Instantiating Holism:
Beyond ERM's Half Measures

Christopher Perrin, FSA

Abstract

This paper focuses on the role of holism in enterprise risk management (ERM). The performance measure known as market consistent embedded value is discounted for being overly reductionistic. The concept of an economic capital charge or credit is seen here as a potential bridge from the reductionistic to the holistic approach. It shows some promise for achieving holism, but falls short. An alternative performance measure, the Equivalent Logarithmic Utility Value (ELUV), is then described and argued to be a more complete embodiment of holistic risk management. ELUV can be seen as the expected value of a transformed probability measure, and this transformation sheds new light on the economic capital concept. In addition, to the extent the underlying probability models can be relied upon, ELUV can be used to define the concept of optimal capital. However, holism also encompasses how we face uncertainty, and so the extent to which probability models can be relied upon is an issue holistic ERM must address.

1. Introduction

Holism seeks the understanding of an object in its entirety and without the filter of mediating signs (or models); it is an ideal. Models are reductions of reality. We are incapable of experiencing reality immediately.[1] We require the reductions afforded by models. This is so much the case that reducing reality is an ingrained (and perhaps biologically based) habit; we are all reductionists by nature. But, while they are absolutely essential to us for interacting with reality, reductions run the risk of distorting or even ignoring important dimensions of it. Even if we could completely dispense with reduction, there would remain the issue of the economy of information processing. If every aspect of a decision choice were given full consideration, it may simply take too much time and energy to calculate the best action to take in complicated situations. Still, it would be very valuable to the decision-making process to resist the habit of reducing things *too much*; to capture as many dimensions of a reality as we can. The best pragmatic instantiation of holism that we can muster is what we should try to realize.

Enterprise Risk Management (ERM) is a movement that takes this (pragmatically) holistic view of decision making. The whole is not merely explainable in terms of the parts alone; one must also consider how the parts relate to the whole. The traditional approaches to decision making that ERM would supplant have been rather reductionistic in regards to this part/whole relationship. For all traditional approaches, the parts of the firm are treated independently, and maximizing the performance of each part is *presumed* to be the best way to ensure the overall success of the firm. These approaches should be displaced because optimization of the parts will not optimize the whole in general.

It is often said by its advocates that the reason for pursuing ERM is to increase the firm's overall value to its stakeholders. The two major stakeholder groups for insurance companies are the policyholders and the equity owners. In the case of policyholders, increasing the ability of the firm to meet its product guarantees adds value to them; increasing surplus or reducing risk will serve this end. For stockholders, there is a case to be made that there is such a thing as too much capital relative to risk, and so a return to them of the excess is in order. This situation is more likely to occur when the firm's growth opportunities are somewhat limited. For other firms, free surplus is a good in that it provides needed growth capacity. In either case, the concept of economic capital has become a focal point of the ERM movement. In order to accomplish the goal of ERM, the concept needs to be integrated into decision-making at many if not all levels of the enterprise.

Now, decision- making is shaped by performance measurement. Therefore, it is important that the performance measure reflects what is worthwhile; that it tracks value. The meaning of the phrase "economic capital" varies, although its purpose, the idea that it represents the minimum surplus an insurance firm needs to have, is common to all the variations. The specific form it takes is invariably in reference to an underlying performance measure. So, it only makes sense to consider the two together.

[1] Read as much C.S Peirce as you can. *Peirce on Signs* (1991) is a very good anthology of his most accessible work on the topic of signs.

This paper is meant to be an elementary exposition of these issues. But while it is intended for the uninitiated, it is hoped that the experienced practitioner can find some value in it as well. I will only mention a few rather general ideas and try to relate them to the paper's overall theme of holism. My main presupposition is that holistic approaches to ERM are preferable. I contend that the performance measures being touted today are not holistic at all, and that attempts to retrofit some degree of holism through the constraints imposed by economic capital measures are only partially effective in achieving holism. I then propose a measure that is more thoroughly holistic in its approach. Once we have reduced the enterprise to a robust (enough) probability distribution of financial present values, the approach relates not only the parts of the enterprise to the whole, but also the parts of the distribution to the whole distribution.

The remainder of this section describes very briefly the fundamentals of the models used and some elementary measurements that are needed for ERM. Section 2 criticizes the performance measure known as "market consistent embedded value" on holistic grounds.

Section 3 describes a couple of versions of economic capital. It discusses how economic capital can be integrated by making a simple adjustment to the existing performance measurement. This is meant to convey the appropriate consideration for enterprise-wide risk in decision-making that is geared toward maximizing the performance measure. Section 4 examines a concept that goes beyond economic capital. It is a full-blown alternative to existing performance measures or accounting methods. In addition, it subsumes the notion of capital. It is a natural extension of economic capital and a purer embodiment of the holistic view. It also opens up another way to view capital; not only from the perspective of the policyholders, but also from that of the equity owners.

1.1 Stochastic Scenarios

We generate a large number, n, of economic scenarios to represent all possible future environments in order to handle the problems the future presents. That is, we reduce the world to this model; these n scenarios. Our belief in the model is enhanced by increasing the number of scenarios generated. A smaller choice of n means a greater likelihood that the description of the future is (perhaps tragically) incomplete. Of course, if n is too large, the point of reduction is lost, along with our capability of taking swift action that will increase the value of the enterprise.

The idea here is that 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; that is, we naturally treat each as having a 1/n chance of occurring.

Considering the whole enterprise, the probability distribution will be restricted to nonnegative results. There is no need for us to consider negative results because either result is indicative of ruin. A negative result would have the same pragmatic consequences as the zero result. Therefore, we can lump them together into the same state.

