (*qua* insurer) provides. The risk individuals face that insurers mitigate is generally that their current savings will not persist in the future due to (for example) capital loss or inflation. To the extent that consumers believe that insurers do not value the stability of growth or its ability to pay claims, then the products and services offered by insurers will surely be of diminished value to them; and so, consumers will tend to avoid those products and services. The principal role of the insurer's product is to provide a long-term, stable foundation for wealth accumulation. Having built a foundation, customers can then prudently take more risk in their other investments. The insurer's main reason for operating is to provide this foundation, but it cannot create such a foundation by merely being large or lucky; it must value and demonstrate risk avoidance.[3]

To successfully fulfill its role, an insurer must track in some way the values of the policyholders where their interests are mutual. Consequently, one risk (the most extreme) an insurer faces is the possibility that guarantees made to the policyholder are not met; i.e., solvency risk. The avoidance of solvency risk should be valued explicitly by an insurer (qua insurer). More accurately, it should be seen as a vitally important aspect of value. In general, all risk should be a dimension of value and potentially fatal outcomes should be de-valued to the degree that such possibilities are avoided at all cost.

Beyond meeting its policyholders' needs, the insurer also seeks to succeed financially. This success will not only ensure that the fundamental requirements of existing policyholders will be met, it will also allow for growth in its business. In that way the insurer can increase its role by insuring more policyholders. But more policyholders will only be desirable if it is believed their business will be profitable and will add to capital or free, distributable surplus. Without defining it technically, (free or stand-by) capital is the insurer's degree of wealth. For the insurer, it corresponds to what we are calling value.

---

[3] It should be noted that our use of the term "insurer" has a fairly narrow meaning. Firms engaged in providing insurance often provide other financial services as well; in that capacity they are not strictly insurers, but rather intermediaries engaged in providing instruments that are risky from the buyers' view. They will often add some insurance component in order to differentiate what is otherwise a pure commodity. So, to the extent it is not being an integral part of their product offering, risk avoidance is not valued by these intermediaries nearly as much as it is by pure insurers. Most firms that are commonly grouped under the classification "insurance companies" are not pure insurers (under this definition), but are partially insurers to varying degrees.

While it is quite difficult (if it is possible at all) to pin down the ultimate goals of individual people, the goals of an insurer are rather simple. This goal is the *permanence* achieved by maximizing the sustainable growth rate of the firm's wealth without sacrificing the basic objective of its policyholders. An ineluctable fact is that this goal is sought in an environment of uncertainty and fallibility. Nevertheless, simple belief in solvency is insufficient. The method used for forming belief needs to be honest and as reliable as possible.

2.5 Employment of Cash Flow Testing

Cash flow testing methods based on realistic stochastic economic scenario sets constitute a big improvement over and generalization of the notion of present value. Uncertainty is directly taken into account, so the choice of a risk spread for purposes of discounting should fall in a lower and narrower range than one used within a deterministic setting. Theoretically, a thorough enough model will be able to use the short-term risk-free rates (or something close to them) within a given scenario. By lending partial belief to each of many possible futures, we can have more confidence in taking action.

The product of the simulation is a probability distribution. To be a usable performance measure, it itself needs to be valued. In dealing with probability distributions, one must then be able to measure them; that is, be able to reduce them to a number for preference-ranking purposes. Many will immediately jump to the expected value of the distribution to realize that reduction. However, another method (PMEU) will be shown to more accurately measure a distribution's proper value.

The major issue regarding the role that is played in the decision-making process of insurance companies by cash flow testing and asset liability management (and risk management, in general) is whether risk be treated as a side-constraint or its avoidance incorporated into the objective of the firm. The former seems to be the dominant view held currently, as can be judged by observing current practice among insurers. The regulatory emphasis given to cash flow testing has been a contributor to the subject's development in recent years. However, this has also led to the view that risk management be seen as a constraint on the actions of life companies. If it is truly a constraint, then it cannot be a dimension of value; and, consequently, some probably view the role of asset and liability management as being a kind of gadfly. Only if the cash flow "tests" are not satisfactorily passed (and affect the accounting measure) will cash flow testing exert much of an influence on decision-making and its influence will be negative, sub-optimal.

It can argued that side constraints should succeed in avoiding solvency risk (or what is from the policyholders' view capital risk) if the constraints are strict enough. But, this does not address the desirability of policy performance above the bare minimum guarantee level. There will remain the risk that the minimum guarantee will not keep pace with long run inflation and that the policy performance will not keep up with that of the insurer's competitors. Besides, the side constraints are unnecessary; if managers properly value their firms' performance, they will naturally avoid solvency risk while tracking all the other interests of their firm, including those they hold mutually with their policyholders. This will be made clear later on.

The objective of the firm was mentioned above: maximize the growth rate of capital or wealth without sacrificing the basic objective of the policyholder. This sounds like a goal with a side constraint. But, as we will show below, to the extent we can reduce uncertainty to probability, this goal is not only achievable, it also does not involve the apparent side constraint. That is to say, the insurer's goal of growth is quite compatible with the policyholders' goal of stability. We advocate the development and use of performance measures and pricing models that reflect overall value by directly incorporating as many of its dimensions as we can. Risk is one important dimension of value; ethical behavior is another. The firm is best thought of as an irreducible whole. It will be shown that such a holistic approach is needed to employ utility theory properly and so employed, the principle of maximizing expected utility provides a vast improvement in performance measurement.

Cash flow testing technology and actuarial practice are critical in providing the proper perspective to the decision-maker. Their principal role is to be a reliable tool for the formation of belief regarding future events. It is through building reliable models of behavior in the context of changing economic environments and models assessing the probabilities of those environments' occurrence that provides our best chance of assessing the impact of each of our alternative action choices and selecting those that will lead to ultimate success.

The proper relationship between decision-making and overall value is one of tracking. What is wanted, then, is to value decision. Given the preponderant use of risk management techniques and measures as mere side-constraints, this is an apparently provocative proposal. It is a very common-sense approach, nevertheless, that has been in the (philosophical) literature for a long time. The development of asset and liability management technology (made possible by advances in computer technology and OPT) will play a central role. It is what enables us to better instantiate the PMEU as a valuation method.

3. Philosophical Basis for a Normative Theory

The ideas developed in a set of papers written by Frank Plumpton Ramsey during the 1920s speak directly to issues related to the nature and purpose of actuarial practice. In particular, "Truth and Probability" proposes a framework for making rational decisions in a general setting where belief is (of necessity) distributed among various mutually exclusive propositions regarding future outcomes; that is, under the condition of uncertainty or risk. Better than any other single work, it forms the foundation of the subjective interpretation of probability theory. Thus, "Truth and Probability" (1926) is an extremely important work for actuaries. It has not received the attention it deserves within the actuarial community, perhaps because of its apparent emphasis on the subjective aspects of probability theory (and this may be anathema to many actuaries). Further, Ramsey tends to be lumped in with more extreme subjectivists like di Finetti for whom objective probability did not exist. Ramsey allowed for an objective view that was approachable subjectively so long as reliable methods were in operation.[4]

Robert Nozick has also written quite extensively and insightfully on the subjects relevant to our discussion. Two of Nozick's books (which are very highly recommended), *Philosophical Explanations* and *The Nature of Rationality* are particularly enlightening. The main ideas about decision tracking value expressed herein are due principally to those two works.

But even more fundamental than these are rather abstract ideas due to Charles Sanders Peirce. He is generally thought of as the first pragmatist, but the term "pragmatism" has lately taken on a different meaning for some people; one that associates it with positions that deny the independent existence of an objective point of view (one that is real rather than culturally imagined). Peirce was definitely not of this latter day school; he was an objective realist. Although his "pragmatic" works have been well known and influential since their publication in the 1870s, Peirce was a philosopher of such astoundingly original insight that it seems as though the world is just recently catching up.

3.1 Belief as the Basis for Action

Ramsey was greatly influenced by Peirce's "belief-doubt" doctrine. Two early papers by Peirce that are of particular importance for our present discussion, "The Fixation of Belief" and "How to Make Our Ideas Clear," set forth his basic tenants regarding belief and action. These two papers are primarily responsible for his being

---

[4] For an excellent and very lucid exegesis of Ramsey's work see Sahlin's *The Philosophy of F. P. Ramsey*.

deemed (by William James) the founder of pragmatism, and they were a direct influence on Ramsey in "Truth and Probability." Originally two parts of a series published in *Popular Science*, these papers are his best known and accessible works.

Beliefs are the very basis of whatever action the believer is prepared to take. A rational person will not take an action or make a decision that is in conflict with his beliefs. For example, one would never bet against the sun rising tomorrow. In this example, the decision-maker has (practically) full belief in the proposition that the sun will rise. He or she would likely bet that the sun would rise even at very short odds. Another example of more interest is the set of assumptions used to price life insurance products. These assumptions embody the beliefs concerning the relevant determinants of profit performance.

According to Peirce, doubt is an irritant. The purpose of fixing belief is to relieve the irritation of doubt. Peirce uses the term "inquiry" (even though he remarks that it is sometimes not very apt) to describe the struggle whereby the irritation of doubt is relieved and a state of belief is reached. He discusses four methods of fixing belief: 1) the method of tenacity, 2) the method of authority, 3) the *a priori* method, and 4) scientific inquiry. The first of these methods is to simply believe anything we want. Anyone using this method will be confronted with the fact that other people disagree with his belief and this will cause doubt (and the subsequent struggle) to reappear. The method of authority elevates orthodoxy as a virtue and has been too successful in controlling the conduct of people and thereby standing in the way of progress. According to the third method (the *a priori* method) is intellectually more satisfying than the method of authority. One believes what is "agreeable to reason," but not necessarily experience; it amounts to believing what one is inclined to believe.[5]

These three methods ultimately fail. In the *a priori* method and the method of authority, one refers to only one entity outside itself to fix belief; so-called "reason" and the state, respectively. In this sense these two are really of the same species. The method of tenacity refers to nothing outside itself. All three of these methods are closed. The beliefs generated by any one of these methods are completely settled *by the method's own lights*, even if believing them results in unsuccessful action. Once a belief so obtained fails in the sense that it produces an unsuccessful action, there is nowhere to go and nothing to do within that method. There is no provision in any of these for *learning*.

The fourth, the scientific method, connects with the external; and Peirce champions the scientific method. It requires two things outside the observer: a source

[5] The efficient market hypothesis seems to fit into this category. Even PMEU, when taken as a description rather than a prescription, has found acceptance on the grounds of its rationality.

for hypotheses however obtained (as in the other methods) and an external, objective test; a reality that is independent of our minds.

Inquiry proceeds as follows: the vague (doubt) evolves into the specific (belief or action). Truth is ultimately converged upon by this kind of process within the scientific community. Although the scientific method is fallible, short-term, we will all come to believe what is the truth in the long run. Some may interpret this as being Peirce's definition of truth (a pragmatic definition of truth as fiat), but that is not correct. Rather, the statement is best interpreted as a consequence of scientific inquiry. It expresses his faith in the reliability of the scientific approach in tracking an independent, real world.

3.2 The Categories

This idea about the convergence upon the truth is inscribed in Peirce's later evolutionary cosmology. An excellent and accessible source for this is his set of lectures presented in *Reasoning and the Logic of Things*. The world is its past, present and future. The world contains facts accumulated in the past; the past is marked by specificity. The future is continuous where continuity typifies infinite ("abnumerable") possibility; it is vague by nature. The present is the mediator of the other two via inquiry and action.

> Looking upon the course of logic as a whole we see that it proceeds from the question to the answer, from the vague to the definite. And so likewise all the evolution we know of proceeds from the vague to the definite. The indeterminate future becomes the irrevocable past.[6]

This is best understood as an instance of Peirce's categories of being which he referred to as firstness (being that is in-itself and independent of anything else, pure quality, vague, potential, a sign), secondness (that of being relative to or in reaction with something else; in contrast, it is concrete, specific, definite, an object), and thirdness which relates or mediates a first and a second in a triadic relationship (generality, a law, regularity, an intrepretant).

> My view is that there are three modes of being. I hold that we can directly observe them in elements of whatever is at any time before the mind in any way. They are the being of positive qualitative possibility, the being of actual fact, and the being of law that will govern facts in the future.[7]

[6] See *Reasoning and the Logic of Things, The Cambridge Conference Lectures of 1898*.
[7] See "The Principles of Phenomenology"; *Philosophical Writings*, p. 75.

This rather strange sounding idea is central to all of Peirce's writing. It stems from his work in logic and his logic of relatives, in particular. His work (along with that of Boole and Frege) marks the beginning of modern quantified logic. His logic of relatives is actually a first version of quantified logic and is as robust as that developed later and used today. He developed a special graphical notation to form logical expressions that is very interesting but not very compact. Here, we will use the better-known notation to try to explain his view.

Consider an n-place relation $R(x_1,x_2,\ldots,x_n)$. One may be tempted to describe such a relation as follows:

$$(x_1)(x_2)\ldots(x_n)\quad \big(R(x_1,x_2,\ldots,x_n)\quad .\equiv.\quad S_1(x_1)\,\&\,S_2(x_2)\,\&\ldots\&\,S_n(x_n)\ \big).$$

This expresses the idea that the n-place relation R can be reduced to the n 1-place relations S_1 through S_n. Now, it is clear that many n-place relations can be reduced in this manner. Indeed, one is tempted to start with the monadic (i.e., 1-place) relations and, by definition, build all the higher ordered relations up from them using the conjunction, disjunction and negation operators. Call this an "atomist" approach to the subject. All higher ordered relations so constructed are clearly reducible to those 1-place components. But that is not to say *all* n-place relations are reducible to 1-place relations. When they set out to construct all n-place relations, the atomists would miss a great many. To see the general case, we need to introduce the notion of mediation. So consider

$$(x_1)(x_2)\ldots(x_n)\quad \big(R(x_1,x_2,\ldots,x_n)\quad .\equiv.\quad \exists w(S_1(x_1,w)\,\&\,S_2(x_2,w)\,\&\ldots\&\,S_n(x_n,w)\big).$$ Here, R is reducible to some 2-place or dyadic relations (S_1 through S_n). We intend reduction to be nontrivial (and non-question-begging), so we should stipulate that membership in each Sj must not refer back to relation R. Again, many relations will be reducible to dyadic relations, but the role the variable w plays in the above expression is much too strong, in general. Might there not be an n-place (where n>2) relation for which no w exists that does all the work of pair-wise relating to each argument? Even if we write it as

$$(x_1)(x_2)\ldots(x_n)\quad \big(R(x_1,x_2,\ldots,x_n)\quad .\equiv.\quad \exists w(S_1(x_1,x_2,\ldots,x_{n-2},w)\,\&\,S_2(x_{n-1},w)\,\&\,S_3(x_n,w)\big),$$

we reduce the n-place relation to an (n-1)-place relation and two dyadic relations, but still the role w plays is too strong for a completely general description.

Now consider,

$$(x_1)(x_2)\ldots(x_n)\quad \big(R(x_1,x_2,\ldots,x_n)\quad .\equiv.\quad \exists w_1\exists w_2(S_1(x_1,x_2,\ldots,x_{n-2},w_1)\,\&\,S_2(x_{n-1},x_n,w_2)\,\&\,T(w_1,w_2)\big)$$

When we add two mediating variables (w_1 and w_2) there must also be a relation (T) between them. This demonstrates how a general n-place relation can be reduced to lower order relations; an (n-1)-place, a triadic and a dyadic relation. The process can be reapplied to the (n-1)-place relation until all of the component (sub-) relations have at most three arguments.

Notice that once we get down to triadic relations, we cannot, in general, reduce them to relations with at most two arguments. For if we try to reduce a general triadic relation R to monadic and dyadic relations, we must arrive at

$$(x)(y)(z) \quad \big(R(x, y, z) \quad . \equiv . \quad \exists w_1 \exists w_2 \exists w_3 (S_1(x, w_1) \,\&\, S_2(y, w_2) \,\&\, S_3(z, w_3) \,\&\, T(w_1, w_2, w_3)\big)$$

We need three mediating variables (w_1, w_2, and w_3), because if we try to get by with only one or two, one of them will have to play too strong a mediating role and will fail to provide for complete generality. Also required is a mediating relation T to hold the three mediating variables together; of course, it being triadic, we are back to "square one" and so have not accomplished a reduction to simpler relations.

The general relations are expressed diagrammatically in Figure 1. The lines represent the variables or terms of the relation, and the number of unconnected (or open) lines is the number of terms in the relation. By stringing together combinations of these three figures, we can form complex relations with any number of open terms. For example, Figure 2 indicates how a 4-place relation can be constructed from two 3-place relations and one 2-place relation.

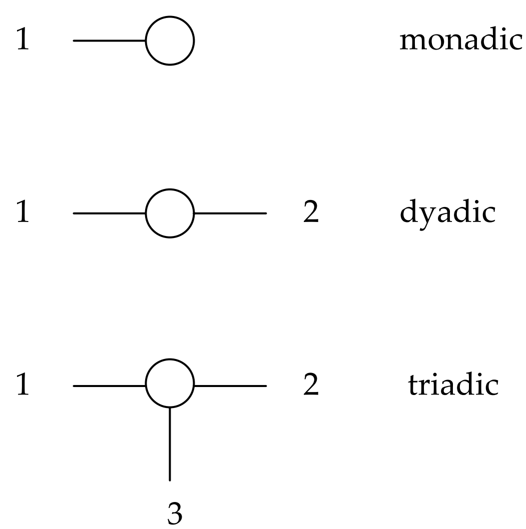
Triadic structure is found at the base of all Peirce's philosophy. The doubt-belief theory mentioned above is an instantiation of it that can be diagrammed as shown if Figure 3. The process of inquiry mediates in time the movement from the vague state of doubt to the specificity of belief. The method of fixing belief through tenacity is characterized by the monadic; it fails when confronted by brute reality (secondness). The method of authority is characterized by the dyadic; it is merely a power relationship between two people or institutions. It ultimately fails in light of the success of a disciplined inquiry; that is, of the scientific method (thirdness).

Another important triad for Peirce was the sign relation (Figure 4). A sign will represent its object only in the presence of an interpretation (or "interpretant"), which gives meaning to the sign. This triad, in turn is capable of subsequently being a sign in a broader context (Figure 5). One can view the sign as a reduction or model of the object; one that is serviceable to the interpretant. The interpretant, in turn is capable of itself being interpreted as a sign whose object is the relationship between the original sign and object. This sets up an infinite (continuous) chain.

A more detailed account of Peirce's categories is well beyond the scope of this discussion; I am not the one to provide it, anyway[8]. The proposed solution to the problem of valuation will have this triadic structure: a decision-maker (a third) takes that action (a second) believed to maximize the value of his or her (future) world (a first). The valuation requires that the world be modeled; the model is a sign of the world (as it might become) that results from the observer's decisions.

Action (or decision) is definite; any contemplated act is either committed or not committed. In many (perhaps most) cases, doubt is not completely eliminated; indeed, there is often a great deal of residual doubt. Full belief in a substantive proposition cannot always be formed, yet *some* action is still required. Indeed, in many situations it would be foolish to wait for enough information that would allow one to come to full belief. Even if time is available, one must decide if the cost of getting the additional information will be worthwhile. The real-world need to act or decide quickly and economically will cause us to cut Peirce's process of inquiry short. So, how do we finally choose? How do we sort out the possibilities and pick the best action under these circumstances?

Figure 1



[8] One may refer to the list of references to Peirce at the end for further reading. Hookway's book is very thorough.