1.2 Model Risk and Sensitivity Tests

In addition to a sufficiently large number of possible environments, the work requires assumptions regarding policyholder behavior. These must be obtained from as reliable a method as possible; for example, actuarial experience studies and inductive methods used in the science of statistics. Reliable methods of developing assumptions will keep the discounting of the results to a minimum, but it seems there will always be areas in which these assumptions can be questioned. It would seem imprudent to come to the belief that all doubt is extinguished by the stochastic model. There will always be some degree of residual doubt regarding the behavioral assumptions employed, and so we need some way of addressing this potential inadequacy.

If "scientifically" determined by statistical analysis, the assumptions comprising the model will be based on past observance; and for all the virtues this manner of establishing belief may have, this may also be a source of the model's inadequacy. It is likely that the probabilities of important events (scenarios) that would negatively impact the firm's overall value are underestimated (perhaps as zero) in the model.[2] Acknowledging this, we should consider what extreme contingencies might occur that would most harm the company's value without necessarily fixating on the probabilities for the events. To the extent we can adequately incorporate this skepticism into the stochastic scenario set using non-inductive methods, we should. This is yet another source of residual doubt regarding our probability model.

To address this doubt, the enterprise might establish a *policy* of proactively addressing those residual risks. Each key risk identified will lead to the development of another set of scenarios the enterprise will want to test in its management of risk. We will refer to these special

[2] See Taleb's books *Fooled by Randomness* and *The Black Swan* for thorough exploration of the extremely important issue of skepticism.

scenarios as sensitivities. For each baseline scenario, there will be the sensitivity's scenario that assumes 100 percent chance of the event occurring in the first period, but following the baseline scenario thereafter. Examples include the following obvious contingencies:

1. a sudden 30 percent drop in the S&P 500 Index,
2. a 100 basis point increase in interest rates,
3. a spike in mortality due to an exogenous event such as influenza.

For each sensitivity, we can re-run the model on the condition that the extreme event occurred, estimate what consequent, negative impacts on the enterprise would result, and then consider actions that would most economically reduce the event's potential impact. We would not attempt to associate a probability of the extreme events happening because what we are attempting to do is to acknowledge that we do not really know. We form beliefs (or rather degrees of partial belief) from past observations; inductive methods are usually reliable but they are not always so; particularly when the events are rare in occurrence and substantial in their impact. The problem is that we do not have infallible methods of forming belief. So we must to some degree adopt a skeptical stance.

1.3 Measurement

The model of the enterprise consists of a detailed projection of each line of business over the n scenarios. Several mathematical techniques or functions can be used to measure results. One is the present value of future cash flow (PVFCF), an age-old actuarial reduction technique. Another is the notion of the additional assets required (AAR). The recent changes to the risk based capital formula affecting fixed and variable annuities (RBC- C3 Phases I and II, respectively) utilize the AAR concept.

AAR is defined as the least amount of capital needed for a given scenario to prevent an accumulated loss at any time-point (normally year-ends) in the future. AAR relates to the enterprise's liabilities. It is generally performed without surplus assets; the idea is that the calculated AAR is to be compared to actual surplus. (To check for solvency, one could incorporate existing surplus assets in the AAR calculation; this would, of course, reduce the resulting AAR.) Under a scenario, one begins the projection with surplus equal to zero. Some of the firm's assets are selected to back the liabilities so that the accounting value of each is the same (for example, life insurers typically use amortized cost for fixed income assets). Over time, accumulated surplus grows with interest, realized capital gains and with revenue (earned interest from the assets, contract fees and charges). It is reduced by realized losses, expenses, taxes and benefits. At certain time-points of the scenario, the present value of the accumulated surplus is noted. AAR for the scenario is defined for this discussion as the lowest present value among the time-points multiplied by minus one. Thus if the lowest present value is negative, the AAR is positive; and if the lowest present value is positive, the AAR is negative.

AAR is to be calculated for *all liabilities together* while noting at which point in time the present value of accumulated surplus is lowest; we will denote this as scenario's "low-point." With this information, we can decompose the AAR for a given scenario into component parts (aar_k) for each sub-line. For a given scenario, and line of business (k), aar_k is the present value of the line's surplus as of the scenario's low-point. That is, the AAR for the whole company is then expressible as the sum of the individual aar_ks.

PVFCF measures the scenario's result over the entire projection. Rather than picking out to low-point of accumulated surplus in the projection, it looks at the present value of accumulated surplus at the very end of the projection's observation period. Also, it does not require an association of assets with particular liabilities. One can begin with any assets assigned to liabilities one sees as relevant. Any embedded surplus at the beginning of the projection will reveal itself as a higher PVFCF.

PVFCF 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, risk); it directly indicates how much capital is needed. Both measurements are made at the "scenario" level. One is likely to completely ignore scenarios in which the AAR is relatively small; all the focus is on the worst cases. On the other hand, in considering PVFCF, all scenarios will weigh in to some degree. 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.

2. Market Consistent Embedded Value

A performance measure that is gaining in popularity is called market consistent embedded value (MCEV). It is simply the market or fair value of assets minus the "market" value of liabilities. It is a PVFCF method. Theoretically, the market value of an asset or liability is the expected value of the present value of future cash flows. However, expectation is taken with respect to a risk-neutral probability measure. One would as a practical matter observe the market value of the asset or liability as of the valuation date. If no such observation is possible (as when no market exists) then one resorts to theory and calculates it in such a way that is consistent (enough) with the actual market value of close continuers using a risk neutral probability measure.

MCEV is the opposite of holistic. Under the asset pricing model, something's value is independent of the observer's circumstance. Adding a new policy to the set of liabilities and a new asset (purchased with the premium received) will either add or subtract "value" regardless of how much enterprise risk is increased or decreased. But, in reality, adding to the overall risk (without adding enough expected return) would actually detract from the true value added by a new piece of business *for* the enterprise. Market value is context free, but context actually matters.

Now if the only *use* of a given financial instrument is to buy or sell it, then its market value may well be equivalent to its true value. This is the perspective of a trader whose only concern is how much his cash increased or decreased after trades are finally settled. But (even) a trader is (or should be) concerned with his own survival as a trader. The result *is* path dependent because the trader can possibly run out of cash before the trades are settled. If no one gives him more cash, he is through *qua* trader. So even in the rather cut-and-dried world of the trader, market valuation misses some aspect of true value. By taking such a completely reductionistic approach to valuation, MCEV fails to capture important aspects of an enterprise's value, those aspects that are sensitive to the observer's circumstances.

I am not saying that the asset pricing models are wrong; they seem to be good enough at explaining market prices to the extent the prices of things that have actual markets are reducible to the inputs of the asset pricing model.[3] The point is that what I would call "value" is not simply price. Of course, an ideal valuation method will not ignore the fact of prices if it is to be rational and reliable. The price of an asset I might purchase will bear on my decision of whether or not to purchase it, but the decision ought to be grounded in what I believe most increases overall value; in order to decide to purchase, the price really ought to be less than its value *for me*. My success will depend then on the reliability or luckiness of my belief-fixing process.

If one views the prospect as having a (real-world) probability distribution, the market value can be realized as the expected value of a *transformed* probability measure that has the risk-neutral (or martingale) property. To illustrate, for several classes of probability distributions (including the ubiquitous lognormal), the Esscher Transform, suitably calibrated, will produce the required martingale property.[4] So, if p(x) is the "real-world" (i.e., believed) probability for the random variable x (where x represents the PVFCF), we can transform it into:

$$q(x) \quad = \quad p(x) \quad * \quad \frac{e^{hx}}{E[e^{hx}]} \quad \text{for the suitable value of h.}$$

We will then have:

$$MV \quad = \quad \sum x \quad * \quad q(x)$$

Typically, the value of h that satisfies this will be negative (although this does not seem to be necessarily the case). A negative h means that the transformed probability measure q(x) gives more weight to the lower outcomes relative to the measure p(x) and less relative weight to the higher outcomes. So, the market value "risk adjusts" relative to the expected value of the outcomes with respect to the measure p(x); only a fool would value a prospect (with a positive variance) as the expected value with respect to measure p(x).

We can see from the Esscher Transform that MV is a reductionistic valuation method. Say there are n possible outcomes x_j where j varies from 1 to n. The transformation of $p(x_j)$ to $q(x_j)$ does not depend on the other possible values. The denominator appears to so depend, but it is really just a constant; the only constant that would make q(x) a probability distribution (so that adding all values of q(x) will result in unity). Multiplying by any other constant would not change the relative weight of the scenario, but it would fail to make the resulting q(x) a probability measure.

[3] However, this success might be explained by most market participants' continued *belief* in the asset pricing models; the model seems to be subject to what Taleb calls the "narrative fallacy"; explaining things retrospectively and so conveying no empirical content.

[4] See Gerber and Shiu's "Option Pricing by Esscher Transforms" (TSA, 1994).

For another way to see the independence, consider the relative weight of any two outcomes (x_j and x_k). The ratio of their weights is a simple function of their difference; any mention of the other scenarios' results is eliminated.

MCEV (as we have defined it thus far) is not only overly reductionistic (by leaving out the context of the observer); by ignoring risk, it can be somewhat dangerous. Adding some business (an asset purchased with the premium collected when a policy is sold less the liability that policy represents) with its own MCEV will only add to the whole firm's MCEV. This will be how the performance is measured even if this new bit of business increases risk substantially.

Now if two prospects A and B have the same MCEV that is greater than zero, they will both add the same amount to the firm's valuation. Suppose the firm must pick either A or B to be part of its portfolio, but at the exclusion of the other. The firm would apparently (by the lights of its own performance measure) be indifferent even if A would produce an increase in the variance of the firm's total prospect and B would decrease it. One way to redress this inadequacy is to introduce capital charges that are higher for the less desirable alternative. First, we need to develop a definition of capital.

3. Enterprise Risk Management

3.1 Economic Capital as MCEV at Risk

MCEV is a member of the family of present value of future cash flow (PCFCF) measures. Over a rather short period (such as a year), the loss of MCEV (=$MCEV_0$ – $MCEV_1$) will not exceed the fixed amount A with a probability of 95 percent. The value of A that makes this true (according to the model) can be one's definition of required or economic capital.[5] Simply project the MCEV one period hence under a set of real-world scenarios, then rank the results and note the 5 percentile amount. Subtract this from today's $MCEV_0$ to get the capital requirement. Because MCEV is linearly decomposable, this value at risk measure is easily attributed to the various lines of business; one simply needs to know what scenario produced the 5 percentile result for $MCEV_1$. But the price for this clarity is a very short-sighted measure. It potentially ignores the possibility of a bad result sometime after the (one year) observation period. This may be fine if all liabilities are short-term.[6] If final profit or loss is determined within one year, then one year *is* the long term. But if they are not, it is logically possible that there are several ruin scenarios and that in each of them the $MCEV_1$ is *greater* than zero (and perhaps even greater than $MCEV_0$).

This deficiency is a general one for *all* PVFCF based capital constructs. While they are linearly decomposable (assuming assets for each part are siloed off from the others), they will miss the potential insolvency. The PVFCF can be thought of as the present value of accumulated surplus *at the end of the projection period* or "horizon." If surplus would be wiped out before reaching the horizon, PVFCF will gloss over this.

[5] Equivalently, setting A to be the actual (accounting value of) capital, one can then solve for the probability and this becomes the risk measure.

[6] Guarantees made by casualty insurers are generally short-lived at least as compared to life insurance guarantees. Even so, the impact of possibly making dramatic changes to renewal premiums is an issue for them and not one that is separate from the issue of their long-run viability.