Figure 2

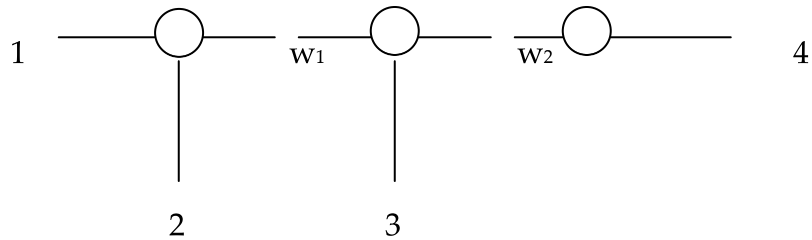


Figure 3

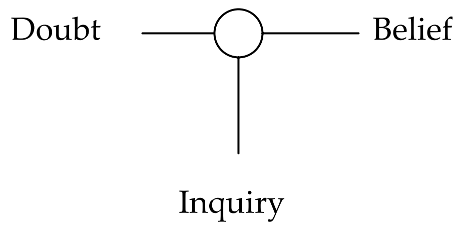


Figure 4

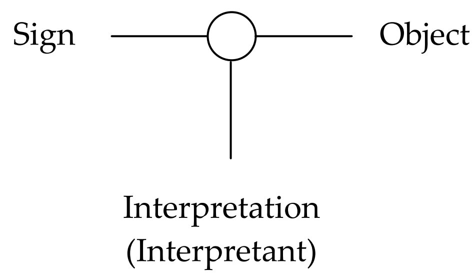
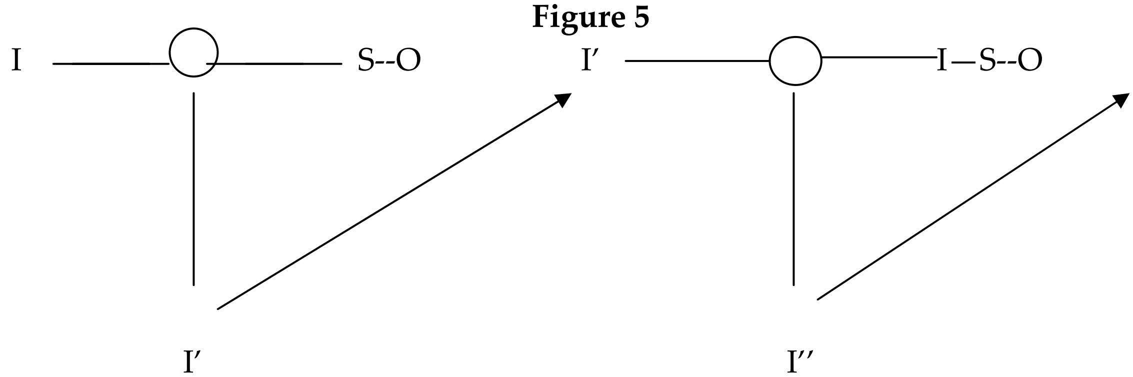


Figure 5



3.3 Subjective Probability and Expected Utility

It being rare that a rational agent will have full belief in a given proposition (particularly one regarding the future), the best one can usually accomplish is a partial belief in a given proposition. For example, consider the proposition that the next flip of a (fair) coin will yield heads. Of course, we do not know that it will be heads but we will lend partial belief to the proposition to the degree of one half. This is not the same as believing that out of 100,000 tosses very nearly 50 percent will be heads. We may well believe the former to the degree one-half *because* we believe the latter to the degree of nearly 100 percent. Here we are just considering one toss of the coin. We would, in other words, be willing to bet that the next coin flip will produce heads as long as we value the prize for winning to be at least as great as twice the value of what is bet. That is because we have a 50 percent belief in the proposition's truth and a 50 percent belief in the truth of its negation. In fact this notion of betting on a proposition's truth is a useful way of measuring the degree of partial belief one has in it, provided one knows the decision-maker's (subjective) valuation of the payoffs of the gamble. Ideally, partial belief is measured in terms of the smallest odds the decision-maker will accept.

Ramsey's "Truth and Probability" examines an interpretation of probability that differs from the more orthodox frequency interpretation. The frequency interpretation treats probability as being an objective (even inherent) property that is scientifically discoverable; it is the basic theory that has been normally taught in school. Ramsey saw the frequency theory as having great value in scientific (descriptive) endeavors. He was not trying to supplant the frequency theory, though his complete theory can be seen as subsuming it.

Ramsey's approach to the subject was based on the above-mentioned betting method. Assuming one can order one's preferences among all possible states of affairs with a cardinal measure, he showed that a probability measure corresponds to the degrees of partial belief of a *coherent* decision-maker. The degree of partial belief one has in a given (true or false) proposition is roughly measured by the minimum odds (stated in terms of preferences) one would accept in betting on its truth. By "coherent," Ramsey means that the subject will not engage in a series of such bets that would result with certainty in a loss:

> If anyone's mental condition violated these laws, his choice would depend on the precise form in which the options were offered him, which would be absurd. He could have a book made against him by a cunning bettor and would then stand to lose in any event. [9]

[9] "Truth and Probability" (1926).

This is known as the Dutch Book Theorem. It has a great deal of support as a principle of rational action. Avoiding such books has much appeal as a principle of rational action. For example, consider a coin toss game: 'tails' will pay the player $2, while 'heads' will pay nothing. Suppose someone believes to the degree 50 percent that the toss will result in 'heads' and to the degree 60 percent that the result will be 'tails.' Based on the first, he would sell up to a certain number (s) of these chances for $1.10 each believing he will make money in the long run. Based on the second, he will also be willing to buy up to a certain number (b) of chances for $1.15 each. Now, let n be the smaller of b and s. He would simultaneously accept an offer to first buy and then sell n changes and lose n*.05 with certainty.

In "Truth and Probability," Ramsey focuses on something his more famous successors seem to have missed or, at least, under-emphasized. He explicitly saw the task of the decision-maker as the appraisal of **"ultimate goods"—i.e., the possible worlds or irreducible wholes.** To illustrate the importance of this, consider the following (extreme) example: You are guaranteed a $1 million payoff five years into the future, provided that you give up (or invest) $1,000 now. Your choices are to accept or reject the offer. Acceptance represents an annual effective rate return of 398 percent, which is considered a favorable rate because you happen to believe that all other alternatives will yield 5 percent per year, at best. Acceptance seems to be the clear choice. But, suppose that you also believe your acceptance would somehow cause an inflation rate of 500 percent annually, and that your rejection of the offer would result in a mere 3 percent annual inflation rate. Clearly, under these circumstances, rejection of the offer seems to be the most rational choice.

Another example that is more down-to-earth is the following: Prospect A is by itself preferred to B by itself, but A is positively correlated with C and B is negatively correlated with C; the conjoined prospect of (A and C) might not be preferred to (B and C). Viewed in isolation, these preferences are not even transitive. But if A, B, and C are *totalities,* then the notion of (A and B) or (A and C) do not make sense (except perhaps in second-order set-theoretic sense).

So, the objects of comparison are to be complete and total descriptions. In general, we must (try to) evaluate the totality of (possible) consequences of a given decision. We must (as best we can) evaluate the possible worlds that we believe would spring from a given decision. We should fuse all the prospects into one big coherent picture. Our choices result from seeking the most valuable whole world. This evaluation is in fact the measure of our wealth. Our primary objective is to increase this measure.

Now, evaluating the whole world is not merely a daunting task, it is too complicated to be considered humanly possible. Still, those forced to make decisions are forced to attempt such a valuation, even if the task is ultimately Sisyphean. Fortunately, most of the decisions we need to make on a day-to-day basis will have (it is believed) only a limited impact. We can make the assumption that everything else is unchanged. We will describe the world in as much detail as we can, but then say "ceteris paribus" with regard to everything else. Normally, the substantial impacts will not be so far reaching into the world and we will be able to take account of them. As for any other more obscure impacts, we will have, after all, very little basis for believing them and will tend to ignore them on pragmatic grounds. The advice here is that when we can incorporate whatever material impacts (we believe) there are, it is important to do so. This is trite, but much decision-making is currently being driven by measures of performance that are grounded in accounting models that fail to reflect this basic wisdom.

Provided one adheres to certain "rationality" and mathematical axioms, Ramsey demonstrated that an agent's preferences (taken as whole worlds) could be measured by a utility function. He also showed that a coherent set of partial beliefs would satisfy the axioms of probability theory. Thus he gave us an interpretation of probability that is subjective. Rational decision is therefore subjective as well, because the rational action is to pursue the prospect that maximizes one's expected utility relative to one's subjective probability measure.

Probability is thus interpretable as the internally coherent system of degrees of belief of a given agent; and this, in turn is a coherent model of the future. When confronted with an unknowable future, the agent distributes his or her beliefs among several (or many) competing propositions according to the agent's degree of belief in those propositions. In this context, probability is seen as being subjective.

3.4 Chance

Although belief is fundamentally subjective, some beliefs are (or should be) preferred because they lead to successful action. Such beliefs tend to be taken for granted because they support successful action. Ramsey describes (in "Chance") these types of beliefs as beliefs imbedded in a theory or system of beliefs. Provisionally accepting a theory (to the point of basing action on it) means accepting certain consequent beliefs. For example, if a pair of dice is thrown. Consider the potential result of two sixes (or "12"). One might partially believe that the result will be "12," and that is subjective. It is considered a true proposition *within the theory* (which has not led people astray, and has, in fact, been useful) that there is a 1/6th *chance* of each face on each die ending in the "up" position (and a zero chance it will land on one of its edges). The theorized and provisionally accepted independence of the dice then leads to the conclusion that the occurrence of "12" will have a 1/36 probability.

> Hence, chances must be defined by degrees of belief; but they do not correspond to anyone's actual degrees of belief; the chances of 1000 heads, and 999 heads followed by a tail, are equal, but everyone expects the former more than the latter...
>
> Chances are degrees of belief within a certain system of beliefs and degrees of belief; not those of any actual person, but in a simplified system to which those of actual people, especially the speaker, in part approximate...
>
> The chances in such a system must not be confounded with frequencies; the chance of φx given ψx might be different from the *known* frequency of ψ's that are φ's. E.g. the chance of a coin falling heads yesterday is 1/2 since 'yesterday' is irrelevant, but the proportion that actually fell heads yesterday might be 1.[10]

The system of beliefs is a theory that is neither true nor false; it is rather a road map for acting.

> A *theory* is a set of propositions which contains [the conjunction] (p and q) whenever it contains p and q, and if it contains any p contains all its logical consequences. The *interest* in such sets comes from the possibility of our adopting one of them as all we believe.
>
> A *probability-theory* is a set of numbers associated with pairs of propositions obeying the calculus of probabilities. The *interest* in such a set comes from the possibility of acting on it consistently.[11]

One could of course take the position that the future is known; that is, one could give 100 percent belief to some proposition about the future, and act accordingly. That is one way of fixing belief, but it is one we are likely to discard as the future unfolds differently. Realistically, we would expect that tenaciously holding to this theory would lead to ultimate failure because it is unreliable. We observe from the past that risk is a reality of which we must take account while making our decisions. And it is not necessarily the case that we get full belief in a probability theory, either. We might

[10] See Ramsey's "Chance" (1928).
[11] See Ramsey's "Probability and Partial Belief" (1929).

tentatively form a working approximation to such probability statements, but can have residual doubts even about that.

3.5 Reliability

It is simply a brute fact of reality that rational decisions need to be based on subjective beliefs that are *fallible*. One cannot know that actions taken in accordance with such beliefs will result in success. There is an evolutionary aspect to this, too. We act on our (system of partial) beliefs regarding certain propositions (we cannot avoid this and be rational), yet if those beliefs do not correspond (closely enough) to reality, the decisions made or the actions taken will tend to fail often enough to endanger the viability of the enterprise. It is not enough to simply have a coherent set of degrees of belief regarding what will transpire in the future. The beliefs have to ultimately track truth. It is important to one's long-term success that the beliefs be formed using as reliable a method as possible.

Ramsey did not endorse a purely subjective account (as did di Finetti and Savage). Indeed, he contended some beliefs are irrational (at least, they become irrational in light of certain evidence). One can believe whatever one wants, but he will still need to deal with the brute facts of reality whenever contrary facts confront the belief (this confrontation is the Peircean notion of secondness). While belief is the basis of action, belief obtained in an unreliable manner would (in the absence of extraordinary luck) tend to result in ultimate failure. In a very short paper (more of a note, actually),[12] Ramsey expresses the idea that for a proposition to be *known* it must be a) true and b) believed; however, c) the method of obtaining belief must be reliable. This is an idealized notion that may never be perfectly instantiated. The power of skepticism is found in its effective undermining of condition c).

Robert Nozick explores an account of knowledge very much along these lines in his *Philosophical Explanations*: a proposition is known if

a. it is true;
b. it is believed;
c. if it were not true, it would not be believed; and
d. if it were true, it would be believed.

The traditional criterion for knowledge had been that it is true, *justified* belief. But that view is undermined by the Gettier Problem.[13] Here is a counter-example. It may be true, for example, that one of the two men in the room owns a Mercedes Benz. I may be

---

[12] See Ramsey's "Knowledge" (1929).
[13] *The Philosophy of F. P. Ramsey*, by Sahlin (1990).

justified in believing that, as well, because I've seen one of the men drive it to work each day for the last year. But, if I am unaware of the fact that the man I saw driving the car over the last year just sold it to the second man in the room, I cannot be said to *know* that the proposition is true. The traditional view that knowledge is true, justified belief has been supplanted by Nozick's view that knowledge encompasses a *tracking* relationship with belief and truth across possible worlds. The manner in which belief is formed is thus crucial for knowledge. The act of fixing belief needs to connect to the fact. The tracking by means of subjunctives (c and d) is an adaptability condition for knowledge that emphasizes the Ramseyian notion of reliability. In the same work, Nozick examines an ethical theory based on the idea that a similar tracking relationship exists between action (or decision) and value.

An extremely important aspect of knowledge's *value* is in the way in which we come to know. Likewise, an important quality of a firm's success is the manner in which it operates. Its decision process should track the value it seeks to enhance. For example, consider the following proposition: "The actuarial profession's success during the 20th century is largely a matter of pure luck. Were it not for the continuous advances in medical fields that extended life expectancy, actuarial science would not have survived, let alone flourished." If this is true, the actuary's work is truly devalued. If our past pricing decisions have been (on the whole) correct (or at least non-disastrous), it would have been for the wrong reasons. What would be missing would be the tracking. While it is impossible to be completely reliable, in an idealized way, the professions stand as tracking intermediaries between belief and truth (as knowledge-seekers). Likewise, decision-makers stand as tracking intermediaries between actions and value.

The employment of reliable methods is the hallmark of a profession. Their use sharpens belief and minimizes the irritation of uncertainty. Within the context of providing a way to form a coherent set of partial beliefs measured by a probability distribution, reliable methods transform an uncertain decision into a risk situation (replacing degrees of partial belief with chances) that is more manageable. This is a kind of reduction, but the reliability of the method justifies it. As we already mentioned, for Peirce, truth is what we would all come to believe in the long run (having employed appropriate methods of obtaining belief, namely the scientific method). This is not his definition of truth as some have suggested; rather it is a result of Nozick's tracking relationship between belief and truth.

Ramsey, here, was not as committed as Peirce. He took a more instrumentalist approach. The standard of reliability for Ramsey was not necessarily truth, but rather, success (success conditions establish and track truth). There is no guarantee that truth will necessarily be converged upon in the long run. One may <u>feel</u> that successful action will correspond to absolute truth well enough (we *believe* that is the case or else we would not bother with inquiry), but, logically, it need not be the case. We need not fall prey to the skeptic's argument, however. The differences between realism and instrumentalism are quite academic to us. The relevant point is that, whether or not what we end up with is truth, we are compelled to make sound decisions and to seek the truth. Even if it turns out that pure truth (and therefore, knowledge) is sometimes unobtainable, we must still make the attempt, perhaps settling for the most coherent road we see to success.

For Ramsey, causal laws (i.e., subjunctive statements of the form "if p then q") and, more generally, theories are neither true nor false. They are what he termed "variable hypotheticals."

> Variable hypotheticals or causal laws form the system with which the speaker meets the future. ... Variable hypotheticals are not judgements but rules for judging 'If I meet a ϕ, I shall regard it as a ψ'. This cannot be *negated* but it can be *disagreed* with by one who does not adopt it.[14]

Peirce, in his later years, espoused as his chief doctrine (Synecism) that the world is characterized by continuity. Now continuity meant for him infinite possibility. This may be all true, but there is no way of us *knowing* that. Whether the world is infinite or finite, we are extremely limited in our ability to perceive it. We must reduce it to a description or a system of beliefs (i.e., chances) that is finite. This is what Ramsey has in mind when he writes

> So too there may be an infinite totality, but what seem to be propositions about it [that are either true or false] are again variable hypotheticals and 'infinite collection' is really nonsense.[15]

Within the limited context of the finite system of (partial) beliefs, such laws (or variable hypotheticals) shadow propositions and we apply logic to the shadow in order to develop further coherent beliefs that we provisionally accept.

As observers of the world, each of our points of view is necessarily subjective. However, that is not to say that there is no part of reality that is objective. Our experience in having to contend with brute reality is good evidence that there are many objective aspects to the world, but there are necessarily subjective aspects as well. Our

[14] See Ramsey's "General Propositions and Causality" (1929).
[15] Ibid.

conception of an objective world needs to include (and leave room for) the subjective aspects of it. Our perspective on the world is essentially subjective; we see it from the "inside." In order to perceive objective reality, we use imagination. We see our own particular view as an instantiation of a more general point of view. That we seem to be successful in objectifying the world is evidence that the world indeed has objective aspects; that truth exists rather than mere meaning. By the same token, objective reality includes all the subjective perspectives. This can be stated in Peircean terms. All our perception of the real world is accomplished through signs or representations. We cannot directly perceive firstness directly, but only through the mediation of signs. There exists an objective (aspect of) reality and we (qua interpreters of signs) stand in a triadic relationship with it along with representations of it. These signs are *reductions* of the world they represent.

Obtaining these reductions by more or less reliable methods bolsters the observer's belief that she faces risk rather than uncertainty. One's degree of partial belief (a subjective thing) so obtained comes closer to being classified as chance (an objective or objectified thing). It is useful, even necessary, to model our system of partial beliefs as a more simple chance situation. This type of substitution is the essence of modeling.

Urn models describe chance situations. In Section 5, we will employ such a model to construct a general risk-averse utility function. The output of a carefully constructed cash flow or simulation model is equivalent in form to an urn model. These results represent the distribution of results that follow from the assumptions made. To the extent that one's decisions affect future results, one can associate with each set of actions a probability distribution. The best action (set) is thus the one that produces the best distribution of results. But how does one judge which distribution is best? The coherent way of sorting this out, of measuring probability distributions, is to take expected utility as the measure of value.

4. General Criticisms of PMEU

Ramsey's "Truth and Probability" not only preceded the work normally credited with the discovery or creation (von Neumann and Morganstern), it also made the explicit point that the theory is valid to the extent that it is applied on a global scale. He intended the theory to a) be seen as "a useful fiction" or approximation and b) be applied to what he called "ultimate goods." The usual treatment of the subject focuses on the mere representation of preferences among a given isolated set of alternative gambles involving sums of money. We want to point out, in particular, problems encountered by these more popular versions of the theory that stem from employing PMEU in too narrow a context. These problems are presented by various paradoxes. These paradoxes are discussed in order to address the criticisms head-on and to better understand PMEU as I think it should be applied. While being criticisms, one should keep in mind that paradoxes (by themselves) could not replace the theory; PMEU continues to be the leading decision theory, at least from a normative perspective. The paradoxes are presented and discussed in order of increasing cleverness.

4.1 Some Paradoxes

4.1.1 A Zeno-like Paradox

The first paradox we will discuss is not particularly well known. It is not particularly threatening, either. It is included to clarify a point. It might be offered as a counter-example to the utility theory used to explain why rational policyholders will buy insurance that costs more than the expected value of the claim (it being well known that that is the only basis on which insurance will be offered). It is reminiscent of Zeno's Paradox in some ways.