We can repair this if we redefine the term "cash flow" to pick up any intervening insolvency (such as "cash flow is set to zero after any ruin takes place"). But then the linear decomposability of PVFCF is destroyed. What is more important, the insinuation of a solvency monitor into the projection process changes the very nature of the measure from one that is PVFCF based to one involving AAR.

Therefore, I will treat all risk measures based on purely PVFCF ideas such as value at risk and its kin (MCEV at risk and TVAR) as non-starters for economic capital and consider only those that are more attuned to the longer term survival of the enterprise. Consequently, the risk measure will need to involve the more encompassing concept of AAR.

3.2 Economic Capital as the Amount Needed to Survive (with a high probability)

A method for evaluating the distribution of AAR is the conditional tail expectation (CTE). This is the basic method for the NAIC's RBC C3 Phase I and II, only these are performed on certain isolated blocks of business and then later incorporated into the total company capital requirement. For enterprise risk management, the intent is to apply the method to all blocks of business at once.[7]

First, rank and enumerate the n scenarios according to their AARs in descending order. Next choose α to be rather small—10 percent, 5 percent or 1 percent so that $\alpha*n$ is an integer. Then define the sub-set of scenarios (call it the tail), as those scenarios with AAR among the highest $\alpha*n$ values. We will refer this subset as $T_{1-\alpha}$.

$$T_{1-\alpha} = \{j \;;\; AAR_j \text{ is in the top } \alpha\}$$

Note that the set $T_{1-\alpha}$ is set globally—that is, for all product lines aggregated. Risk is viewed from the whole enterprise's perspective. If for some scenarios, the AAR is positive, then there is some risk that capital should be present to prevent insolvency.[8] Enterprise risk is revealed as a *relationship* between the enterprise, its identified lines of business and the economic environment in which the lines of business generate future returns.

In order to measure the probability distribution of the AAR, we use the CTE.

$$CTE = \sum_{j \in T} AAR_j * \Pr ob(j) \Big/ \sum_{j \in T} \Pr ob(j)$$

[7] Note that this also entails the sharing of asset liquidity: if one line of business has to meet a demand for cash, all asset cash flows from across the legal entity are available to meet it; this reduces the need to forcibly liquidate assets.

[8] By the term "insolvency" I refer simply to the condition in which the AAR for at least one of the scenarios is greater then current capital (or if, equivalently if AAR is positive when reflecting all assets and all liabilities together). I do not mean to include under it any regulatory action that may transpire well before the assets are exhausted in the worst-case scenario; that would be probably too conservative if not redundant. If, for example, a life insurer had risk based capital at the company action level because of C3 Phase I, there may arguably be no more than a 3 percent or so probability that a complete regulatory takeover of the company is necessary. Of course, the regulator in such a situation would be obligated to at least constrain the actions of the firm until it is rehabilitated.

This is the average value of additional assets required given that the scenario is among the worst cases defined by the set $T_{1-\alpha}$.

Since the probability of a given scenario occurring, Prob(j), is equal to 1/n, we have:

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

By the term "line of business" or "product line" we refer to not only a collection of similar policies, but also to the company's behavior regarding those policies: marketing, investment and renewal rate strategies, along with price structure, etc. So, for example, we might regard a potentially unhedged block of policies as distinct from the same policies with hedges associated with them. Now when there are m existing and 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,* $T_{1-\alpha}$*, the same*. The subset, $T_{1-\alpha}$, represents the tail of the distribution of AAR for the *whole* enterprise rather than that of its *parts*. This definition allows for the following:

$$CTE = \sum_{k=1}^{m} cte_k$$

Where:

$$cte_k = \sum_{j \in T} \frac{aar_{k,j}}{n * \alpha}$$

is the CTE for the k^{th} product or sub-line. Note that $aar_{k,j}$ is *not* the additional assets required for the line on a stand-alone basis. It is the present value of the line's surplus at a particular time for that scenario; the scenario's *low-point* that is calculated in reference to the whole enterprise. Like CTE, a sub-line's cte can be negative or positive. If it is less than zero, that means that the sub-line *contributes* capital; whatever are the "risks" that might be said to "reside" within that sub-line, they are not presently (big enough) risks to the enterprise. On the other hand, if cte_k is greater than zero, the sub-line *uses* capital. Here we have a natural, and highly meaningful, capital allocation. Once the level of α is set, $CTE_{1-\alpha}$ serves as a form of economic capital. . I will use the term "EC" to denote this version of economic capital (and its upcoming variations) to distinguish it from non-starters.

Economic capital as defined above is not constant. It is highly variable with things such as

a. the mix of business,
b. the strategies employed in the administration of the products in the mix of business,
c. beliefs or assumptions regarding mortality rates, lapse rates and other policyholder behavior, and
d. economic factors like interest rates, equity levels, volatility, etc.

Sudden changes in the level of interest rates or the level of the equities markets will potentially cause a significant change in the value of EC. Each product line has a call on all the assets of the enterprise; so besides the assets allocated to it, each line has a call option with corporate as the counterparty.[9]

3.2.1 A Possible Refinement of the EC Concept

Some will contend that the above definition of EC does not go far enough. The bad results (those falling in the "tail") are all given the same weight (i.e., the probability of 1/(αn)) in determining EC. Some believe that the probability weights ought to increase from scenario to scenario as the results get worse; reflecting the idea that the seriousness of the scenario (the pain associated with experiencing it) increases more than linearly with the size of the loss. The transformed probability measure technique enforces this idea.[10]

First, transform the probability of each real-world scenario (1/n) to one in which unfavorable results get more relative weight than do the favorable results. The Esscher Transform (with negative h) discussed earlier is one way to do this. If one is willing to restrict one's view of the possible underlying probability distributions to those for which the Esscher Transform produces (an approximate) market value, one gets in the bargain a tidy bridge back from the real world to the risk neutral world where MCEV resides. Of course, if one begins with a risk neutral set of scenarios each with probability 1/n, one need not bother with the Esscher Transform since the over (under) weighting of bad (and good) scenario results was already baked-in.

The concept of probability weighting has merit, but the Esscher Transform can be improved upon. While it does increase the weight of bad results, it fails to do so in a way that reveals enterprise-specific risk.[11] Later, I will present another transformation that will be a better guardian of the firm's survival.

3.3 Capital Credits and Charges in Performance Measurement

A product line that uses capital should be charged something related to the cost to the enterprise of providing that capital. This would reduce that line's adjusted return. If the product is still profitable enough to compete with the other line's products, then no action would need to be taken. But if the capital charge reduced profitability enough, this would encourage (a combination of) higher prices, lower commissions or a change in some other company behavior with regard to that product. That action would doubtlessly reduce sales (if it did not, one should wonder why the change wasn't made sooner), but that product's reduced profile would be a positive for the enterprise as a whole; for that reason, it would be a sensible decision. Attempts to maximize the adjusted performance measure would promote the most sensible action; this is the key property of any good performance measure.

---

[9] See Mango (2005) for a more thorough discussion.

[10] See Venter (2006) for a more thorough discussion.

[11] It also, ironically, fails as basis for (expected) valuation. For a given (negative) value of the parameter h, there exists a non-negative prospect A, and another prospect B, such that 1) B dominates A and so is always preferable to A and yet 2) B has a lower expected value (with respect to the Esscher transformed probability measure) than does A. So, valuation using the Esscher Transform produces unacceptable results, in general.

On the other hand, a product line that contributes capital should be credited with an amount that is commensurate with the cost of capital. Its return for a given period would be enhanced by this credit. More exposure to that line's product may be beneficial to the entire enterprise because its returns are positive in the scenario set $T_{1-\alpha}$ (otherwise the product's cte_k would not be negative). More of that product in the overall mix would further reduce the overall EC which means less capital would need to be set aside. Provided it were sufficiently profitable relative to the other products, the improved performance (current period's profit adjusted for the capital credit) would provide management with the proper signal to take steps that encourage growth in that line relative to the other lines of business. Price sensitivities (to the degree they're known) could be exploited (by cutting the price or increasing commissions) so that the increased sales would materialize.

Without the above-mentioned capital definition, allocation process and performance measure adjustment for the cost of capital to govern its growth, a fast growing product line might get to be so large as to present a risk to the enterprise that is too great relative to its return prospect. It could even become so large as to present a solvency risk (and this kind of risk could arguably be rejected for any prospect of return). The proposed framework controls this by signaling what price increases and decreases increase the whole company's risk-adjusted return.

Suppose we identify a product line that contributes EC and there we take action intended to grow it relative to the others. Successful growth of such a product line can ultimately cause the scenario set $T_{1-\alpha}$ to change so that new risks emerge and old risks become more tolerable. After a point, further growth of the product line will become counterproductive to the enterprise, and so other lines of business might then warrant attention. In other words, the framework proposed encourages prudent growth and balance of product mix by dynamically adjusting incentives.

3.4 Additional Capital Charges

Because a sudden change in the economic environment may cause a product line's cte to increase significantly, and because the model may underestimate such an event's probability of occurrence, some additional charge is appropriate. Because this is prompted by residual doubt regarding our model, we have no way of formally quantifying the additional charge. Management judgment will be needed to establish the charge.

Recall the sensitivity tests; these are the additional scenario sets that reflect sudden large environmental changes. One way to develop the charge is to calculate the change in EC that would result under each of the sensitivities. This EC will be decomposed into each line of business' cte as before; some positive and some negative. For each sensitivity and line of business, allocate the price of an option that would hedge the overall EC to those lines that would have caused the *increase* by the degree to which each such line would have caused it. A positive increase in cte would lead to an additional charge.

The line of business may even choose to invest in these options directly; having been charged for the option it will be better to be able to get credit for any gains that might eventuate. Also note that if the hedges were purchased as one of the line's allocated assets, this would obviously reduce the line's cte. This would, of course, reduce or eliminate the amount of additional hedges to purchase.

A *decrease* in a line's cte from sensitivity test would *not* result in an additional credit, however. Such a line is not providing the option to the other lines and should not be compelled to do so; that is corporate's responsibility just as it is corporate's decision of whether to purchase a hedging instrument from the capital markets or not.

3.5 Allocation of Free Capital

After allocating the capital to the lines of business and either hedging or holding back enough assets to cover any potential calls from the lines, what remains is free capital. The ultimate goal is to grow free capital at the highest possible rate. One natural way is to grow the businesses in which the enterprise has particular expertise. Other alternatives may present themselves too, and this definition of capital might help identify them. Still, internal organic growth will likely be seen as the best approach and the decision to allocate free surplus to the various products will be influenced by the prospective returns those products will bring to the enterprise. How new sales of those products will shape the EC will need to be considered. Products that will benefit from EC credits will have some advantage over users of capital; all else being equal, capital providers will attract more free capital than will capital consumers.

3.6 Economic Capital and the "Holistic Approach"

We have briefly described how an EC allocation framework can assist in managing risks by influencing the pricing of products so that a good mix of business results and by informing us as to what and how much to hedge. Of course, something more is needed. We have cogently defined what our capital requirement should be (EC), but have left as undetermined the proper interpretation of the phrase "profitable enough." The next section attempts to provide guidance.

The focus of ERM is on the corporation as a whole. It takes a rather holistic view of proper decision-making. Typical performance measurement systems have countenanced a de-centralized, perhaps entrepreneurial, approach to decision-making. The designated lines of business are in a competition for growth capital. Performance measures are geared toward treating them as independent and as pursuing their individual interest. The goal of each part is to maximize the value of its particular "silo."