> Our subject has $10,000 dollars today. Let us assume that is all she has. There is a 50 percent chance that the $10,000 will be lost in a very short time period—say one day. After this period is over (and if her fortune survives), the risk renews for another period, etc. Suppose her utility curve is the square root function (which is acceptable under the classically known utility theory). She can buy insurance for the period that costs 51 percent of the face amount of coverage or $5,100. Option A is to forgo the insurance. It has an expected utility of $(10000^{.5} * .5 + 0 *.5) = 50$. Option B has an expected utility of $((10000\text{-}5100)^{.5} * 1.00) = 70$. The theory indicates she should buy the insurance; so $5100 goes from her pocketbook to the insurance company.
>
> The next period produces the same basic decision: buy insurance at 51 percent of the remaining $4,900. This will continue until she has (in the limit) paid all of her fortune over to the insurer. Her expected ruin will occur a little sooner, as well. This is hardly an argument for buying the insurance; at least, PMEU doesn't justify it.

The above is an attempt to apply PMEU in a vacuum. Suppose that by spending her $10,000 in cash before the risk had time to eventuate, she could enjoy what was purchased and that would <u>not</u> be subject to the 50 percent chance of loss. In that case PMEU would indicate that she do just that and forgo insurance; a sound decision given the alternatives. On the other hand, suppose anything purchased did not escape the 50 percent chance of total loss during the period, which was followed by more periods with the same risk of complete loss. Then the fallacy stems from (and this is the important point) a confusion regarding the proper conception of wealth. We took for granted that her wealth is measurable as $10,000. The truth of the matter is that she does not really have $10,000. Her wealth, considering the probability distribution of results is zero with 100 percent probability. Her wealth should be measured as zero. Insurance cannot protect nonexistent wealth. That is hardly a criticism of PMEU or insurance.

4.1.2 Allais Paradox

This is a rather famous refutation of PMEU. It tends to show that PMEU fails as a descriptive account of how people actually make decisions under conditions of uncertainty. The following experiment was conducted:

The subject is asked to choose between the following two prospects:

A_1 : 100% chance for $1 million (in 1953 dollars) or

A_2 : 10% chance of $5 million, 89% chance of $1 million, and 1% of $0.

Next the subject is asked to choose between:

B_1 : 11% chance of $1 million, and 89% chance of $0 or

B_2 : 10% chance of $5 million, and 90% chance of $0.

The prizes refer to the incremental addition to the subject's wealth. That is, a zero prize does not mean a complete loss of wealth but rather that there is no increase in it. Most people chose A_1 and B_2. However, the PMEU would require that if A_1 is chosen over A_2, then B_1 should be preferred to B_2.

One possible explanation is that people do not tend to act rationally; and those wanting to apply expected utility as a descriptive theory rather hoped that they would. PMEU is not descriptive; that was established scientifically by Allais' experiment. However, under the normative interpretation, PMEU is not a scientific proposition. The Allais paradox could undercut the normative view if it could be established that the subjects in the above experiment were rational; but that cannot be clearly established.

Even if the subjects could be judged to be (*a priori*) rational, Allais paradox might be explained by a distinction made by Ramsey in "Truth and Probability" between ordinary goods and ultimate goods; in general, money is not an ultimate good, although for business it seems to be an acceptable proxy under all but the most extreme conditions. Nozick's discussion of symbolic utility in his *The Nature of Rationality* sheds light on this as well. People are quite complicated in terms of what they desire; what is important to them cannot always be represented so conveniently in monetary terms.

Suppose the subject is, financially speaking, ordinary. Situation A_1/A_2 represents placing the subject into or very near to a state of financial bliss. Having been offered a choice of A_1, he can be seen as *having* A_1. The decision that he needs to make is whether or not he should give up A_1 for A_2. For most individuals, the idea of risking *that state* for even more money (even if the odds are actuarially favorable) is morally suspect; greedy. In the grand scheme of things, the *meaning* of having $1 million is not much different from having five times that much. On the other hand, a large financial institution might find A_2 more valuable than A_1 particularly if its wealth before acquiring A_1 as a part was worth much more than 1,000,000 in 1953 dollars. It would not likely see A_1 (along with its other forms of wealth) as a state of bliss. If the prizes were much larger but in the same relative proportions, the institution *might* then face a similar situation as does our human subject.

The choice between B_1 and B_2 is quite different. Here the subject starts in a much more modest financial position and is to compare an 11% chance of $1 million to a 10% chance of $5 million. B_1/B_2 as presented does not represent a gambling situation.

The large prize will either materialize or not under situation B_1/B_2. One does not *risk* anything one already has. There is a very significant chance of becoming very rich, however. The trade-off for settling with a chance for $1 million is a 1 percent higher chance of realizing it. If $1 million (in 1953) represents bliss, then that 1 percent higher chance is compelling and taking B_1 is a rational choice. On the other hand, it is natural to want more rather than less. And $5 million is worth more from an expectations perspective. That is not being greedy; it is merely rational. Your viewing someone else as being greedy requires their apparent willingness to risk something that you see as extremely valuable (e.g., bliss, a better world, etc.). Greed is not wanting too much, rather it is risking too much; it is a kind of foolishness. While greed is (for most subjects) a dimension of A_1/A_2, it is not (as significant) a dimension of B_1/B_2.

On the other hand, for a business firm engaged in profit-seeking activities or one that wants to persist in providing financial security to its customers, money (actually, the present value of distributable earnings) is (at least very close to being) an ultimate good. That is, money seems to track what are ultimate goods for such a firm rather closely in those decisions we are likely to face. The PMEU can function well in this kind of business situation as a normative theory (and maybe even as a descriptive one). But, it can break down even at this corporate level in extreme situations. The problem posed by Allais' Paradox comes to light whenever it is invalid to reduce the world to a financial or numerical outcome. As Ramsey would say, PMEU is a "useful fiction." The paradox shows that PMEU is not always useful with respect to individual choice. That is not to say it is never useful there. Nor does it mean that it cannot be dependable (even most dependable) in almost all business decisions faced by financial intermediaries.

Colin Ramsay (1993) argues that the Allais paradox implies expected utility is a useless tool for pricing and so he rejects PMEU. That seems to throw the baby out with the bathwater. He infers too much from the Allais paradox. His real objection is that utility theory provides no objective answer for pricing an insurance risk. We mentioned earlier that part of the world is objective; that one's beliefs get tested against the brute facts of reality. One of those brute facts is that a good deal of what actuaries and decision-makers do is subjective and dependent on highly mediated and fallible sources. That is, they employ more or less Bayesian methods and estimate their degrees of belief (probabilities) of relevant future contingencies as best they can and use feedback to refine those degrees of belief; and never with a completely satisfactory causal, objective explanation of how the future will unfold. Even though an individual may choose differently, an insurance company ought to select B_1 if it would select A_1 and B_2 if it would select A_2.

4.1.3 Prisoner's Dilemma

Much has been written about this. In many ways it is a paradigm for many business situations where two or more competing firms cooperate with each other or firms cooperate with their customers. Mr. A and Ms. B find themselves in jail, separated and about to be indicted for some crime. Each can either confess or not confess. Suppose the following describes the result of each (joint) action and that both A and B are cognizant of them:

		B	
		Don't Confess	Confess
	Don't Confess	(2,2)	(12,1)
A	Confess	(1,12)	(10,10)

The first entry of each couple is the number of years A receives and the second entry is the number of years B receives. Supposing B doesn't confess, then A's best action is to confess (1 year is better than 2 years); and supposing B does confess, then A's best action is again to confess (10 is better than 12). Mr. A reasons that the result of confessing *dominates* the result of not confessing, and so, he *should* confess. The situation from Ms. B's perspective is the same. Her dominant act is to confess as well. The end result is that they both get 10 years. But would not a better result occur from a joint decision to not confess? Then they would each receive only a 2-year sentence.

We can generalize this dilemma as follows:

		B	
		B_1	B_2
A:	A_1	A	b
	A_2	C	d

where a, b, c, and d represent the "state of affairs" produced by the joint action of A and B; and where, for A, $U_A(c) > U_A(a) > U_A(d) > U_A(b)$;

and for B, $U_B(b) > U_B(a) > U_B(d) > U_B(c)$.

Since it is their joint action that determines the outcome, *and the joint action is a part of the outcome,* the possibility and probability that they will cooperate should be evaluated. The outcome "a" is the world that includes the fact that the two parties cooperated with each other and avoided the mutually undesirable world "d." The perceived probability of outcome "a" may be insufficient for Mr. A to take action A_1. The utility of "a" might not be high enough either. But in many cases the utility will be significant and, *therefore,* the probability of the two agents cooperating will be significant as well. This is particularly true if the generalized Prisoner's Dilemma situation is repeatable. It often makes sense to not take full and immediate advantage of an option you have because doing so might hurt opportunities in the future; killing the goose that lays golden eggs. Indeed, it is often the case that the only way for certain goods to exist is with such cooperation. The act of cooperating is often instrumental in *creating* value. Perhaps the Prisoner's Dilemma can be used to differentiate the productive from the merely transactional activities.

Nozick (1993) writes that being a cooperative person can be a value of the individual in and of itself. Living according to that principle has (symbolic) value that should be weighed in the decision process. Likewise, ethical behavior should be seen as more than a means to an end; it is an end-in-itself, to be valued. If you could cause (or make it more likely for) a highly desirable (from your perspective) future possible world to come about by either one of two actions (one unethical the other not), would you really be indifferent to which act you commit? The worlds resulting from each act may have identical futures, but they would not be identical worlds because their *pasts* would differ by the two acts that brought about each of them. Would not one have more value to you than the other, a higher utility? The action or decision that (partially) causes the whole is itself a part of that whole.

Consider a complete (finitist) description of the actions contemplated. If there are n mutually exclusive decisions to make and each has m possibilities, then there are m^n possible worlds to consider. Included in the description of the possible world is the action taken by you and the other individual. Then assess the probability (i.e., your degree of belief) of each world obtaining (on the hypothesis that your action is taken). Assess the utility of each world. Find the expected value of utility for each of your potential actions.

The apparent dominance situation in the Prisoner's Dilemma is seen to be an illusion once we consider wholes.

4.1.4 Newcomb's Problem

This is a very interesting problem and is an excellent paradigm of actuarial practice. Discussions of it can be found in Nozick (1974 and 1993). It has been said that seeing the solution to this problem as being obvious is an indication that one has failed to understand the problem. On a table in a room are two boxes. One is transparent; it plainly contains $1,000. The other box is opaque. Behind the table there is a gentleman.

Before you entered the room, he either put $1,000,000 in the opaque box or he left it empty.

You have one of two choices: 1) take the opaque box or 2) take *both* boxes. You know that the gentleman has predicted your choice beforehand. You also understand that if he predicted that you will select only the opaque box, he placed $1,000,000 in it; and if he predicted your choice will be to select both boxes, he left the box empty. Finally, you have observed the following: of the 100 people that went before you, about half of them selected both boxes and, of these, a very high percentage (nearly all) came out with $1,000. The other half selected only the opaque box and nearly all came out with $1,000,000. What should be your choice?

Based on the evidence, the Expected Utility Hypothesis would indicate taking only the opaque box. Even if a minority of those selecting one box were successful, PMEU will lead you to taking only one box (depending on your utility function, your wealth in the absence of this choice, etc.). Based on logic or common sense (or rather strong opinions that preclude backward causality), this conclusion seems quite wrong. In fact, once the money is or isn't in the opaque box, the choice of both boxes *dominates* the other choice. One will always get $1,000 more.

There have been attempts to refine expected utility theory so that it handles Newcomb's problem.[16] A distinction is made between probabilities that are *causal* rather than *evidential* (or conditional). The probability that your choice *caused* the gentleman to place $1,000,000 in the box is zero. This allows us to use PMEU in deciding to take the dominant action of selecting both boxes.

Nozick's later analysis of Newcomb's Problem (in *The Nature of Rationality*) introduced his notion of "decision value," which gives weights to both these forms of utility (evidential and causal) as well as to a third form he named "*symbolic* utility." Symbolic utility refers to the utility of acting according to principles. Each act of a given kind represents symbolically all such acts. I interpret this approach as a way of viewing the outcomes as wholes (i.e., as Ramsey's ultimate goods). Symbolic utility can be seen as a way of incorporating one's actions as an aspect of the entire outcome; *how* something comes about is a part of its value. Nozick's result can be obtained by treating the weights as the observer's degrees of belief in propositions that report what method of fixing belief is more reliable; such propositions are a part of the long conjunctions of propositions that describe the relevant, possible worlds.

---

[16] See Gibbard and Harper (1978).

The world set in Newcomb's Problem is not one with which I am remotely familiar, unless it is equivalent to a Prisoner's Dilemma situation. In that case, it is handled by expected utility. Otherwise, it portrays a rather bizarre world, perhaps one that is not logically possible. If I should learn that it is the actual world, my beliefs regarding the impossibility of backward causality might be shaken. One possibility is that something earlier caused both the gentleman's prediction and your choice. Such would be a blow to those who believe in free will. Something operating in this world is not well understood, and humility is often extolled as a virtue.

The paradox boils down to what is the best pragmatic approach for fixing one's belief, and not so much one of making a rational decision *in light* of one's beliefs. Rational decision is still acting coherently, given belief. Of course, fixing one's belief is a decision too, and it is that kind of decision that is at the heart of Newcomb's Problem.

Those who pick only the one box do not believe they face a dominance situation. Perhaps they should see it that way, either logically or causally. But, in the world of Newcomb's Problem, the evidence strongly suggests that typical causality is not holding. Is this an impossible world? The evidence *could* be a fluke, but induction is usually very reliable. Some weight should be given to it. Nozick argues that if instead of $1,000 in the clear box, there was only $0.50, then (at least) some proponents of taking both boxes would then switch their answer. This suggests that some weight should be given to the evidential view.

Fixing belief is a prerequisite for a coherent action. The scientific method, for example, has proven itself to be (evidentially) a reliable method (in the actual world, at least). It is not omni-competent, however. But that poses no special problem. The same can be said of deductive logic. Logic is reliable as far as it goes; it simply is not sufficiently edifying in some situations. Statistical inference (or inductive logic) is fallible; however, it may be the most reliable method of fixing belief (in the realm where deductive logic and empirically validated causal models shed no light) available to us. While causal explanations are (inherently) more satisfying (in terms of relieving doubt), contradictory evidence will undermine belief established by causal explanation. Newcomb's Problem is a paradigm of the agent facing the future; that is, the central issue of actuarial practice.

If instead of an experience base of 100 predecessors, suppose we had 1,000,000. That would not change the essence of Newcomb's Problem. It would be clearer, on the other hand, that we are describing a world that is quite different from the one we think we actually experience; one where some methods of fixing belief, heretofore reliable, are no longer serviceable for dealing with the world. In that context, I would choose the one box.

In the bizarre (possible?) world of Newcomb's Problem the two ways of fixing belief overlap but contradict each other. It seems that ignoring dominance is the more adaptable approach. This conclusion need not rest on mere opinion, however; in such a world, it would be a matter of observable (brute) fact. If believing that belief X reliably (enough) leads to an enriching action, while another, belief Y, does not, then believe X. What more rationality (in fixing one's belief) does one need? Well, we say enriching, but enriching in what sense?

Interestingly, Nozick, in a *Scientific American* article,[17] reports that Isaac Asimov responded to Newcomb's Problem. In his response he stated that even though he was a determinist, he would take both boxes because he wanted to "cast a vote" for free will. To translate into utility terms, although he would assign a higher probability to there being a deterministic world, the *value* of worlds that exhibit free will are sufficiently high to maximize expected utility. For Mr. Asimov, the validity of evidentiary methods for solving Newcomb's Problem is not what Ramsey called an "ethically neutral" proposition. Perhaps one's answer to Newcomb's Problem partially depends on the kind of world one desires: one where causal considerations trump evidential considerations or *vice versa*.

4.1.5 Ramsey's Decision Theory in Peircean Terms

The paradoxes were developed to suggest that utility theory does not function well or that it is missing something. While they failed to do that, the paradoxes do suggest that the theory (as normally construed) may not function well in a radically reductive setting. I say "radically," because one cannot perceive the world in its entirety (or immediately for that matter), and so one *must* reduce the world to some model. In Ramsey's pedestrian example in "Truth and Probability," the world is not otherwise affected by the decision to walk, perhaps a bit farther in order to get (possibly) good directions. Thus the required model can be successfully simplified by very liberal use of *ceteris paribus* clauses. The ability to use such clauses helps pick out what goods are (at least serviceably) ultimate.

The paradoxes discussed in the above order begin with solutions that require small extensions to the world model and end up requiring models that entail metaphysics. This modeling process can become quite daunting and is utterly fallible. It is nonetheless ineluctable. A model amounts to the agent's set of beliefs and the manner

[17] See Nozick (1974) and reprinted in Nozick (1997).

in which those beliefs are fixed. With regards to future events, it entails the agent's degrees of partial belief.

Fallibility may be a reason for Ramsey to state that (normative) PMEU is "a useful fiction." To the extent the model fails to reflect aspects of the world that are material to a given decision, it will introduce a tracking error to the measured value of the world. But any such tracking error is due to our inadequacy in describing the world and not PMEU *per se*. Ramsey's point regarding "ultimate goods" in this context is that the more robust the model is, the more useful is PMEU. That is, since one must deal with a sign rather than the world directly, the sign should be sufficiently robust so that the *object* of the valuation is worthwhile. Confusions often arise from mistaking the sign for the object when the sign is somehow inadequate.

The above ideas fit together to form a normative decision-theory that we can extend to a theory of value. In the specific instance of actuarial practice, we can develop performance measures that better track and integrate the various functions of pricing, investing, marketing and planning. By projecting future results in as reliable a way as is available, we reduce the world to a probability distribution (say of the present value of free cash flow that would result from a given course of action). If this reduction of the future course of the world is not too radical, it will serve us. Tying decision alternatives to their (believed) probability distributions allows us to measure the value of the decision (or course of actions) with expected utility.

Value is thus not the result of some dyadic (cause-and-effect) relationship. Rather it is triadic and 1) relates the vague, inchoate possible future to 2) the definite and specific action mediated and taken by 3) the rational, value-seeking agent or decision-maker. We diagram this in Figure 6.

Figure 6 can be expanded into a more detailed diagram; see Figure 7. The model is a (Peircean) sign of the whole world that is the ultimate good. An agent (or decision-maker) stands in a triadic relationship with the world and his or her (probabilistic) description of it. Each of us is a value-seeker and expected utility is our *valuation* of our world (past, present and future included). Our actions (i.e., the present) mediate the progress of the world as it evolves in time. Our aim is to make the world more valuable. In the diagram, the circle denotes a relation and open lines denote the parties to it. In each case, the argument on the upper left is of (Peirce's) first category, the upper right is of the second, and the bottom is of the third. The three relations are connected by the intermediating arguments to form a larger triadic relationship among 1) the world, 2) a specific action taken to appreciate the world and 3) the interpreting decision-maker.

Also, we can now add to the diagrams a temporal aspect. The action taken at time t will in some (probabilistic) way transforms and molds the future possibilities; see Figure 8.

The observers (at the two different times) are further engaged in a triadic relationship. Refining the diagrams to reflect all time the observer is conscious, it will converge toward an infinitely detailed diagram with three open values: 1) the world at the beginning of the process, 2) the world at the end that has been affected by the continuous actions taken, and 3) a continuous observer guided by principle and rationality.

Figure 6

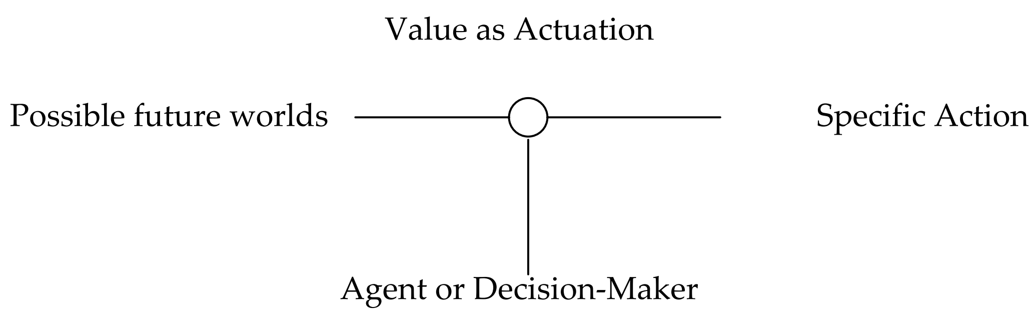


Figure 7

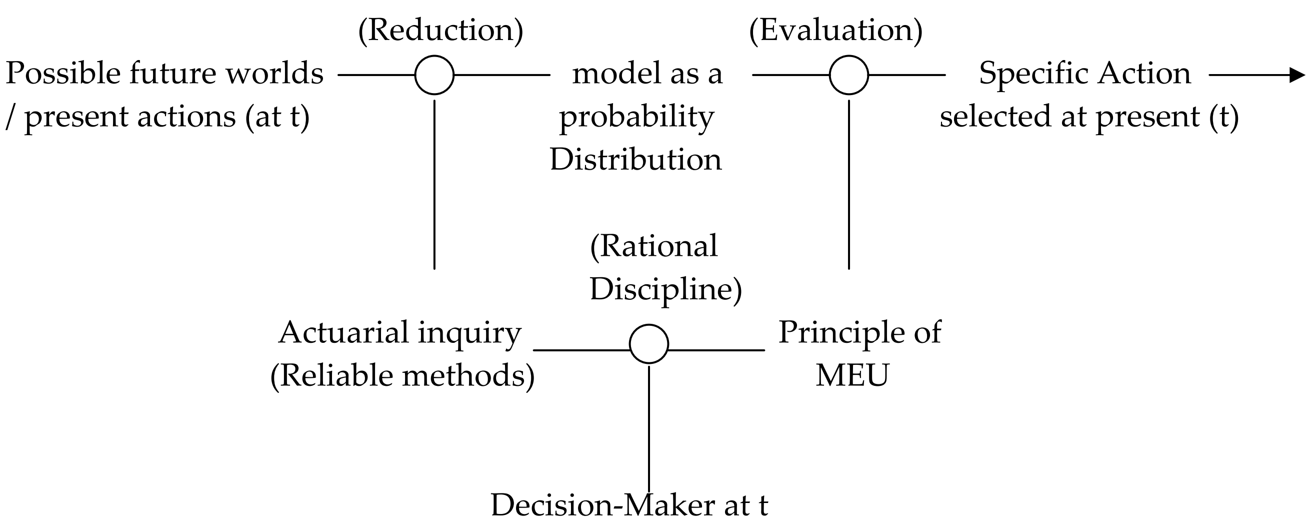
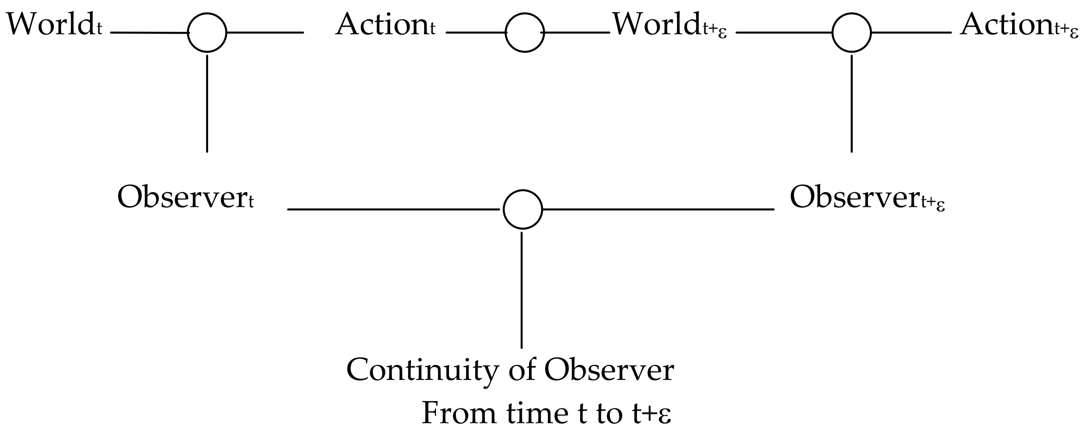


Figure 8



The decision-maker has no choice but to reduce the (vague) world to something comprehensible (i.e., specific). We are warned to avoid, on the other hand, the inaccuracy of descriptions that reduce too much. Is (normative) PMEU too reductive? That question is misplaced. It is an aspect of the valuation rather than the description. Once the utility function is set, the valuation part of the process is rather cut and dried; infallible in a sense. The real, interesting work is in forming the probabilistic description of future.

The composite triadic relation embodies Ramsey's pragmatism. But do not confuse this pragmatism with subjectivism or relativism (both of which can be represented as two-place relations). Although we experience (even objective aspects of) the world subjectively, the third (sub-) relation (rational discipline) connects us to the objective aspects of the world. It is how we can reliably track both the world and its value.

What are the wholes or ultimate goods to which we refer? That varies considerably among individuals. But, for an insurer, its goals are relatively straightforward. For insurers, monetary profit (or more accurately, what is often called the free cash flow stream) is close to being its ultimate good. When that is combined with how the profit was made (i.e., what decision or action influenced its realization), we are closer still. Since money is a quantitative entity, the existence of the appropriate model of preferences seems within grasp, but we are still a long way from specifying that model for a given insurer. How does one actually construct one's utility function? We will use a bootstrapping approach. We hope to develop the details so that the motivation for the method is better appreciated as a normative valuation process.

5. The Internal Structure of the Utility Function

This section develops the notion of utility function for a risk-averse decision-maker in somewhat formal detail. It is intended for those readers for whom the notion of expected utility maximization is unfamiliar and for those who may think that the subject of utility theory is mysterious or even overly artificial. The purpose of this section is to explain the motivation behind the theory, thus showing it to be more natural (or at least, less mysterious). We are specifically concerned with a risk-averse financial institution whose goals are more transparent than those of individuals.

We will build up the theory by means of simple, very hypothetical, risk situations and focus on the decision-making process. Coherence will be an organizing principle here. This process will reveal a microstructure that will suggest how to **construct** the utility function. The mathematical machinery will be kept to the minimum needed (elementary calculus), but it does require a substantial number of technical or theoretical terms and concepts.

The first assumption we make is

> **Axiom 1**: any two prospects (A and B) can be compared and either A is preferred to B, B is preferred to A, or neither; that is, we are indifferent. We can write these as A>B, B>A, or A=B. We say that A is weakly preferred to B if either A is preferred to B or if the decision-maker is indifferent (A>=B).

This is a **completeness axiom**. Theoretically, this may be a rather strong assumption, but we will not get far without it if we have any hope of associating worlds with quantities.

The next two axioms regarding preferences are transitivity axioms:

> **Axiom 2:** if A is preferred to B and B is preferred to C, then A is preferred to C. If A>B, B>C, then A>C.

> **Axiom 3:** if the decision-maker is indifferent to A and B, and is also indifferent to B and C, then the decision-maker is indifferent to A and C. A=B and B=C, then A=C.

Axiom 2 is motivated by the "money-pump" argument. Suppose $A > B$, $B > C$, but $C > A$. It seems reasonable to assume that there exists a C' that is identical to C in all respects except that the subject has ε less money in his pocket. Clearly $C > C'$. It is also reasonable that if ε is small enough, then C' > A. Given that you have C in your possession, you would trade C for B when that trade is offered, then trade B for A. Finally, you would trade A for C'. After that, suppose you are given no more offers to trade. Thus you would end up with ε pumped out of your pocket without anything else to show for it. If one restricts the set of possible world alternatives to those differing only in some monetary outcome, Axioms 2 and 3 are not very strong or controversial

Indifference is an equivalence relation. The realm of possible (and humanly comprehensible) prospects can thus be partitioned into equivalence classes. Two members of the same indifference class will have the same value.

Now, consider the following **urn game**:

> There is an urn containing only red and green balls. Drawings of one ball each are made from the urn and then replaced. At any given time the players know what percentage of the total number of balls in the urn are green, the winning color. To play, one pays an amount of money. If the next ball drawn is green, the player will receive the amount twice the amount paid; that is, the payoff is even-money. If the ball drawn is red (the only other possibility) then the player receives nothing; that is, he or she loses the premium paid to play the game. Assume that there are no other collateral consequences of playing or not playing the game.

The situations depicted here are suggestive of decision-making under conditions of uncertainty. We have simplified matters considerably by using what Ramsey referred to as *chances* in place of probabilities that measure degrees of belief. Again, chances are degrees of belief under a (hypothetically accepted) system of beliefs; i.e., a theory. In this sense, they are objective. If only x percent of the balls are green, then there is little if no controversy that the probability of winning is x percent. In the real life business situations with which we are most concerned here, we will have degrees of partial belief in place of chances. There is likely to be a fair amount of disagreement among people as to what probability function to assign to the set of possibilities in a realistic situation; there might even be disagreement regarding what the possibilities are. How to arrive at the degrees of belief adds a considerable dimension to the problem; one of which much more will ultimately need to be said, later. The endeavor to obtain such degrees of belief, *in a reliable (enough) way,* is the hallmark of the professions.

We are assuming the decision-maker (firm) in question has as its primary objective the increase in wealth (even though we will not have a clear notion of what we mean by the term “wealth” until we are finished. Characterize it, for now, as our valuation of what Ramsey referred to as ultimate goods). It is also assumed that the world will not be different in any way other than the agent's wealth differences that result from playing the urn game. This is an important simplification. We need to evaluate (as best we can approximate) the various outcomes as whole **possible worlds**.

For simplicity, we will restrict our class of possible worlds to those in which the only differences are the various outcomes of the games. That is, no outcome will have any collateral impact on the world, only that the player will be richer or poorer relative to everyone else by her wins or losses. In addition, the worlds that include as parts certain acts (e.g., the unethical ones) are (or ought to be) sufficiently de-valued by us to eliminate such acts from consideration as feasible solutions. Presumably, these constitute all (or at least an interesting range for financial intermediaries) of the equivalence classes.

Let us now return to our simple urn game. If a given player is averse to taking risk, she will be reluctant to bet anything. However, if the percentage of balls that are green (and consequently the chance of winning) were large enough, she would be compelled to play; she would be willing to make a small bet at the very least. Indeed, refusing to bet anything, no matter the chances of winning, seems irrational, *ceteris paribus*. (Of course, she might be in a state of bliss; we shall assume she is not. Also, she may hold as most valuable abiding by the principle of not gambling. In that case, she would avoid making financial decisions under uncertain conditions; and so she isn’t a financial intermediary.)

If we also assume that money or wealth is infinitely divisible, then for any urn game whose probability of winning exceeds one-half, there exists an amount for which a given decision-maker will play. There is a least upper bound for such amount. This is a continuity axiom that we will define more clearly later.

The degree to which you are willing to bet, given the belief of your chance of winning, is a quantitative indication of your attitude toward taking risk. Presuming for now that all of your wealth is in cash, you might be willing to bet as much as 5 percent of it on one drawing if the percentage of green balls in the urn is 60 percent, but in order for you to bet more (say 10 percent) you might require an even better chance of winning (say 70 percent). We postulate that for any individual and any given amount (up to that individual's level of wealth) there is a minimum chance of winning that would compel that individual to bet that amount of money in the urn game; call this the **minimum**

required probability of winning, **p**. Since the probability distribution associated with the urn model is the binomial distribution, we have the corresponding **maximum required probability of losing**, **q** (note that p and q will always add to unity and that q is a least upper bound at or below which, the gamble is taken). Finding someone willing to pay you even-money at your required minimum probability of winning may be impossible, but that is beside the point. The point is that, in order for you to risk a given amount in the urn game, the chance of winning needs to be large enough, otherwise you will prefer not to gamble. We can measure your degree of risk aversion with the notion of the minimum required probability of winning.

If, for all amounts, you require at least a greater than 50 percent chance of winning, then you are, by definition, **risk-averse**. Obviously, when asked to risk a greater amount, your minimum required probability of winning could not decrease. The required chance of winning will generally increase with the size of the bet, *ceteris paribus*. For a given bet, the required chance of winning will generally decrease (or remain the same) as your beginning wealth increases; we refer to this situation as **decreasing (or constant) absolute risk aversion.**

A large bet can be viewed as a series of smaller bets under the following procedure. Play the game repeatedly for a smaller amount until one of two results is obtained: either you cumulatively have lost the larger amount or you have won the larger amount. To illustrate, suppose the larger amount is your entire fortune of 1,000 units. Among the choices you have are to bet 100 percent of your wealth at one time or to bet 1/nth (1,000/n units) in a series of draws that ends only when you have either lost the entire 1,000 units and can play no longer or when your fortune has increased to 2,000 units, and you choose to stop playing. The one large bet is in a sense equivalent to series of small bets with the stopping rule.

5.1 Decision Lattice

Let us rewrite the minimum required probability to reflect the agent's level of wealth. At the same time, we will generalize it a bit more to reflect payoffs that are not necessarily even-money. The expression we will use from hereon is **p(w,a,b)**. This is to be interpreted as the **minimum required probability of winning** an urn trial where the agent's wealth level is w, and the pay-off is win a or lose b. For the rest of this exposition, we will refer to one of these **possible gambles** by the triplet **{w,(a,b,...),(p,q,...)}**. The first argument of the triplet is the level of the agent's wealth, the second is a pay-off vector that specifies all possible results, and the third argument is the vector of probabilities that correspond to each possible result. These probabilities sum to unity, of course. We are defining p(w,a,b) as the smallest value for which the gamble

{w,(a,-b),(p(w,a,b),(1-p(w,a,b))} is acceptable. Any offer to play when the probability of winning exceeds p(w,a,b) will be accepted and any offer when it is less than p(w,a,b) will be rejected. It should be noted that values of w, a, and b are no less than zero, and b is further constrained to be no greater than w. For each w, a, and b, p(w,a,b) will take values between zero and one. To be risk-averse, p(w,a,b) is greater than one half. Our next axiom is:

Axiom 4: for a given decision-maker and for every meaningful combination of w, a, and b, p(w,a,b) exists.

We also have the dual function **q(w,a,b)** which represents the **maximum required probability of losing** the same game. Define

$$c(w,a,b) = q(w,a,b) / p(w,a,b) \qquad [6]$$

Fixing a, the function c(w,a,a) can be used as a measure of **risk tolerance**. Whenever c(w,a,a) is zero, the decision-maker demands a 0 percent chance of losing as a maximum, and therefore has no tolerance for risking the quantity 'a.' Whenever c(w,a,a) equals unity, the decision-maker will be indifferent to a 50 percent chance at even-money odds. This is the least upper bound of the degree of risk tolerance for those who are risk-averse. Any value of c(w,a,a) greater than unity indicates a willingness to gamble under unfavorable conditions (and in a discussion that focuses on normative aspects of decision theory, little more will be said regarding such willful gamblers).

Also, the p function is set so that

$$p(w,0,b) = 1 \qquad \text{for any } b > 0, \qquad [7a]$$

and

$$p(w,a,0) = 0 \qquad \text{for any } a > 0. \qquad [7b]$$

It would not be rational to accept the possibility of losing b without some chance of winning something. Likewise, if there is no possibility of a loss, then there is no reason not to accept the terms of gamble.

Any given a gamble {w,(a,-b),(p,q)} it is possible to view it as an equivalent network of smaller **connected gambles**. Two gambles, $\{w_1,(a_1,b_1),(p_1,q_1)\}$, and

$\{w_2,(a_2,-b_2),(p_2,q_2)\}$, are said to be connected if either

$w_2 = w_1 + a_1$ or $w_2 = w_1 - b_1$. The combined result of the two connected gambles is

$\{w_1,(a_1+a_2,a_1-b_2,-b_1),(p_1*p_2,p_1*q_2,q_1)\}$ if $w_2 = w_1 + a_1$ and

$\{w_1,(a_1,a_2-b_1,-b_1-b_2),(p_1,q_1*p_2,q_1*q_2)\}$ if $w_2 = w_1 - b_1$.

By connecting gambles with n-1 or less possible results or payoffs, we can generate all gambles with n possible payoffs; this is true of any integer, n.

We now state the following (temporal) **Axiom 5** or the **Coherence Constraint on rational choices: If each of two connected gambles would be rejected by a rational agent when presented separately, their combined result will also be rejected.**

This constraint is needed to establish our result. It is an assumption. It appears, to most people, to be quite natural and reasonable. It is a somewhat strong assumption for some in that it requires behaving according to a strict principle. Some may feel the condition is merely an attempt to *reduce all* choice to a simple case. It does that, but it should be seen as fallout from adhering to principle; as *specificity*. Recall (Peirce's trinity and) that value is the relationship in which a decision-maker (a third) mediates between an action or decision (a second) and the world (a first that includes the decision-maker and his or her actions). The decision-maker must (as a practical matter) reduce the description of the world to something comprehensible. We are warned to avoid, on the other hand, the inaccuracy of descriptions that reduce too much. When the Expected Utility Hypothesis is taken as (part of) the description, it is radically reductive. But for normative purposes, (i.e., as a relation) it is not. Indeed, it is a prescription for what one *ought* to do. As such, it is neither accurate nor inaccurate; neither true nor false because it is neither a simple fact nor a (long) conjunction of facts. It is a principle, a map for future action; what Ramsey referred to as a variable hypothetical. One cannot validate or invalidate this principle; one can only agree or disagree with it.

For purposes of pricing and investment decision-making in a quasi-fiduciary role, it provides us with a useful method under all but the most extreme circumstances; that is to say it is serviceable. Also, the Coherence Constraint has normative appeal; if we actually abide by it we will be rational even if it turns out not to be the *only* way. It imposes discipline in a reasonable, direct and understandable way. If a principle that more broadly encompasses the idea of coherence is presented, it will supplant this Coherence Constraint.

Theorem 1: for a given (w,a,b) and for any positive a_1 and b_1 such that

$$a_1 \leq a \quad \text{and} \quad b_1 \leq b \quad \text{we have}$$

$$p(w,a,b) = p(w,a_1,b_1) * p(w+a_1,a-a_1,b+a_1) + q(w,a_1,b_1) * p(w-b_1,a+b_1,b-b_1) \quad [8]$$

Equation [8] holds trivially for $a_1 = a$ or $b_1 = b$. We need only concentrate on $a_1 < a$ and $b_1 < b$.

We know that the right-hand side of [8] exceeds or equals the left-hand side; it follows from the definition of p(w,a,b). If we can also argue that right-hand side cannot exceed the left-hand side, we would establish equation [8]. We can make the argument as follows.

Suppose you are lost in the desert outside Las Vegas. You have $w in your pocket and no other wealth. You find a magic lamp. Upon rubbing the lamp, a genie appears. He is not a traditional genie, however. He either cannot or will not simply grant wishes. Instead, he offers his coherent master sequences of gambles in the urn game varying the payoffs. Being omniscient, the genie offers a gamble the outcome of which is either to win 'a' or lose 'b.'

Suppose the probability of winning is arbitrarily close to but greater than your particular minimum requirement for the chance of winning, i.e., p(w,a,b)+ε. By definition, you would accept these terms. If the genie instead set the chance of winning as p(w,a,b)-ε, then you would prefer not playing.

Now, suppose there is an a_1 and b_1 with $a_1 < a$ and $b_1 < b$ and for which

$$p(w,a,b) < p(w,a_1,b_1)*p(w+a_1,a-a_1,b+a_1) + q(w,a_1,b_1)*p(w-b_1,a+b_1,b-b_1).$$

The genie now proposes the following compound bet:

Step 1 is $\{w,(a_1,-b_1),(prob1,1-prob1)\}$ with probability of winning

$$prob1 = p(w,a_1,b_1) - \varepsilon_1.$$

If the trial is a green ball (a win), then

Step 2 is $\{w+a_1,(a-a_1,-b-b_1),(prob2a,1-prob2a)\}$
with probability of winning

$$prob2a = p(w+a_1,a-a_1,b+b_1) - \varepsilon_2.$$

But, if the first draw is not a green ball (a loss), then
Step 2 is $\{w-b_1,(a+b_1,b_1-b),(prob2b,1-prob2b)\}$ with

$$prob2b = p(w-b_1,a+b_1,b-b_1) - \varepsilon_2.$$

You would not accept any of these sub-gambles, so you would also reject the proposed compound gamble. This is true even if the genie chose ε_1 and ε_2 to be small enough (and he can always find them) so that

$p(w,a,b) < (prob1 * prob2a) + ((1-prob1) * prob2b)$. You would have rejected a gamble you should have accepted according to the definition of p(w,a,b). Because this is a contradiction, the supposition of a strict "less than" relation is false. This completes the proof.