The addition of EC charges and credits to these performance measures adds a holistic dimension, but it misses the mark of holism in the following respect. Separate lines of business may be encouraged to over-hedge (from the enterprise's perspective) their particular cte positions. For example, suppose Line A would be hurt by a drop in the level of interest rates because its future profits would then be decreased. Line A might find interest rate floors valuable to it and contend that they should be purchased; suppose, in addition, that this purchase will increase EC but reduce Line A's component cte. At the same time suppose the equally risk-averse Line B's future profits would be helped by a drop in interest rates but hurt by an increase in them. It might find interest rate caps valuable (to itself) and contend that it should be allowed (as a matter of fairness) to purchase some. However it might be economically better for the enterprise to purchase neither caps nor floors. EC is a step in the holistic direction, but, by itself, it will not remove silos.

Some may believe that a competitive environment emphasizing diversity of approach is vital to the firm's success even with its built-in friction costs; and that holistic considerations undermine that. But competition is compatible with holistic decision making. The grounds for the coordination of the parts are undeniably present. Even in a competitive environment, any one of the parts may need the strength of the others at some point in the future and all are strengthened by the presence of others. ERM need not require centralized decision-making. What is needed is a focus on the whole enterprise with the single goal being its betterment. The advisability of making decisions that enhance the enterprise's overall value (even if they are at the expense of one of the parts) is the core message of ERM and is embodied in the concept of Equivalent Logarithmic Utility Value (ELUV) which is explored next.

4. Enterprise Value Management

4.1 Defining Equivalent Logarithmic Utility Value

Let x_j be the present value of cash flows (revenue less benefits, expenses and taxes) from all sources under the j^{th} scenario. Unlike the calculation of EC, we include the surplus assets. We also impose *the proviso that x_j is zero if the AAR for the scenario is positive*. If the AAR for a given scenario and for the entire enterprise (*including all its surplus assets)* is positive, that means that enterprise ruin occurs under that scenario at some point in the projection period. Let $p(x_j)$ be the (actual or real world) probability of the j^{th} scenario; we are only concerned with x for which p(x) is greater than zero. We can write ELUV, provided $x_j > 0$ for all j, as:

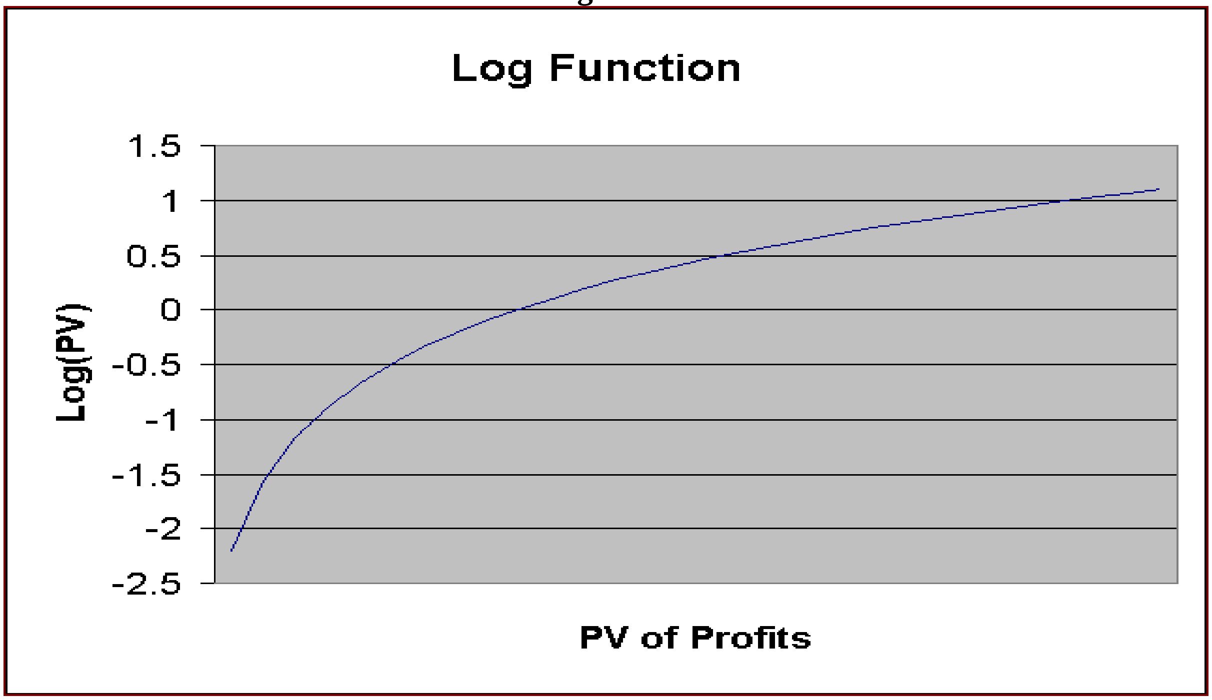
$$ELUV = u^{-1}(E[u(x)]) = e^{\sum_j \ln(x_j) * p(x_j)} = \prod_j x_j^{p(x_j)} = G$$

and ELUV equals zero, otherwise. This is the geometric mean (G) of the distribution. Note that the exponent in the second term in the above equations refers to function inversion and does not refer to the reciprocal. Also note that G is a function of all the x_j and their probabilities of occurrence; these arguments have been dropped here for the sake of brevity.

As with the definition of EC, it is important to emphasize that present value refers to the projected returns for the entire entity (enterprise); that is, for all lines of business combined along with surplus assets. As with the EC calculation, "j" refers to the scenario, only with ELUV we take into account all scenario results (not just the worst cases). However, if any one scenario produces a non-positive (i.e., negative or zero) present value, ELUV is evaluated as zero; the lowest value it can attain. Very low singleton scenario results tend to take over the measure of the whole distribution. Such very low results signify the possibility of (measurable and believable) insolvency of the whole enterprise. Any action that could be taken that would produce positive values under all scenarios will be preferable, no matter how much it reduces the present value of the otherwise positive scenarios.