Suppose we restrict, for now, our examination of minimally acceptable gambles. First, we will set a sufficiently small monetary unit. We will consider only those gambles with payoff amounts that are whole integer multiples of the unit measure. A **decision lattice** is the set of all minimally acceptable urn games of the form

$\{w,(a,-b),(p(w,a,b),q(w,a,b)\}$ where w, a, and b are non-negative integers.

We can always refine the lattice to any degree of resolution we desire by dividing the unit in half enough times. By taking this process of refinement to its limit we can arrive at results good for the continuum of non-negative real numbers. Actually we can absolutely stop refining the lattice once the unit measure is down to the smallest unit that concerns us. But the continuous results are handy approximations of what we want. That is why we will "complete the program." So, without loss of generality we can assume the values w, a, and b are integers.

The most obvious interpretation of the decision lattice is that the minimal probability of winning takes as its arguments monetary amounts; numbers. But, 'w', 'a', and 'b' could denote possible worlds (or equivalence classes thereof) where they all share the same past and w "includes" w+a and w-b as yet possible futures. We start with an ordinal enumeration of those possible worlds (or rather the countable number of equivalence classes induced by the indifference ordering) and assume there is a worst (or sufficiently bad) world. We can then associate non-negative integers or rational numbers with them.

Lemma 1: For all integer w >0 we have

$$q(w,1,w) = \alpha(w) \Big/ \sum_{t=0}^{w} \alpha(t) \qquad [9]$$

and consequently,

$$p(w,1,w) = \sum_{t=0}^{w-1} \alpha(t) \Big/ \sum_{t=0}^{w} \alpha(t), \qquad [10]$$

where

$$\alpha(w) = \prod_{t=1}^{w} c(t,1,1)$$

for w>0 and $\alpha(0) = 0$.

Proof:

We will prove the above by the method of induction.

A. For w=1,

$$q(w,1,1) = \frac{c(w,1,1)}{1+c(w,1,1)} = \frac{\alpha(1)}{\alpha(0)+\alpha(1)}$$

B. Now suppose equation [9] holds for w-1. From the Theorem 1, by taking a=1, b=w, a_1=1,b_1=1 and noting that p(w+1,0,w+1)=1, we have

$$p(w,1,w) = \left[p(w,1,1)*1\right] + \left[q(w,1,1)*p(w-1,2,w-1)\right]. \qquad [9a]$$

Also, from Theorem 1,

$$p(w-1,2,w-1) \quad = \quad \left[p(w-1,1,w-1) \quad * \quad p(w,1,w)\right] \qquad [9b]$$

Therefore,

$$p(w,1,w) = p(w,1,1) \quad + \quad \left[q(w,1,1) * p(w-1,1,w-1) * p(w,1,w)\right]. \qquad [9c]$$

Also, because p(w,1,1) > 1/2 > 0, q(w,1,1) <1 and so q(w,1,1)*p(w-1,1,w-1)<1.

This along with [9c] implies that

$$p(w,1,w) = \frac{p(w,1,1)}{\left[1 \quad - \quad (q(w,1,1) * p(w-1,1,w-1))\right]} \qquad [9d]$$

Therefore,

$$q(w,1,w) = \frac{\left[1-(q(w,1,1) * p(w-1,1,w-1)) - p(w,1,1)\right]}{\left[1-(q(w,1,1) * p(w-1,1,w-1))\right]}$$

$$= \frac{\left[q(w,1,1)-(q(w,1,1) * p(w-1,1,w-1))\right]}{\left[1-(q(w,1,1) * p(w-1,1,w-1))\right]}$$

$$= \frac{\left[q(w,1,1) * q(w-1,1,w-1)\right]}{\left[1-(q(w,1,1) * p(w-1,1,w-1))\right]}. \qquad [9e]$$

Now,

$$q(w,1,1) = \frac{c(w,1,1)}{1 \quad + \quad c(w,1,1)} \qquad [9f]$$

and by hypothesis,

$$q(w-1,1,w-1) = \frac{\alpha(w-1)}{\sum_{t=0}^{w-1} \alpha(t)}. \qquad [9g]$$

Substituting these into [9e] we get

$$q(w,1,w) = \frac{\dfrac{c(w,1,1)}{1+c(w,1,1)} \quad * \quad \dfrac{\alpha(w-1)}{\sum_{t=0}^{w-1}\alpha(t)}}{1 \quad - \quad \dfrac{c(w,1,1)}{1+c(w,1,1)} \quad * \quad \dfrac{\sum_{t=0}^{w-2}\alpha(t)}{\sum_{t=0}^{w-1}\alpha(t)}} \qquad [9h]$$

which reduces to

$$q(w,1,w) = \frac{c(w,1,1) \quad * \quad \alpha(w-1)}{\sum_{t=0}^{w-1}\alpha(t) \quad + \quad c(w,1,1) \quad * \quad \alpha(w-1)} = \frac{\alpha(w)}{\sum_{t=0}^{w}\alpha(t)} \qquad [9i]$$

QED

Lemma 2: For all positive integers n we have

$$p(w,n,w) = \frac{\sum_{t=0}^{w-1}\alpha(t)}{\sum_{t=0}^{w+n-1}\alpha(t)} \quad , \qquad [11]$$

and consequently,

$$q(w,n,w) = \frac{\sum_{t=w}^{w+n-1}\alpha(t)}{\sum_{t=0}^{w+n-1}\alpha(t)} \quad \text{for all positive integers w.} \qquad [12]$$

Proof: We will use induction on n.

A. For n=1 we have the result from Lemma 1.

B. Suppose equation [11] holds for m < n

$$p(w,n,w) = p(w,1,w) \quad * \quad p(w+1,n-1,w+1)$$

$$= \frac{\sum_{t=0}^{w-1} \alpha(t)}{\sum_{t=0}^{w} \alpha(t)} * \frac{\sum_{t=0}^{(w+1)-1} \alpha(t)}{\sum_{t=0}^{(w+1)+(n-1)-1} \alpha(t)} = \frac{\sum_{t=0}^{w-1} \alpha(t)}{\sum_{t=0}^{w+n-1} \alpha(t)}$$

QED

Theorem 2: For positive integers w, n, m (with m ≤ w), we have

$$p(w,n,m) = \frac{\sum_{t=w-m}^{w-1} \alpha(t)}{\sum_{t=w-m}^{w+n-1} \alpha(t)} \quad , \qquad [13]$$

and consequently,

$$q(w,n,m) = \frac{\sum_{t=w}^{w+n-1} \alpha(t)}{\sum_{t=w-m}^{w+n-1} \alpha(t)} \qquad [14]$$

Proof: For each m ≤ w, we have

$$p(w,n,w) = p(w,n,m) \quad + \quad (q(w,n,m) * p(w-m,n+m,w-m)) \qquad [13a]$$

Substituting for q(w,n,m) and expanding we get

$$p(w,n,m) = \quad p(w,n,w) - \quad p(w-m,n+m,w-m) + \quad (p(w,n,m) * p(w-m,n+m,w-m))$$

Therefore we have

$$p(w,n,m) = \frac{p(w,n,w) - p(w-m,n+m,w-m)}{q(w-m,n+m,w-m)} \qquad [13b]$$

$$= \frac{\dfrac{\sum_{t=0}^{w-1}\alpha(t)}{\sum_{t=0}^{w+n-1}\alpha(t)} - \dfrac{\sum_{t=0}^{w-m-1}\alpha(t)}{\sum_{t=0}^{w+n-1}\alpha(t)}}{\dfrac{\sum_{t=w-m}^{w+n-1}\alpha(t)}{\sum_{t=0}^{w+n-1}\alpha(t)}} \qquad [13c]$$

which simplifies to

$$p(w,n,m) = \frac{\sum_{t=w-m}^{w-1}\alpha(t)}{\sum_{t=w-m}^{w+n-1}\alpha(t)} \qquad [13d]$$

QED.

We could have fixed our units once and for all in the development above and let the size, ε, of allowable, incremental moves on the decision lattice be a small as desired. The arguments would be identical. We would have defined

$$\alpha_\varepsilon(x) = \prod_{s=1}^{x/\varepsilon} c(s*\varepsilon,\varepsilon,\varepsilon) \equiv \prod_{s=\varepsilon, by\, \varepsilon}^{x} c(s,\varepsilon,\varepsilon) \qquad [15]$$

and the result of Theorem 2 would be

$$p(w,n,m) = \frac{\sum_{t=(w-m)/\varepsilon}^{(w/\varepsilon)-1}\alpha_\varepsilon(t*\varepsilon)}{\sum_{t=(w-m)/\varepsilon}^{(w+n)/\varepsilon-1}\alpha_\varepsilon(t*\varepsilon)} = \frac{\sum_{t=w-m,\varepsilon}^{w-\varepsilon}\alpha_\varepsilon(t)}{\sum_{t=w-m}^{w+n-\varepsilon}\alpha_\varepsilon(t)} \qquad [16]$$

So, we can express any minimally acceptable gamble in terms of a series of arbitrarily small single gambles with stopping rules. The fineness of the lattice can be measured by 1/ε; call it the decision lattice's **degree of resolution.**

From equations [15] and [16] we can specify the decision lattice in terms of c(t,ε,ε). If we specify c(t,ε,ε) for every positive t that is a multiple of ε, we know c(w,a,b) for every w, a and b (where w is greater than or equal to b) and each are whole multiples of ε. We do not know values of c(x,a,b) if a or b < ε. But this will be of no concern if ε is small enough; smaller units will be practically meaningless. On the other hand, it may be computationally convenient to specify lattices with arbitrarily high degrees of resolution; and overly specific lattices will be of no concern either.

We constrain c(t,ε,ε) as follows: For all *rational* ε, and for every t a positive integral multiple of ε.

Condition 1. $0 < c(t,\varepsilon,\varepsilon) < 1$ **Risk Aversion**

Condition 2. $c(t,\varepsilon,\varepsilon) \leq c(s,\varepsilon,\varepsilon) \Leftrightarrow$ t < s **Non-Increasing Absolute Risk Aversion**

By making the inequality in 2 a strict inequality, we could eliminate any c-function that is independent of "wealth" (we would also eliminate step functions).

If we want to stop at the discrete case, these suffice. Our functions will only be defined on the sparse domain of the whole multiples of rational ε however small we want to fix it. In order to extend results to the domain of the real numbers, we will require the following **continuity** conditions, as well.

3.a. $\forall t,\varepsilon > 0, \quad \forall \lambda > 0, \quad \exists \delta \quad s.t. \quad \|x-\varepsilon\| < \delta \rightarrow \|c(t,x,x) - c(t,\varepsilon,\varepsilon)\| < \lambda$

3.b. $\forall t,\varepsilon > 0, \quad \forall \lambda > 0, \quad \exists \delta \quad s.t. \quad \|x-t\| < \delta \rightarrow \|c(x,\varepsilon,\varepsilon) - c(t,\varepsilon,\varepsilon)\| < \lambda$

Lemma 3: For all ε >0 and all t >0 and a multiple of ε ,

$$c(t,\varepsilon,\varepsilon) = c(t-(\varepsilon/2),\varepsilon/2,\varepsilon/2) * c(t,\varepsilon/2,\varepsilon/2) * \frac{(1 + c(t+(\varepsilon/2),\varepsilon/2,\varepsilon/2))}{(1 + c(t-(\varepsilon/2),\varepsilon/2,\varepsilon/2))}$$

Proof: $\dfrac{q(t,\varepsilon,\varepsilon)}{p(t,\varepsilon,\varepsilon)} = c(t,\varepsilon,\varepsilon) = \dfrac{\alpha_{\varepsilon/2}(t) + \alpha_{\varepsilon/2}(t+\varepsilon/2)}{\alpha_{\varepsilon/2}(t-\varepsilon) + \alpha_{\varepsilon/2}(t-\varepsilon/2)}$ from Theorem 2.

$$c(t,\varepsilon,\varepsilon) = \frac{c(t-\varepsilon/2,\varepsilon/2,\varepsilon/2) * c(t,\varepsilon/2,\varepsilon/2) * (1+c(t+\varepsilon/2,\varepsilon/2,\varepsilon/2))}{(1+c(t-\varepsilon/2,\varepsilon/2,\varepsilon/2))}$$

Utility Functions

We define a dual concept for the decision lattice. Define the set of functions, U_ε (relative to a decision lattice with resolution $1/\varepsilon$), to consist of those functions f(w,x) that satisfy the following:

$$(p(w,a,b)*f(w,w+a)) \quad + \quad (q(w,a,b)*f(w,w-b)) \quad = \quad f(w,w) \qquad [17]$$

and

$$f(w,x) \quad \neq \quad b_w \qquad \text{for every x,} \qquad [17a]$$

where b_W is a constant with respect to f's second argument.

The set U_ε consists of the decision-maker's **utility functions.** We have defined the concept of utility function in terms of expected values; the notion of expected value is "programmed" into the concept of utility function.

If f(w,x) is in U_ε, then so is c*f(w,x) + d, because p and q in [17] must add to unity.

For each ε there designate a special member of U_ε,

$$u_\varepsilon(w,x) \quad = \quad \varepsilon\sum_{t=e}^{x-\varepsilon}\alpha_\varepsilon(t) \qquad [18]$$

where e is an arbitrary, non-negative constant selected independently of the first variable, w. If e is greater than w then $u_\varepsilon(w,x) < 0$. Note that changing e merely adds or subtracts a constant from $u_\varepsilon(w,x)$. It is interesting that this function is independent of w, the level of wealth. Therefore, we will from here on drop its first argument and refer to it unambiguously as $u_\varepsilon(x)$.

Lemma 4:

If f(w,x) is a member of U_ε then, f(w,x) is a nontrivial linear transformation of $u_\varepsilon(x)$ in its second argument. That is

$$f(w,x) \quad = \quad a(w)*u_\varepsilon(x) \quad + \quad b(w) \qquad \text{where } a(w) \neq 0. \qquad [19]$$

To see this, note that the property of one function being a linear transformation of another function is an equivalence relation.

1. any function is a linear transformation of itself, and

2. if f is a linear transformation of g then g is a linear transformation of f, and

3. if f is a linear transformation of g, and g is a linear transformation of h, then f is also a linear transformation of h.

Denote the corresponding equivalence class of f(w,x) by [f]. The set U_ε is the union of the mutually exclusive equivalence classes. Now suppose f(w,x) is <u>not</u> a linear transformation of $u_\varepsilon(x)$. Then there are two mutually exclusive equivalence classes, [f] and [u].

From each of these two classes, select a representative, f_1 and u_1, so that $f_1(w,w)=0$ and $u_1(w)=0$. We are assured these representatives exist because <u>all</u> linear transformations of f and u are members of U_ε, and therefore of [f] and [u], respectively. So, we have for every $y < w < x$,

$$p(w, x-w, w-y) * f_1(w,x) \quad + \quad q(w, x-w, w-y) * f_1(w,y) \quad = \quad 0$$

and

$$p(w, x-w, w-y) * u_1(x) \quad + \quad q(w, x-w, w-y) * u_1(y) \quad = \quad 0$$

which is true if and only if

$$\frac{u_1(x)}{u_1(y)} \quad = \quad \frac{f_1(w,x)}{f_1(w,y)}$$ (we know each numerator and denominator are non-zero).

Fixing y (as small as we want) we get $f_1(w,x) = a_W * u_1(x)$ for x>w and this is trivially true for x = w. Similarly, $f_1(w,y) = b_W * u_1(y)$ for y <w. But $a_w = b_w$. So, f_1 and u_1 are and are not from the same equivalence class; a contradiction. Thus the supposition of non-linearity is false. This completes the proof.

Lemma 5: If f and g are (possibly the same) members of U_ε and if , for some p, q=1-p, w, w', x, and y,

$$p*f(w,x)+q*f(w,y)=f(w,w) \quad \text{and}$$

$$p*g(w',x)+q*g(w',y)=g(w',w') \quad \text{then } w=w'.$$

Proof: Suppose w' > w. From Lemma 4, f(w,x) = a(w)*u(x) + b(w) and g(w,x)=c(w)*u(x) + d(w). Therefore

p*u(x) + q*u(y) = u(w). Likewise, p*u(x) + q*u(y) = u(w'). That is, u(w) = u(w').

If w' > w then

$$\varepsilon\sum_{t=w}^{w'-\varepsilon}\alpha(t) = 0.$$ This implies that for s such that w < s < w',

α(s) = 0. This implies that α(t) =0 for all t >w'. It must be that for some positive t less than or equal to w, c(t,ε,ε) =0. But by condition 1, c(t,ε,ε) > 0 if t >0. ⇒⇐
Q. E. D.

Lemma 6: For any rational ε > 0 we have for every t >ε and a multiple of ε,

$$0 < \alpha_{\varepsilon/2}(t) <= \alpha_{\varepsilon}(t) \qquad [20]$$

proof: $\alpha_{\varepsilon/2}(t) =$

$$c(\varepsilon/2,\varepsilon/2,\varepsilon/2)*c(\varepsilon,\varepsilon/2\varepsilon/2) \cdots c(t-(\varepsilon/2),\varepsilon/2,\varepsilon/2)*c(t,\varepsilon/2,\varepsilon/2) \quad . \qquad [20a]$$

Now for each s that is a multiple of ε,

$$c(s,\varepsilon,\varepsilon) = c(s-(\varepsilon/2),\varepsilon/2,\varepsilon/2) * c(s,\varepsilon/2,\varepsilon/2) * \frac{[1 + c(s+(\varepsilon/2),\varepsilon/2,\varepsilon/2)]}{[1 + c(s-(\varepsilon/2),\varepsilon/2,\varepsilon/2)]} \qquad [20b]$$

(this follows from Lemma 3).

$$\text{So, } \alpha_\varepsilon(t) = \alpha_{(\varepsilon/2)}(t) * \frac{[1 + c(t+(\varepsilon/2),\varepsilon/2,\varepsilon/2)]}{[1 + c(\varepsilon/2,\varepsilon/2\varepsilon/2)]} \qquad [20c]$$

From Condition 2, we have $c(t+(\varepsilon/2),\varepsilon/2,\varepsilon/2) \geq c(\varepsilon/2,\varepsilon/2,\varepsilon/2)$

Therefore $\alpha_{\varepsilon}(t) \geq \alpha_{(\varepsilon/2)}(t)$

Lemma 7: For w >e +ε

$$u_{\varepsilon}(w) \geq u_{\varepsilon/n}(w) \geq 0. \qquad [21]$$

Proof: Let $\beta_{\varepsilon}(t) = \frac{1}{2}\left(\alpha_{(\varepsilon/2)}(t) + \alpha_{(\varepsilon/2)}(t+(\varepsilon/2))\right)$

$$= \frac{1}{2}\alpha_{(\varepsilon/2)}(t)\left[1 + c(t,\varepsilon/2,\varepsilon/2)\right]$$

$$\leq \alpha_{\varepsilon/2}(t,\varepsilon/2,\varepsilon/2) \qquad [21a]$$

since $c(t,\varepsilon/2,\varepsilon/2) < 1$ (Condition 1)

Now,

$u_{\varepsilon/2}(w) = \varepsilon * \sum_{t=e,\varepsilon}^{w-\varepsilon}\beta_{\varepsilon}(t)$. Therefore $u_{\varepsilon/2}(w) \leq \varepsilon * \sum_{t=e,\varepsilon}^{w-\varepsilon}\alpha_{\varepsilon/2}(t,\varepsilon/2,\varepsilon/2)$
where the t-terms inside the sigma are incremented by ε.

So, from Lemma 6 we have $u_{\varepsilon/2}(w) \leq \varepsilon * \sum_{t=e,\varepsilon}^{w-\varepsilon}\alpha_{\varepsilon}(t,\varepsilon,\varepsilon) = u_{\varepsilon}(w)$.

Theorem 3: For every rational ε>0, and rational w > 0,

$$u_{\varepsilon}(w) \to c + \int_a^w \alpha(t)dt \equiv u(w) \text{ as } \varepsilon \to 0, \text{ where } \alpha(t) = \lim_{\varepsilon \to 0} \alpha_{\varepsilon}(t) \qquad [22]$$

and ‘a’ and ‘c’ are constants.

Proof: Recall that the definition of $u_{\varepsilon}(w)$ depends on an arbitrary constant, e. Consider the case where w > e+ε, first, then by making the selection of e as small as we want the result will obtain for arbitrarily small w.

So, we can define the following sequences:

$$\varepsilon(k) = \frac{\varepsilon}{n^k}, \quad a_k(t) = \alpha_{\varepsilon(k)}(t) \text{ for } t>\varepsilon, \text{ and}$$

$$I_k(w) = u_{\varepsilon(k)}(w) \text{ for } w > e+\varepsilon.$$

From Lemma 6, $a_k(t)$ is a monotonic decreasing sequence that is bounded below by zero. Therefore, it converges. Call its limit $\alpha(t)$.

Also, from Lemma 7, $I_k(w)$ is a monotonic decreasing sequence that is bounded below by zero. Therefore it converges. Its terms are the Riemann integrals for step functions and so its limit is the Riemann integral and is written as $\int_e^w \alpha(t)dt$. QED

Corollary 1: For any w, a, and b (all positive rational numbers) and decision lattice of resolution 1/ε (where ε is rational and a common divisor of w, a and b), we have

$$p(w,a,-b) = \frac{u(w) - u(w-b)}{u(w+a) - u(w-b)} \text{ and} \qquad [23]$$

$$q(w,a,-b) = \frac{u(w+a) - u(w)}{u(w+a) - u(w-b)}. \qquad [24]$$

Proof: since for every ε(k) in the above proof of Theorem 3 we have

$p(w,a,-b)*u_{\varepsilon(k)}(w+a) + q(w,w-b)*u_{\varepsilon(k)}(w-b) = u_{\varepsilon(k)}(w)$,

$$p(w,a,-b) = \frac{u_{\varepsilon(k)}(w) - u_{\varepsilon(k)}(w-b)}{u_{\varepsilon(k)}(w+a) - u_{\varepsilon(k)}(w-b)}$$ for every k. Therefore, the sequence

$$P_k = \frac{u_{\varepsilon(k)}(w) - u_{\varepsilon(k)}(w-b)}{u_{\varepsilon(k)}(w+a) - u_{\varepsilon(k)}(w-b)}$$ is constant so it converges (trivially) to that constant p(w,a,-b). From Theorem 3, we get the desired result.

Theorem 4:

Given axioms 1 through 4, for an arbitrary, two-outcome gamble {(x,y),(p,q)} where x and y are positive rational numbers (x > y) and p and q are greater than zero and p+q=1, then

there exists a real w, and sequence $\{w_k\}_{k=1}^{\infty}$ and $\{\varepsilon(k)\}_{k=1}^{\infty}$ such that

$\{w_k\}_{k=1}^{\infty}$ converges to w , $\{\varepsilon_k\}_{k=1}^{\infty}$ converges to 0, and

$\{g_k(w_k)\}_{k=1}^{\infty}$ converges to 0 where

$$g_k(w) = u_{\varepsilon(k)}(x) * p + u_{\varepsilon(k)}(y) * q - u_{\varepsilon(k)}(w)$$

and consequently u(w) is defined as the limit (as k increases to infinity) of $u_k(w_k)$

and

$$u(x) * p + u(y) * q = u(w). \qquad [25]$$

Proof: If x=y then w_k=x=w and any sequence of ε that converges to 0 satisfies the claim. So, suppose x > y.

Step 1: Let our step index, n=0. Select an $\varepsilon_0=\varepsilon$ so that x and y are both multiples of ε_0 and note that

$$g_0(y) = p * \left(u_\varepsilon(x) - u_\varepsilon(y)\right) \geq 0 \text{ and}$$

$$g_0(x) = q * \left(u_\varepsilon(y) - u_\varepsilon(x)\right) \leq 0$$

Because $u_\varepsilon(w)$ is a strictly increasing function (i.e. $\alpha_\varepsilon(t) > 0$), there exists w_0 such that $g_0(w_0) >= 0$ and $g_0(w_0+\varepsilon) <= 0$. If either <u>equals</u> 0, we have found our w and related sequences $\{w_k\}$, etc. Otherwise, note that

$$|g_0(w_0)| < g_0(w_0) - g_0(w_0 + \varepsilon_0) = \varepsilon_0 * \alpha(w_0 + \varepsilon_0)$$

$$= \varepsilon_0 \prod_{t=\varepsilon_0}^{w_0+\varepsilon_0} c(t, \varepsilon_0, \varepsilon_0).$$

Step n: Increase step index, n-1, by 1. Let $y_n=w_{n-1}$ play the role of y in step 1 and $w_{n-1}+\varepsilon = x_n$ play the role of x.

Also, let $\varepsilon(n)= \varepsilon(n-1)/m$ for some m>1.

Define g_n as in step n-1 with $\varepsilon(n)$ replacing $\varepsilon(n-1)$. As in Step n-1 find a w_n such that $g_n(w_n) >=0$ but $g_n(w_n+\varepsilon(n)) <=0$. Stop if equality holds.

Otherwise note that

$$|g_n(w_n)| = \varepsilon_n\alpha_n(w_n) \leq \varepsilon_n\alpha_n(y) \leq \varepsilon_n\alpha_0(y) \text{ from Lemma 6.}$$

Therefore, $|g_n(w_n)| \rightarrow 0$ as n increases.

Clearly, ε_k converges to zero while w_k converges to some real number between y and x. Also, $|g_k(w_k)|$ is bounded above by a sequence that converges to zero, so $g_k(w_k)$ converges to zero as well. QED

The continuous version of the utility function, $u(w) = \int_e^w \alpha(t)dt$ has been defined at all positive rational values of w. Also, Theorem 4 establishes that any urn game involving two positive, rational payoffs (to be viewed as the pro-forma results that could occur) has associated with it a unique (see Lemma 5) real value, w' (for wealth). This is a risk-adjusted value of the game. If the game has value only one possible outcome (i.e. x=y), then w is that value (w=x). This value, w, can therefore be seen a "sure-thing" equivalent to the game. For example suppose the game is {(a,-b),(p,q)} and the decision-maker's cash stake is w_0 (and has nothing other than cash to value). If applying the theorem to the gamble {(w_0+a,w_0-b), (p,q) } results in the solution, w_1 of p*u(w_0+a) + q*u(w_0-b) = u(w_1) then if w_1 exceeds w_0, the terms of the game are acceptable; otherwise, our decision-maker will choose not to play.

We have not yet used the continuity axiom. Since the only payoffs of any real world importance are indeed rational, we are, in a way, finished. We can fill in the function, u(w), for its irrational arguments as the limits of sequences {u(w_k)} where w_k are rational and converge to w.

Continuity will ensure the value of u(w) at irrational w is well defined. If $c(t,\varepsilon,\varepsilon)$ is continuous, then so is $\alpha_\varepsilon(t)$, and so it follows that the limit function, $\alpha(t)$ and its integral u(w) are also continuous. We can then adopt values for p(w,a,b) when any of the arguments is irrational as follows:

$$p(w,a,b) \;=\; \frac{u(w) \;-\; u(w-b)}{u(w+a) \;-\; u(w-b)} \qquad [26]$$

This implies

Theorem 5:

For an arbitrary, two-outcome prospect $\{(x,y),(p_1,q_1)\}$ where x and y are positive rational numbers ($x > y$) and p and q are greater than zero and $p_1+q_1=1$, then there exists a unique, real w such that $p_1 = p(w,x-w,w-y)$. In other words, any two-outcome gamble or prospect has a unique one-outcome or sure-thing prospect that is of equivalent value.

Finite Probability Distributions

We can now move on to the general case of finite distributions.

Theorem 6: For every finite distribution with n possible pay-off values,

$0 <= a_1 < a_2 < ... < a_n$ with probabilities p_1 , p_2, ... p_n, there exists unique real
$w_1 <= w_2 <= .....<= w_n$ such that
$p_n = p(w_n,\; a_n-w_n,\; w_n- w_{n-1})$, and, for j=2 to n-1,

$$p_j \;=\; [\prod_{s=j+1}^{n} q(w_s, a_s - w_s, w_s - w_{s-1})] * \quad p(w_j, a_j - w_j, w_j - w_{j-1})$$

and $u(w_n) \;=\; \sum_{j=1}^{n} p_j \;*\; u\left(a_j\right)$.

Proof: We proceed by induction on n.

Throughout, set $w_1 = a_1$.

For n=2 the statement is true by letting $w_1 = a_1$,and applying Theorem 5. Let $x = a_2$ and $y=a_1$ to get $w_2=w$.

$u(w_2) \;=\; \sum_{j=1}^{2} p_j \;*\; u\left(a_j\right)$ by definition of $p(w,a_2-w,w-a_1)$.

Suppose it is true for all distributions with n-1 elements. Consider the n-1 term distribution $\{a_j\}$ j=1 to n-1 with probabilities $p'_i = p_j / q_n$. By supposition, there exists unique $\{w_j\}$ for j=1 to n-1. There also exists a unique w such that $p(w, a_n\text{-}w, w - w_{n-1}) = p_n$. Let $w_n = w$.

$$u(w_n) = p_n * u(a_n) + q_n * u(w_{n-1})$$

$$= p_n * u(a_n) + q_n * \sum_{j=1}^{n-1} [(p_j/q_n) * u(a_j)]$$

QED.

This shows that any finite non-negative distribution (of a complete outcome) is equivalent to some one-outcome distribution or sure-thing, in that the decision-maker is indifferent (that is, would not trade one for the other). Furthermore, the expected utility of the distribution equals the utility of the equivalent sure-thing.

5.1.1 Examples

Suppose we are given $c(t,\varepsilon,\varepsilon)$ at the bottom level of our construction. Consider $\alpha'(t)$.

$$\alpha'(t) = \lim_{\varepsilon \to 0} \left(\frac{\alpha(t+\varepsilon) - \alpha(t)}{\varepsilon} \right)$$

$$= \alpha(t) \lim_{\varepsilon \to 0} \left(\frac{c(t+\varepsilon,\varepsilon,\varepsilon) - 1}{\varepsilon} \right). \qquad [27]$$

as refinement to the lattice occurs.

1. Suppose $c(t,\varepsilon,\varepsilon) = 1-b\varepsilon$, where b is a positive real number small enough so that $c(t,\varepsilon,\varepsilon) > 0$. In this example, the c-function is independent of t; a rather peculiar state of affairs. This violates the variation on Condition 2 that requires decreasing absolute risk aversion (Condition 2 allows for constant absolute risk aversion). Substituting this into [27] yields $\alpha'(t) = -b*\alpha(t)$. This implies that

   $\alpha(t) = e^{-bt}$. Therefore,

$$u(w) \;=\; \int_1^w e^{-bt}\;dt \;=\; \frac{1 \;-\; e^{-bw}}{b}$$

or some (equivalent) linear transformation thereof. This has been a popular choice for a utility function because it is not only simple in form, but also it has the property that the sure-thing equivalent (or value) of a normal distribution with mean, μ, and variance, σ^2, is simply μ - .5*b*σ^2. For this reason, it is referred to as the "quadratic" utility function. But, because expected utility is only meaningful as a measure of a subject's entire wealth (rather than an isolated gamble or a part of that wealth), the fact that the normal distribution takes on negative arguments makes it a very questionable choice for describing someone's wealth. Negative wealth in the present context has no meaning. Therefore, the reason for choosing this function (i.e., the mean-variance feature) seems invalid.

In "Risk-Adjusted Economic Analysis" (*NAAJ* vol. 2, no.1), Alastair G. Longley-Cook has the right idea to employ expected utility to value lines of business but errs, I think, when he seeks to remedy this particular function's lack of decreasing absolute risk aversion by transforming x to its return, $\frac{x \;-\; w_0}{w_0}$. What is w_0? That approach begs the question of what wealth is. There appears to be confusion with a traditional "accounting" sense of the term "wealth." Our holistic account is that expected utility is *characteristic* of wealth. I want to define wealth (under a reformed accounting system) as

$$w \;= u^{-1}\;\left[E(u(x))\right].$$

2. Suppose c(t,ε,ε) = t / (t+ε). Then

$$\alpha'(t) \;=\; \alpha(t)\;\lim_{\varepsilon\to 0}\left(\frac{\frac{t+\varepsilon}{t+2\varepsilon} \;-\; 1}{\varepsilon}\right)$$

$$=\; \alpha(t)\;\lim_{\varepsilon\to 0}\left(\frac{-1}{t+2\varepsilon}\right) = \frac{-\alpha(t)}{t} \quad .$$

Therefore $\alpha(t) = 1 / t$, and u(w) is log(w).

3. Generalizing example 2 somewhat, consider

$$c(t,\varepsilon,\varepsilon) \;=\; \left(\frac{b*t}{b*t \;+\; a*\varepsilon}\right)$$ where 'a' and 'b' are positive real numbers.

There are three cases.

a. First if a=b, we have example 2 and u(w) = log(w).

b. Second, if a >b, we have

$\alpha'(t) \;=\; \alpha(t)*\left(\frac{-a}{bt}\right)$. This implies that $\alpha(t) \;=\; t^{-a/b}$, so that

$$u(w) \;=\; 1 \;-\; \left(\frac{1}{w}\right)^{(a-b)/b}.$$

c. If b >a then $u(w) \;=\; w^{(b-a)/b}$.

This family (including all three cases) comprises the power utility functions.

4. If $c(t,\varepsilon,\varepsilon) \;=\; 1 \;-\; \left(\frac{\varepsilon}{t}\right)^{a}$ for a >1, then we have

$$\alpha'(t) \;=\; \alpha(t)*\lim_{\varepsilon\to 0}\frac{-\;(\varepsilon/t)^{a}}{\varepsilon} \;=\; 0.$$

Therefore, α(w) = bw. This is not a risk-averse utility function. Even though c(t,ε,ε)

a. is less than 1 ,

b. is strictly increasing in t,

it violates the first condition requiring
$0 \;<\; c(t,\varepsilon,\varepsilon)$ for all t a multiple of ε (and this is violated when t = ε).

5. Let us examine

$c(t,\varepsilon,\varepsilon) \;=\; c^{\varepsilon/t}$ where 0 < c < 1. Note that $c(t,\varepsilon/n,\varepsilon/n)^n = c(t,\varepsilon,\varepsilon)$, and so is on the boundary of feasible c-functions.

$$\alpha'(t) \;=\; \alpha(t) * \lim_{\varepsilon\to 0} \frac{c^{\varepsilon/t+\varepsilon} - 1}{\varepsilon} = \alpha(t) * \frac{\ln(c)}{t}.$$

This implies that $\alpha(t) \;=\; t^{\ln(c)} = t^{-b}$ for some b>0. Therefore, u(w) = ln(w) if b=1 and $u(w) = w^{(1-b)}/(1-b)$ if b is other than one. This is the same result produced in example 3 above. The c-function sequences do not characterize (uniquely) the utility function generated, but rather straightforward choices of c-functions (that also recognize one's position on the decision lattice) converge to the power utility functions.

If b<=1, the resulting utility function is unbounded from above on the domain of non-negative real numbers. There has been a fairly well accepted argument against allowing utility functions to be unbounded above.[18] The argument uses the Modified St. Petersburg paradox. The original St. Petersburg paradox provides motivation for the argument to use utility functions rather than expected value when valuing a risky prospect. The St. Petersburg game pays $\$2^n$ where n is the number of fair coin tosses made before the first tail turns up. This game has infinite expected value, yet nobody will pay more than a few dollars to play it. This was explained by using a log utility function.

Assume that u(x) is unbounded above. The Modified St. Petersburg paradox substitutes for the $\$2^n$ payoff, the value $\$x_n$, where $u(x_n) = 2^n$. This will have unbounded expected utility, so one should be willing to pay any sum to play the Modified St. Petersburg game. But there is a limit to how much people will pay to play the game, which is a contradiction.

The argument is unconvincing. First, the number of things in the world is finite; that includes the world's money supply. So the objection that a utility function for money cannot be unbounded above at impossible (and therefore meaningless) values is a singularly academic complaint.

Second, even if the money supply were infinite (suppose some had currency with Cantor numbers written on them and that the holders of the currency would thereby be entitled to produce new currency for any finite amount), both the number of things that money could buy and the people who buy things would still be finite. In

[18] See Karl Menger (1934).

that case, anyone with only a finite amount of money would be very poor indeed, having practically speaking nothing to spend on the game. On the other hand, someone with an infinite amount of money (and at least one would necessarily exist) would find any finite price for playing the Modified St. Petersburg game quite trivial. In the realm where the modified game would have any applicability, I could justify paying to play it. Let us say that I specify what I would pay based (in part) on my knowledge or opinion of how much money there is in the world. If we then make the supposition that there is much more money in the world, I should, *ceteris paribus*, increase my supposed price! As the supposed money supply approaches infinity, there is no reason my supposed price for playing cannot also approach infinity.

6. Practical Application

6.1 Pragmatist's Paradox

There are two other paradoxes that we need to address.

The first is that PMEU is to be applied to wholes that are incomprehensible because of their complexity. It is humanly impossible to consider every last detail in describing all possible worlds. That makes the valuation we seek an unobtainable fiction. We only experience the objectively real through the intermediary of signs, or models. We cannot experience it immediately, nor can we jump to a meta-level to know it. To the extent the signs are incomplete, they are fallible, and the success of PMEU is then necessarily approximate. But if we can identify a meaningful subset of possible worlds for which we can state well-ordered preferences and probabilities (and describe their relevant differences in complete detail), then PMEU will be quite serviceable. As a practical matter, we need not concern ourselves with describing all possibility; we merely take into account only what are believed to be the relevant factors that can be economically reflected in the decision. In most cases, our actions are of such modest impact in the grand scheme of things that there are only a few factors with which we are willing to characterize the possible consequences of a given action. All other aspects of the world are (fully) believed to be unaffected; that is, we will use a "*ceteris paribus*" clause in making our descriptions of the possible futures. Thusly, we partition the collection of possible worlds into a finite number of equivalence classes (where the equivalence relation is indifference). Our descriptions are fallible because of the possibility that we may believe some aspect lacks relevance to decision while it actually makes a material difference. We ought to make our description as complete as we can to guard against this fallibility. It is the model's ability to track objective (aspects of) reality that allows for the long run success of PMEU. Even if one concludes the idealized valuation is a fiction, it is a useful fiction, and its serviceability is increased with the completeness of the model.

The second paradox concerns our ability to realize the so-called utility function. As was discussed earlier, the Allais paradox implies that PMEU apparently fails as a descriptive theory, at least for individuals. One possible reason for this may be good for those espousing a normative theory. That reason is that people have trouble acting rationally or coherently, particularly in complex situations. If so, there is a real good served by PMEU as a normative theory and we should be eager to utilize it. In order to do so, however, we need to know or be able to rationally construct our utility function. It would seem likely that the main reason PMEU fails as a descriptive theory is our inability to state or know that utility function! It will not be enough to interview the principal decision-maker of the firm and ask his preference between various gambles in a number of situations, and then discern his utility function by piecing together his answers in some clever way. Presumably, if the utility theory is of any use, at least one of his answers will be wrong. So, PMEU, while perhaps being an ideal normative theory, is beyond our reach; it is an impossible theory to apply in the real world.