The logarithmic function serves as a utility function. ELUV is the sure-thing equivalent value of the present value of profit distribution—the sure thing result that has the same expected utility value. The shape of the logarithmic function is such that the worst scenario results are given extra weight and the best scenario results are given less relative weight.

Figure 1



Utility theory is associated with the subjective point of view. This association is historical but it is not a necessary association; it is even somewhat unfortunate and misguided. Nothing requires that the choice of the utility function reflect some personal inner feelings on the part of an individual. Nor is it required that anyone contemplate what are their inner preferences. If the choice of utility function is made rationally, it should be considered objective. For financial intermediaries, there is a strong argument that the choice of the logarithmic function is a rational and objective one. The reasoning behind the choice of the logarithmic function is discussed in some mathematical detail in Appendix 1. There we demonstrate that the strategy that maximizes the expected growth rate of wealth is the same one that maximizes expected utility if and only if the utility function is the logarithmic function (in information theory, this result is known as the Kelly Criterion).[12] Its choice also countenances the proper attitude toward the policyholder by demanding solvency risk be fully addressed before anything else.

[12] Also see *Fortune's Formula* (2005) by William Poundstone for an interesting, popular treatment of this topic.

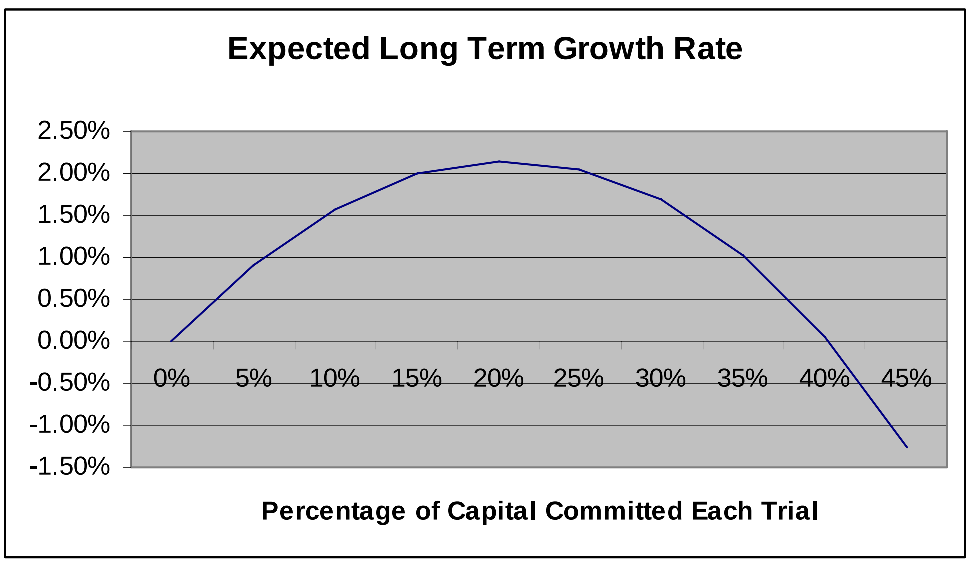
4.2 Expected Growth Rate

Consider this very simple and (hopefully) clear example. You begin with an amount of cash representing your *entire* wealth. You have the opportunity to engage in repeated trials of a game in which there is a 60 percent chance of doubling the amount committed to the trial and a 40 percent possibility of losing the amount bet. You can sign up for a series of 100 trials. Your fortune will grow or decline with each trial depending on how much of it you wager. Your goal is to maximize wealth.

This prospect has an expected ROE of 20 percent per trial, but of course it is quite risky. If one were to risk 100 percent of one's stake each time, this has the highest expected terminal wealth; an extremely small chance (60 percent multiplied by itself 100 times) of an extremely large terminal amount (2 multiplied by itself 100 times the initial stake). But one would ultimately lose everything with near certainty. Obviously the strategy of maximizing expected terminal wealth is not a good one; it is in fact the worst.

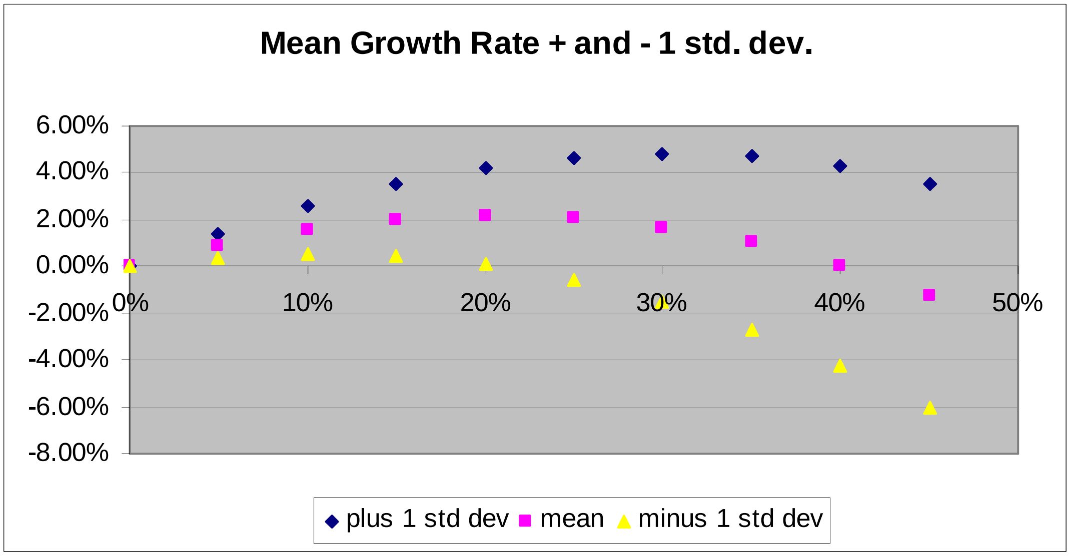
The following graph shows the *expected rate of growth* assuming you wager x percent of existing cash on each trial (we show x ranging from 0 to 45). These are long-term growth rates, so small differences among them translate into huge differences in terminal wealth.