This is a rather harsh analysis. We are at least partially rational (this may be evident from the fact we seek rationality and value it). It is difficult to maintain coherence in the complicated world in which we find ourselves and expected utility is, at least, a good organizing technique. It conveys information about the coherence of our thoughts, and helps us to sort them out in a disciplined way. It provides a framework for reconsidering what beliefs we want to alter in order to remain coherent.

We need to decide on a utility function as rationally as we can. The method of discovering the utility function by eliciting choices among hypothetical alternatives is definitely problematic. Rather, the utility function should be constructed in accordance to principles and in view of rationally determined, worthwhile, long-term goals; that is to say, *normatively.* Let us posit two rational overall goals for an insurer: long-term growth and permanence. There may be other possibilities, but we shall restrict our discussion to these two.

6.1.1 Growth

Suppose, for simplicity, that we start out with a stake of \$w and assume this measures the subject's ultimate good. There is an opportunity to engage in a long series of (n) urn games each with a finite number (m) of outcomes. The pay-off of outcome j is denoted by r_j and its chance of occurring is p_j. In order to play, you must specify what percentage, α, of your stake you will risk on each and every draw. So, for example if you agree to play for 10 percent of your stake, then you are committed to wagering 10 percent of whatever your stake amounts to at the time of a given drawing. Once specified you must continue playing, wagering that percentage each time. The dollar amount of each bet will vary with your fortune but the percentage will not.

One is not allowed to bet enough to more than wipe out one's stake. Let r_{min} denote the minimum payoff among the r_j's and r_{max} be the maximum payoff. If $r_{min} \geq 0$ then we have a prospect that cannot lose and one should invest everything; $\alpha=1$. We will assume $r_{min} < 0$ and $r_{max} > 0$. Now α can vary from $\frac{-1}{r_{max}}$ *to* $\frac{-1}{r_{min}}$.

What is a proper choice for α? One might select the criterion of maximizing terminal wealth. If α is the percentage selected and n is the number of drawings in the series, the expected value of terminal wealth is

$$E[w_n] = w_0 * \sum_{k_1=0,}^{n} \sum_{k_2=0}^{K_2} \sum_{k_3=0}^{K_3} \cdots \sum_{k_{m-1}=0}^{K_{m-1}} \left(\frac{n!}{k_1! * k_2! * \ldots * k_m!} \right) * \prod_{j=1}^{m} p_j^{k_j} * \prod_{j=1}^{m} (1+\alpha * r_j)^{k_j}$$

$$\text{where } k_m = n - \sum_{j=1}^{m-1} k_j \quad \text{and} \quad K_i = n - \sum_{j=1}^{i-1} k_j$$

$$E[w_n] = w_0 * \left(1 + \alpha * \sum_{j=1}^{m} r_j * p_j \right)^n$$

$$= w_0 * \left(1 + \alpha * E[r] \right)^n$$

This is maximized when α is as large as one can make it, $\alpha = \frac{-1}{r_{min}}$, whenever $E[r]>0$ and as small as one can make it, $\alpha = \frac{-1}{r_{max}}$, whenever $E[r]<0$. Let us consider this strategy in more detail. Consider the case where the worst case outcome is losing everything; $r_j = -1$ for some j. The distribution of terminal wealth is such that one ends with \$0 with probability equal to $1 - (1 - \sum_{i \neq j} p_i)^n$.

For large n, this strategy leads to almost certain failure provided $p_j > 0$. It is quite difficult to defend as a rational strategy. In fact, it is the least preferred of all. The criterion is at fault here; it does not respond to the value we place on avoiding risk. In fact, it countenances behavior that is unreasonably risky.

Let us instead select as our criterion the maximization of the *expected rate of growth* over the period of n drawings and let n get very large. We will at least eliminate the all or nothing strategy. Betting 100 percent each time will, in fact, *minimize* the expected rate of growth of our stake; so, not only does the expected rate of growth criterion avoid selecting the all-or-nothing strategy as the best, it properly identifies it as the worst whenever there is any chance of losing all in a given draw.

We want to find the percentage of our stake to be bet on each trial that will maximize (in the long run) the expected rate of growth. For a large number, n, of successive trials the expected rate of growth for a given strategy, α, is

$$g_n(\alpha) = \sum_{k_1=0,}^{n} \sum_{k_2=0,}^{K_2} \sum_{k_3=0}^{K_3} \cdots \sum_{k_{m-1}=0}^{K_{m-1}} \left(\frac{n!}{k_1! * k_2! * \ldots * k_m!} \right) * \prod_{j=1}^{m} p_j^{k_j} * \prod_{j=1}^{m} (1+\alpha * r_j)^{k_j/n}$$

where $k_m = n - \sum_{j=1}^{m-1} k_j$ and $K_i = n - \sum_{j=1}^{i-1} k_j$.

This function is defined on the closed interval $\left[\frac{-1}{r_{\max}} , \frac{-1}{r_{\min}} \right]$. It is continuous and at least twice differentiable.

The expected rate of growth after n draws is $g_n(\alpha) - 1$ whenever α is wagered. If α=0, $g_n(\alpha) = 1$. If $\alpha = \frac{-1}{r_{\min}}$ or $\alpha = \frac{-1}{r_{\max}}$, then $g_n(\alpha) \xrightarrow[n\to\infty]{} 0$. Note that $g_n(\alpha) > 0$ on the interior of the domain in which we are interested. In particular, note that for α>0,

$$(1+\alpha * r_{\max}) > g_n(\alpha) > (1+\alpha * r_{\min}) \geq 0$$

and for α<0,

$$(1+\alpha * r_{\min}) > g_n(\alpha) > (1+\alpha * r_{\max}) \geq 0$$

As n increases to infinity, the sequence of functions converges. To see this, we can put the limit into the form $\sum_{j=0}^{\infty} a_j * b_j$ where the infinite sequences $\{a_j\}_{j=0}^{\infty}$ and $\{b_j\}_{j=0}^{\infty}$ are such that the a_j represent the state probabilities and the b_j represent the returns corresponding to the states. We have $a_j \geq 0$, $b_j \geq 0$, $\sum_{j=0}^{\infty} a_j = 1$ (i.e., converges), and $b_j \leq M$ (i.e., is bounded above).

We will call g(α) the *long term expected rate of growth*.

Now, $g_n(\alpha)$ is maximized on the interior $\left(\frac{-1}{r_{\max}} , \frac{-1}{r_{\min}} \right)$ because it approaches zero at the endpoints of the interval and is positive at α=0.

It is maximized at α on the interior if and only if

$g_n'(\alpha) = 0$ and $g_n''(\alpha) < 0$. After taking the first derivative with respect to α, and some algebra, we get

$$g'_n(\alpha) = h_n(\alpha) * \sum_{i=1}^{m} \left(\frac{r_i * p_i}{(1+\alpha * r_i)^{\left(n-1/n\right)}} \right)$$

where,

$$h_n(\alpha) = w_0 * \sum_{k_1=0,}^{n-1} \sum_{k_2=0,}^{K_2} \sum_{k_3=0}^{K_3} \cdots \sum_{k_{m-1}=0}^{K_{m-1}} \left(\frac{(n-1)!}{k_1! * k_2! * \ldots * k_m!} \right) * \prod_{j=1}^{m} p_j^{k_j} * \prod_{j=1}^{m} (1+\alpha * r_j)^{k_j/n}$$

and $K_i = (n-1) - \sum_{j=1}^{i-1} k_j$ and $k_m = K_m$

for α between 0 and 1. The function $h_n(\alpha)$ is similar to $g_{n-1}(\alpha)$. In fact, when E[r]>0,

$0 < (g_{n-1} - h_n)(\alpha) \leq g_{n-1}(\alpha) * \left[1 - \left(1+\alpha * r_{\max}\right)^{1/n} \right] \to 0$ as n gets large. Thus, $g_n'(\alpha)$ converges and being continuous, it converges to $g'(\alpha)$.

We then have $g'(\alpha) = g(\alpha) * \sum_{i=1}^{m} \left(\frac{r_i * p_i}{1 + \alpha * r_i} \right)$. The solution to this differential equation is $g(\alpha) = \exp\{\sum_{i=1}^{m} p_i * \log(1+\alpha * r_i)\}$

$$g''(\alpha) = g'(\alpha) * \sum\left(\frac{r_i * p_i}{1 + \alpha * r_i}\right) - \left(g(\alpha) * \sum_{i=1}^{m}\left(\frac{r_i^{2} * p_i}{\left(1+\alpha * r_i\right)^2}\right)\right)$$

Whenever $g'(\alpha) = 0$, $g''(\alpha) < 0$ because $g(\alpha) > 0$ on the open interval. Thus, the maximum expected growth rate occurs at the α that solves

$$\sum_{j=1}^{m}\left(\frac{r_j * p_j}{1 + \alpha * r_j}\right) = 0 .$$

We want to find the utility function, u(w), that will always pick out this strategy by the maximization principle. Consider the functions for arbitrary w and $\{r_j , p_j\}$

$f(\alpha) = \sum_{j=1}^{m} p_j * u[w*(1+\alpha * r_j)]$, the expected utility of a single play, risking α*w when the current stake is w.

$$f'(\alpha) = \sum_{j=1}^{m} w * r_j * p_j * u'[w*(1+\alpha * r_j)] \text{ and}$$

$$f''(\alpha) = \sum_{j=1}^{m}\left(w * r_j\right)^2 * p_j * u''[w*(1+\alpha * r_j)]$$

Because any feasible utility function will have a negative second derivative (u'(t) is positive but decreasing with t) any solution, α, to $f'(\alpha) = 0$ will maximize $f(\alpha)$. Furthermore, we want to require that α solve $\sum_{j=1}^{m}\left(\frac{r_j * p_j}{1 + \alpha * r_j}\right) = 0$ regardless of our choice for $\{r_j\}$ and $\{p_j\}$. The only family of functions that solves both for all choices of $\{r_j\}$ and $\{p_j\}$ is the log function and the linear translations of it. We have shown earlier that log(x) is a feasible utility function (it corresponds with c(t,ε,ε)= t/(t+ε)). This property of log utility is mentioned in Rubinstein (1976). I understand that it can be shown using Information Theory.

Thus, there is a rational reason for prescribing the logarithm as our utility function, particularly in financial decision-making situations. Its selection is consistent with and responsive to the notion of long-term success. Bear in mind, however, that the way we rationalized the choice has forced us to adopt a utility function that exhibits

constant relative risk aversion. We might be able to contour the selection to one with increasing relative risk aversion if we impose a maximum wealth level.

6.1.2 Permanence

The insurer has a fiduciary responsibility to do its best to remain solvent. If it is believed that there is the slightest chance that a firm's capital may be insufficient to handle its overall prospect from a contemplated action, it must devalue that prospect sufficiently so that the action is rejected. And if it should find itself in that situation (an adverse change in the present economic environment will have a direct impact on the future scenarios and their probabilities), it needs to value any action that eliminates the perceived risk of insolvency so highly, that such action is unequivocally preferred over maintaining the *status quo*. This includes anything from paying the asking price for interest rate options to merging with another company. Insurers often need to take risks, but with their free capital and never with assets needed to guarantee the minimum policyholder benefits.

The logarithm function is unbounded below as its argument approaches zero. For any finite probability distribution (of a total outcome) that includes an outcome of zero with positive probability (not matter how small), the expected (log) utility is unbounded below and equivalent to a certain zero outcome.

Ethically, this is the proper valuation. Because Option Pricing Theory is a reductive theory that takes the form of a zero-sum game, the sum of the values for the parts must add up to that of the whole (i.e., no value is ever created); and this produces the so-called put option to default. This is inconsistent with the notion of a guarantee. Log utility captures the fundamental properties a decision-maker having rational, worthwhile goals ought to exhibit.

So, not only does the principle of maximizing of expected log utility (PMELU) correspond with maximizing the expected long-term rate of capital growth, it also recognizes the value of actually being able to make the financial guarantees that constitute the product of the insurance firm.

6.2 Performance Measures

The paradigm of our interaction with the world, the triggering of actions that change (and hopefully improve) it is depicted in Figure 7. It envisions the process of valuation of alternative actions (ultimate goods) as the determinant for the action taken.

But valuation is performed on a model or sign of the ultimate good. We want to reduce the ultimate good to something we can manage, without reducing it too far.

In applying this paradigm to an insurer making decisions, actuaries take as (their closest estimate of) the firm's ultimate good the future free cash flow stream. They further reduce this using the concept of present value (discounting can be thought of as being a rational utility function for time). Even though their perspective is of necessity subjective, the adherence to rational discipline helps to keep them on track of success. It is the most reliable way of tracking the objective that is available to them. The two aspects of rational discipline are 1) the use of reliable methods in forming the system of partial beliefs and 2) the principle of maximizing expected utility (PMEU). The first is instantiated by the prudent development of a cash flow simulation model. The second is instantiated in the rational choice and use of logarithmic utility.

6.2.1 Cash Flow Simulation

The sheer size of the effort required to model so much in a single context is daunting, and the model will always be lacking in some way. This will no doubt always be the case, but cash flow testing and computer technology are rapidly making larger contexts possible. One has to make an economic decision regarding the value versus the cost of getting better information on which to base assumptions, as well. Uncertainty may always be with us, but technological advances continue to reduce it in a cost-effective way.

If one fully believes a deterministic model of the future events, we should judge him as being irrational. It is rational to distribute one's belief over a wide range of potential outcomes and assign degrees of partial belief to each; inductive reasoning forms the basis of this wherever possible, and conservative judgment will be substituted wherever evidence is too scarce. Information feedback processes will allow the more tentative assumptions to evolve. Barring the discovery of a causal law explaining how and why interest rates and other economic indicators change, the evidentiary approach seems to be the best method available to us. Causal laws are generally thought to be ideal because they would impart greater understanding (scientifically speaking). However, causal laws may be impossible in this field. Note that the use of any causal model that would accurately predict the future course of interest rates would be so successful as to quickly render the causal law worthless unless it somehow took account of its very use in the future. The changing economic indicators are relationally triadic in nature (the market participants collectively set them through their dynamic choices); causal laws can only explain the dyadic. If causal law is not forthcoming, then we are stuck with the best evidential methods we can find.

We also need to discount properly to take into account the time value of money. Ideally, we use the after-tax equivalent short-term Treasury rates in whatever interest rate environment we find ourselves. In this way short-term Treasuries, being risk-free cash equivalents, will be valued the same as cash; this is a necessary condition for rationality. Also, we have reduced the future to a set of yield curve scenarios. Embedded in each scenario is an inflation scenario (based on short-term, default risk-free rates). So, this method of discounting reflects the relative attractiveness of a sum of money received some time in the future under various conditions. It will reflect the fact that a large sum received in a high inflation future may well be less valuable than a smaller sum in a low inflation future. The use of risk free interest rates connects our valuation to the present (i.e., cash).

Some would recommend that we discount at Treasury yields plus some spread, but this is in the context of evaluating the expected (present) value of profits rather than the expected utility of those (present) values. The only reason to add a spread would be to address uncertainty. But, theoretically, within the context of a given scenario, one has eliminated that uncertainty; one pretends to have an omniscient view of things. Then again, prudence dictates that some uncertainty about what scenarios can obtain and their attendant probabilities, so some argument for a small spread is valid.

The degrees of partial belief need to be coherent in the sense Ramsey (1926) indicates (that is, they must conform to the axioms of probability). The manner in which future economic environments are generated assures this. Additionally, the specificity with which this discipline is practiced by a community of actuaries creates an objectified perspective; Ramsey referred to such probability distributions as a system of "chances."

This is the actuary's methodology in a nut-shell. The product of the process is the probability distribution of the firm's profitability. To the extent the degrees of partial belief are recognized and shared within a community, they are objectified. To the extent they are private, they are subjective. In either case, if the number of economic scenarios is n, then the result is a distribution of n present values. This does not assure a complete ordering of the preferences among a set of alternative actions. The probability distribution itself needs to be valued.

6.2.2 Expected Utility

The valuation is performed by taking the expected value (relative to the system of chances) of the utility function applied to the whole firm's present value of free cash flow, and then determining the unique sure-thing prospect with the same expected value. A subjective utility function may be used. A rationally determined utility function (the logarithm) is recommended.

The logarithm function is a rational choice for the utility function of an insurer. It is consistent with maximizing the long-term rate of capital growth. At the same time, it is unbounded below at zero. Thus, it devalues any action that will risk the solvency of the entire enterprise absolutely. Now, the log function will value a prospect as worthless if it has any measurable chance of insolvency. In such a situation, any action taken to eliminate the probability of insolvency will be highly valued over those that retain some probability of insolvency.

Expected utility is a measure of perceived wealth. Expected utility measures a probability distribution's relative attractiveness to the decision-maker. When reduced to financial outcomes, there will always be a unique sure-thing distribution with the same expected utility value as the one being measured. From the decision-maker's perspective, this sure-thing is equivalent to the decision value of the probability distribution. The value of the sure-thing equivalent is therefore a well-defined measure of the company's perceived wealth. The objective of increasing expected utility is thus the same as the objective of achieving the company's worthwhile goals.

6.2.3 Examples

6.2.3.1 Example 1: Risk Management

Risk is an important dimension of the value of a future prospect. All else being equal, a rational agent will prefer to avoid risk, and only take it on if some other dimension of value is sufficiently enhanced. Risk is characterized here by distributing partial belief over the relevant possibilities according to a probability distribution that has been formulated in as reliable a manner as possible.

The important characteristic of risk as it relates to future financial returns is that it is specific to the firm. The better we can understand the behavior of our liabilities, the better we can shape investment policy to manage risk. We must also distinguish between the risk aversion of the firm and that of its employees. Some hedging strategies employed today may seem unresponsive to overall value if viewed from the (whole) enterprise perspective. Market value based measures address risk directly, but it is generally the wrong risk. Like value, risk is subjective, but with some rational consideration, it can be objectified.

We generate a large number, n, of economic scenarios to represent all possible future environments. That is, we reduce the world to this model; these n scenarios. Our belief in the model (i.e., its reliability) is enhanced by increasing the number n. Of course, if n is too large the point of reduction is lost, along with our capability of taking action that will rationally increase the value of the enterprise.

To gain objectivity, we seek the countenance for the assumptions in the model of whatever outside parties deemed important.

The scenarios are generated in such a way that we believe that they

1. collectively, range over the possibilities and

2. individually, are equally likely.

The model of the enterprise consists of a detailed projection of each line of business over the n scenarios. Several objective functions can be used to measure results. One is the present value of future free cash flow (PV). Another is the notion of the additional assets required (AAR). The latter is defined as the least amount of capital needed for a given scenario to prevent an accumulated loss at any time-point in the future. The amount could be negative. The latest innovation to the risk based capital concept affecting variable annuities (RBC C3 Phase II) utilizes a modification of this idea—it floors the AAR at zero for each scenario. Present value best captures the idea of the value added by decision-making; it measures the decision itself by reference to the resulting change in (present) value. AAR focuses only on capital adequacy (that is, solvency risk); it directly indicates how much capital is needed. With either approach, one is left with n (equally likely) values. Something further is needed to sort out and compare one probability distribution from another.

6.2.3.1.1 Enterprise Risk Management

A method for evaluating the distribution of AAR is the conditional tail expectation (CTE). First, define the subset of scenarios (enumerated from 1 to n),

$T_\alpha = \{ j \ ; \ AAR_j \ \text{is in the top} \ \alpha \}$ and choose α to be rather small—10%, 5% or 1%.

Note that the set T_α is set globally—that is, for all product lines aggregated.

$$CTE = \sum_{j \in T} AAR_j * \Pr ob(j) \Big/ \sum_{j \in T} \Pr ob(j)$$. This is the average value of additional assets required given that the scenario is among the worst cases defined by the set T. Since Prob(j) = 1/n,

$$CTE = \sum_{j \in T} \frac{AAR_j}{\alpha * n}$$

Now if there are m distinguishable lines of business (or parts of the whole enterprise), we can calculate the CTE separately for each, but always keeping the subset of scenarios, T, the same. The subset, T, represents the tail of the distribution of AAR for the *whole* enterprise rather than that of any of its *parts*. This definition allows for the following:

$$CTE = \sum_{k=1}^{m} cte_k$$ where cte_k can be negative (unlike its risk-based-capital analog) or positive. If it is less than zero, then that means that the sub-line *contributes* capital. If it is greater than zero, the sub-line *uses* capital. This is a natural, and highly meaningful, capital allocation. One way to improve ROI is to reduce the required investment without reducing the return as much. This frees up other capital for high return investment. If a particular "risk" does not manifest itself in the abovementioned tail, then the cost of hedging it will certainly reduce future return without reducing the required capital; one might well deem such hedging counterproductive. One could try to hedge the change in CTE induced by the change in the environment (e.g., the change in the S&P 500 index). Here the returns will keep pace with the change in capital requirements and if the strategy can be succinctly described, the model can reflect them (in which case the strategy will reduce the CTE itself). The decision to hold derivatives will boil down to the economical use of capital.

6.2.3.1.2 Enterprise Value Management

Looking now at the present value of future returns (revenue less benefits, expenses, and taxes), we can define the value as

$$EV = Exp \left\{ \sum_{j=1}^{n} \log(PV(j)) / n \right\}$$ (provided PV(j) >0 for all j) and zero otherwise.

The logarithm function serves as a utility function. The expression inside the brackets is the expected utility of the present value of profit distribution. The remaining part of the equation recovers the sure-thing equivalent value—the sure thing result that has the same expected utility value.

The reasoning behind the choice of the log function was discussed in some mathematical detail above. We demonstrated that the strategy that maximizes the expected growth rate of wealth is the one that maximizes expected utility if and only if the utility function is the logarithm function.

Secondly, this choice of utility function places a huge value on avoiding very bad results. Note that if some risk aspect of the whole business was severe enough to pull the PV(j) close to zero for some scenario j, the log utility will penalize that outcome to such an extent that the use of a derivative strategy that brings that scenario's outcome up to a positive value would improve EV, and would so be warranted. If a PV(j) <=0, then EV is set at (absolute) zero and any measure to remedy that outcome (without causing another scenario's outcome to become negative) will infinitely improve EV on a relative basis.

The decision to use derivatives will be *like any other decision* (e.g., marketing a particular product at a particular price): it will need to improve EV the most among the identified alternatives. Risk is normally thought of as a side-constraint. Under the EV framework, risk is built into value. As a conceptual framework, it is objective, consummately coherent and normatively powerful.

Since EV gives weight to the best as well as the worst case scenarios, it might be expected that evaluating derivatives in this framework would result in lower outlays for derivatives. At the same time, the severe de-valuation of solvency-threatening situations assures that we are hedging the proper risks.

An under-capitalized company (that has determined it has a chance of insolvency, or one in which a possible scenario will produce particularly low results) will view commercially available hedging instruments as a real bargain. On the other hand, well-capitalized firms using PMEU (with the log function) may evaluate these instruments as being too expensive. Derivatives are usually offered as a way of hedging bets. But if a bet is a prudent one, why hedge it? Derivatives should only hedge (or undo) imprudent bets. What would make a given prospect imprudent is the existence of another action that will change the prospect to one which is of greater value to the firm. Derivatives should be evaluated in exactly the same way any other decision is valued; namely, according to how much their purchase will increase wealth as tracked by the utility function.

6.2.3.2 Example 2: Mergers and Acquisitions

Suppose that firm A has performed the stochastic projection of its future profits from its in-force block and finds that there is a measurable chance that its existing surplus is inadequate (i.e. under some scenario with a non-zero probability of occurrence future returns have a present value of zero or less). Now, firm B performs the same projection and agrees with firm A's assumptions about everything. Only B values A as though it had merged with B. It may value A's prospect as adding more value than A does. So a market for A might exist. B, depending on how well it is capitalized (i.e., positioned), will bid nearly the expected value A (under the commonly believed assumptions), E[A]. Consider C. Like A, C has a positive probability of insolvency by itself. C may be in a position to bid more than E[A] because of A's hedging properties in the context of the fused entity A-and-C.

6.2.3.3 Example 3: Macro-Pricing

"Prices for the insurer's products are generally set by the market." This is a vague notion. The pricing function needs a way of fixing belief as to how much will sell at a given price and what further volume would be obtained if the prices were somewhat lower. As with any other profession, marketing experts cannot absolutely *know* (in Nozick's sense) the precise supply and demand curve, but they can seek out the most reliable methods available to come as close to knowledge as possible. One would expect some kind of feedback information would modify the assessed supply and demand curve. Perhaps marketing's sales goal should be determined by the company's choice of price (and volume) read off the product's estimated supply and demand curve. Whenever the actual sale result over-shoots or under-shoots the anticipated volume, adjustments to the supply and demand curve can be made. Marketing could have a lower priced product to sell only to the extent they can make up *value* in volume! This method will clarify the notion of "making *it* up in volume" by firmly addressing to what "*it*" refers.

Another important aspect of pricing is the effort and expense of obtaining good assumptions. There is a cost (an uncertain one) to getting reliable information and some (prior and often visceral) belief regarding the payoff of getting such information will clearly need to be made.

Armed with beliefs about nearly every detail, the pricing model can better reflect the value of the decision of pricing the products at each contemplated level. It supposes that one large joint-product pricing decision will be made. Whatever *package* of products and corresponding expense levels and prices produces the highest decision value (in the context of existing blocks of business) represents the most desirable set of prices. This amounts to a plan of operations.

6.2.3.4 Example 4: Attribution

6.2.3.4.1 To Lines of Business

The expected utility method described here does not lend itself too well to analysis—the breaking up of a whole into parts. The sum of the expected utility values of the various parts does not add up to that of the whole. On the other hand, the concept of capital is subsumed by enterprise value (EV). EV is free capital. It is the thing we seek to grow. It is our ultimate good. One is always referring to the impact on the whole enterprise when considering a decision. To attribute performance to a given line of business, one would evaluate the resulting difference in EV with and without that line of business. Call this the line's "relative value."

6.2.3.4.2 Over a Period of Time

Changes in this wealth measure can be attributed (if desired) to various activities of the company and various environmental changes:

1. Current period's profit versus the expected profit that can be analyzed by source.

2. Changes in future expected profits due to updating actuarial assumptions as actual experience unfolds.

3. Changes due to adding new business at variance with sales goals.

4. Performance of asset management in light of liabilities.

   4.a. Yield curve shifts.
   4.b. Investment decisions.

If an action served to increase item 1) but reduce item 2) sufficiently, it would not add value and may waste value. If such an action was describable as a gamble, only to the extent it later succeeded, would it add value in later periods.

If the investment department happens to take a specific view of the future that is contrary in some sense to that assumed in the model (i.e., contrary to the "evidence"), it is free to "bet" on that scenario. Initially, it will appear to be a sub-optimal decision relative to "scientifically" determined assumptions under the expected utility measure. If, as time passes, reality unfolds in such a way that the bet is a winner, credit will be attributable to the investment department at that time whether the reason for the success is a new insight or dumb luck. If some insight is gained, and if enough evidential weight can be given to it, then the model can be revised. If revised, future decision valuations will give due weight to the new insight and not automatically deduct points from performance. As in GAAP accounting, expected utility provides a certain amount of slack. The benefit of the doubt is given to one's own set of partial beliefs; but, that is simply the nature of belief. As time progresses, the reality of cash transactions that transpired will ultimately assert itself; this is also true of GAAP. The chief difference lies in the distribution of belief among many possibilities rather than just one, so GAAP gives somewhat more slack. Under expected utility, only to the extent a decision to bet against a statistically grounded model has proven itself will credit be given. This will occur, ideally with an adjustment to the model moving forward.

6.2.3.5 Regulation

When brute reality asserts itself on an incompetent or fraudulent set of partial beliefs (one that did not track reality well enough), the decision-making may well lead to failure not only of the decision-maker but also its policyholders. A subtlety of the utility theory being espoused here is that it is an ethical theory, pertaining to what one ought to do. For instance, an insurer ought to be risk-averse to the point of never willingly risking the policyholder. This is an aspect of one's world-view that can be highly valued. However, some may not share this view. They may be inclined to see such a risky position as leverage, even a put option. They can always adopt the view (sincerely or otherwise) that what appears to be a real possibility of insolvency to nearly anyone else as having a zero probability of occurrence. Such a stance is, in the general scheme of things, irrational to those who hold that ethical principle in high esteem, but that kind of occasional irrationality is doubtlessly a fact of the world.

So, only the fittest beliefs will survive. This will be little consolation to the victimized policyholders. The chief weakness of this application of utility theory is its lack of skepticism. Reliable methodology is supposed to prevent the worst from occurring. To many that is naive. Even if we could rule out duplicity, fallibility is still a reality. If regulation is needed at all, it is to address these possibilities. But that is no problem peculiar to utility theory. Rather, it is a separate problem that is present with any alternative.

Regulators might adopt a system of *chances* (in Ramsey's sense); a set of degrees of belief within a system of beliefs and partial beliefs; a behavioral model specified for regulatory purposes. Regulators need a tracking device to monitor solvency and standardized cash flow testing with specific prescriptions for redressing potential insolvency would seem to fit the bill. The cost would be little more than the cost of today's cash flow testing requirements. This would put a constraint on the belief set developed by the insurer internally for pricing, but only by introducing a new set of facts about the world. (Indeed, the manner in which formula reserves are prescribed today affects pricing.) A firm that showed a chance of insolvency would be required to take some minimal step to redress the issue. The advisability of having a "strong" firm take over a hopelessly weak one (i.e., one in need of a bailing out), would be much more demonstrable from the regulator's perspective. The action would track a reliably formed belief.

This possible application of utility theory is similar in spirit to the cash flow testing requirements in the current Standard Valuation Law. In some ways it is stronger by prescribing the behavioral and interest rate models. But it is an improvement in the following sense. By putting the onus on an employee of the regulated firm, cash flow testing is subject to the same skeptical criticism that creates the need for regulation in the first place. Such skepticism undermines the regulator's proper goal. Using a communally accepted standard set of assumptions (arrived at in a professional, disciplined and fair inquiry rather than established by method of "authority") reduces skepticism. For regulatory purposes, it seems imprudent to allow a wide range of belief on these matters. If so, a standard model (reliably maintained) fits the bill optimally because its use provides maximum clarity. Perhaps the actuarial profession should work together to bring about an objectified set of partial beliefs (or chances) that are grounded in methods as reliable as are available.

6.2.3.6 Fair Value of an Insurance Firm

EV is an ideal basis for an accounting system or performance measurement. Today, the accounting world seems rather enamored with a similar concept: that of fair value. One merely replaces the log-utility function and economic scenarios with those of the "market." The market's utility function need not be observed; Option Pricing

Theory is the finesse. Were it observed, however, one would find that it lacks the normative power of log-utility. In addition, one would realize that the coherence of the market is not necessarily stable over time; rather, the market's unobserved utility function is likely a capricious one.

This is not to say that Option Pricing Theory is irrelevant; it is an extremely compelling and reliable model for market prices which, of course, represent important facts. The Option Pricing Model ought to be used in the modeling of future prices of securities that one may decide to buy or sell (we, rationally, ought to believe in those hypothetical future prices); but there are better over-arching frameworks of measuring the worth of the enterprise (and decisions made to enhance the value of the enterprise).

It is also not to say that the collective wisdom of the "market" ought to be ignored in general (though that is possibly a good course to take). The other aspect of (enterprise) valuation is the fixation of degrees of partial belief; that is, the descriptive probabilistic model of future contingent events, the behavioral assumptions and the economic scenarios. As individual agents in the economy, we fix these beliefs through inductive methods where possible and otherwise substitute judgment; ideally, we use the most reliable methods available to us , but the result is still fallible. The "market" may be seen as having degrees of belief as well; and if we could discern them (i.e., separate them from the utility function), this *might* be a good way to establish our own. It might be that this method would be a good adaptive approach (but if that is valid at all, it is not *necessarily* so and it would be hard to rationalize such a position). Unfortunately, we cannot separate the market's belief from its utility function. Even if its beliefs would be an improvement over the traditional methods we use, that is an academic point; the market's beliefs are inscrutable.

Fair value is to be informative to other financial decision-makers besides the regulators; people interested in possibly investing in the insurance company, for example. With an ideal, standard model, the valuation would be completely objective as well as being meaningful. It is unlikely that we would *agree* enough about the future to render the perfectly objective model. Too much is expected of objectivity. The general complaint about PMEU is that it is too subjective, even vague. It is no more so than other measuring methods used in life insurance company management and accounting. PMEU, in fact, begs for specificity when properly employed. The argument that PMEU is too subjective for professional use is actually a non-starter. Any method in dealing with the unknowable involves subjectivity; that is, belief rather than knowledge. The real distinguishing factor is how closely the methods track goals that are worth pursuing, or (negatively) how arbitrary they are. PMEU can be instantiated in as objective a manner as any other method. As we have seen, the choice of utility function can be made rationally, even by those possessing imperfect rationality. The details of (i.e., the beliefs underlying) the behavioral models (of capital markets as well as insurance policyholders) can be ratified and maintained though professional (i.e., reliable) methods.

7. Conclusion

The Expected Utility Hypothesis has been around for quite some time, at least as a theory. It remains the leading principle behind modern decision theory since it was developed in some detail by von Neumann and Morgenstern in their *Theory of Games and Economic Behavior*, published in 1947. It was first introduced, and in many ways bettered, by Frank Plumpton Ramsey's non-reductive approach of 1926.

PMEU appears to have not been employed in the above contemplated manner. This has been due partially to difficulties in analyzing results into parts (the measure is not linearly combinable) and most people seem interested in explaining complicated things in terms of their component parts. Such reductionism is necessary at some level; and modeling always reduces reality to a simpler system of beliefs. However, reduction to too low a level causes us to miss a very important aspect of value; that is, the interrelationships among the parts. It also leaves out the decision-maker whose activities create value that would not exist otherwise. Holistic approaches acknowledge these important aspects of value.

Reductive measures tend to close off the various parts from one another in terms of their day -to-day decision-making; each taking care of his or her piece of the puzzle without properly relating to the other parts of the firm. If the activities of the various parts of the firm are coordinated to improve overall wealth as described in the expected utility framework, then the firm's true goals will be coherently addressed by the appropriately delegated actions taken. PMEU will help in making rational choices in a complicated environment. In complex business environments the formal use of such a decision model (e.g., in the performance measures) is a good communication device.

We have described a very general and abstract decision model. Any implementation of this method can only approximate the ideal; and the ideal description may itself be only a useful fiction (i.e., approximate and fallible). Clearly, many of the important assumptions required may not have much credibility in a scientific sense. Perhaps credibility would be gained in time, but whether or not it would really does not matter. We must deal with this uncertainty in as effective a way as we can. If belief is the basis of action and if some action is necessary (even if that is deciding to do nothing), then some kind of belief must be formed; the best belief we can muster. The problem of needing to make a great deal of assumptions in order to employ utility theory is a challenge, but not a problem peculiar to PMEU. The method should be judged not against the ideal, but rather against the real alternatives. Its subjectivity is often cited as a weakness. But subjectivity is a separate matter; if we had perfect knowledge, PMEU would determine objectively the best course of action. We should endeavor at all times to point to the best instantiations of a perfect valuation.

Expected utility is a formal framework for acting in accordance with one's system of partial beliefs. It tracks coherence. By itself, it will not accomplish very much. The actuaries' daunting task is to form beliefs that track facts as reliably as possible. Ultimately, it is in that arena that actuaries can best serve their clients. The framework advocated here, it is hoped, provides the context in which this quest for reliability is most meaningful. For value is made real and present in the world through our actions. Its realization comes only from effort; it is caused. Value is not observed; it is created. Peirce said that (final) truth is what we would all come to believe (as a community) in the long run, not by mere communal fiat, but rather by discovery through relentless inquiry and conscious mediation. This is an evolutionary idea. The truth is what survives to be built upon. Nozick makes the same point in *Philosophical Explanations* about knowledge and value itself. Intrinsic value is merely value-in-potential. Value is made real as the result of active, conscious processes.

References

Gibbard, A., and Harper, W.L. 1978. Counterfactuals and two kinds of expected utility. *Foundations and Applications of Decision Theory* 1. Dordrecht, Netherlands: Reidel Publishing Co. Reprinted in *Decision, Probability and Utility* (eds. Gardenfors and Sahlin) 1988. Cambridge, U.K.: Cambridge University Press.

Hookway, Christopher 1992. *Peirce.* New York, NY: Routledge.

Longley-Cook, A.G. 1998. Risk-adjusted economic value analysis. *North American Actuarial Journal* 2(1): 87-98.

Menger. K. 1934. The role of uncertainty in economics. *Zeitschrift fur Nationalökonomie* 5, as translated in Shubik, 1967, editor, *Essays in Mathematical Economics*. Princeton, NJ: Princeton University Press.

Nozick, R. 1974. Reflections on Newcomb's Problem. *Scientific American* (March) 102-108, reprinted in *Socratic Puzzles.* 1997. Cambridge, MA: Harvard University Press.

———. 1981. *Philosophical Explanations*. Cambridge, MA: The Belnap Press of Harvard University Press.

———. 1993. *The Nature of Rationality*. Princeton, NJ: Princeton University Press.

Peirce, C.S. 1877. The fixation of belief. *Popular Science Monthly*, reprinted in *Philosophical Writings of Peirce*. 1955, edited by J. Buchler, 5-23. New York: Dover Publications, Inc.

———. 1878. How to make our ideas clear. *Popular Science Monthly*, reprinted in *Philosophical Writings of Peirce*. 1955, edited by J. Buchler, 5-23. New York: Dover Publications, Inc.

———. 1903. The principles of phenomenology. Printed in *Philosophical Writings of Peirce*. 1955, edited by J. Buchler, 5-23. New York: Dover Publications, Inc.

———. 1992. *Reasoning and the Logic of Things*, *The Cambridge Conference Lectures of 1898,* edited by K.L. Ketner with an introduction by K.L. Ketner and H. Putnam. Cambridge, MA: Harvard University Press.

Ramsay, C.M. 1993. Loading gross premiums for risk without using utility theory. *Transactions of the Society of Actuaries* XLV: 305-37.

Ramsey, F.P. 1931. Universals (1925). Published posthumously in *Foundations of Mathematics and other Logical* Essays, edited by R.B. Braithwaite, 156-98. London: Routledge and Kegan Paul. Or in *Philosophical Papers*. 1990, edited by D. H. Mellor, 52-94. Cambridge: Cambridge University Press.

———. 1931. Truth and Probability (1926). Published posthumously in *Foundations of Mathematics and other Logical* Essays, edited by R.B. Braithwaite, 156-98. London: Routledge and Kegan Paul. Or in *Philosophical Papers*. 1990, edited by D. H. Mellor, 52-94. Cambridge: Cambridge University Press.

———. 1931. Chance (1928). Published posthumously in *Foundations of Mathematics and other Logical* Essays, edited by R.B. Braithwaite, 206-11. London: Routledge and Kegan Paul. Or in *Philosophical Papers*. 1990, edited by D. H. Mellor, 104-9. Cambridge: Cambridge University Press.

———. 1931. Knowledge (1929). Published posthumously in *Foundations of Mathematics and other Logical* Essays, edited by R.B. Braithwaite, 258-9. London: Routledge and Kegan Paul. Or in *Philosophical Papers*. 1990, edited by D. H. Mellor, 110-1. Cambridge: Cambridge University Press.

———. 1931. General Propositions and Causality (1929). Published posthumously in *Foundations of Mathematics and other Logical* Essays, edited by R.B. Braithwaite, 237-55. London: Routledge and Kegan Paul. Or in *Philosophical Papers*. 1990, edited by D. H. Mellor, 145-63. Cambridge: Cambridge University Press.

Rubinstein, M. 1976. The strong case for the generalized logarithmic utility model as the premier model of financial markets. *Journal of Finance* 31:551-71.

Sahlin, E.-N. 1990. *The Philosophy of F. P. Ramsey*. Cambridge: Cambridge University Press.