Figure 2



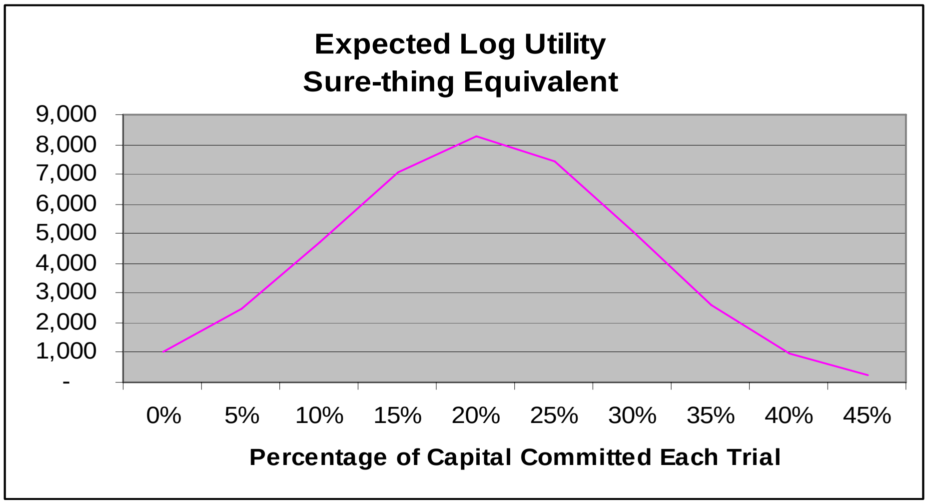
First note that if nothing is wagered, nothing is ever gained; the expected growth rate is 0 percent. Second, one should expect to lose in the long run at wagers above 40 percent. Third, the expected growth rate is maximized at the 20 percent level where the long-term expected growth rate per trial is 2.15 percent; and this is where logarithmic utility of terminal wealth is maximized. It is fair to note that at the 20 percent level, there is a 16 percent chance the long-term growth rate will be negative and a 16 percent chance it will exceed 4 percent. Figure 3 shows the mean growth rate plus and minus 1 standard deviation.

Figure 3



Now consider figure 4 below. This shows the sure-thing equivalent for the expected utility of terminal wealth after the 100 trials assuming the initial stake was $1,000. Any two probability distributions have the same value if their respective expected utilities are equal. The "sure-thing equivalent" refers to the distribution with the same expected utility and which has a zero variance. This is our definition of ELUV. Specifically, for logarithmic utility, it is the geometric mean. In this example, the 20 percent level produces an ending ELUV of over eight times the original stake. The terminal wealth has a 16 percent chance of falling below the initial stake, and there is a 16 percent chance it will exceed 56 times the initial stake. There is also a 0 percent chance of total ruin. Our example is a very simple one with only two payoffs but this property holds in any situation in which the payoffs and their probabilities are understood. See Appendix 1.

Figure 4



4.3 Solvency Risk (And Another Refinement to EC)

The second reason that logarithmic utility is a compelling rational choice is that it places a high value on avoiding very bad results. If some risk aspect of the whole business was severe enough to pull the present value of future cash flows (x_j) down to less than or equal to zero for some scenario j (indicating eventual insolvency under that scenario), logarithmic utility will penalize that outcome to such an extent that *whatever* action that would bring that scenario's outcome up to a positive value (without bringing another scenario's down to zero) would improve ELUV to something higher than zero and so would be preferred to doing nothing.

Let us view ELUV from the following angle. As was mentioned earlier, market value can be written as the expectation with respect to an Esscher Transformed probability measure but only under some very restrictive conditions. For any finite positive probability distribution, its ELUV can be written as the expected value under a specifically transformed probability measure.

Let:

$$q(x) = p(x) * \frac{(G + a(x))}{x} \quad \text{for } x > 0,$$

where p(x) is the real-world probability and:

$$a(x) = (x - E[x]) * \left(\frac{G * E[1/x] - 1}{E[x] * E[1/x] - 1} \right)$$

Then:

$$ELUV = \sum_j (x_j * q(x_j))$$

A mathematical proof of this is given in Appendix 2. We will refer to this transformation as the Geometric Mean Transform.

If there are n real-world scenarios, we distribute our belief in them equally. Therefore for each scenario, j (and outcome x_j), we have $p(x_j) = 1/n$. Now, the above transformation places a weight on the scenario equal to:

$$Weight(j) = \frac{(G + a(x_j))}{x_j}$$

Remember that the components of this probability transformation are calculated from a Monte Carlo projection that includes all liabilities and assets (including surplus).

The striking feature of this transformation is that, at each value of x_j, *$q(x_j)$ depends on all the other possible values in the distribution.* This is in contrast to the reductionistic Esscher Transform. Take any two scenarios (j and k) and their results (x_j and x_k). The relative weight (as judged by the ratio of their respective weights) remains dependent on all of the other scenario results.

Consider the sequence of probability distributions in which the lowest outcome (say x_k) gets closer to zero (while the others are unchanged), $q(x_k)$ will get closer to 100 percent (and the others will approach zero). In the limit, we can extend the domain of our transformation to include distributions that have some chance of a zero outcome as follows:

$q(x_j) = 1$ if $x_j = 0$ and in which case, $q(x_k) = 0$ if $x_k > 0$ (if k does not equal j).

Now $p(x_j)$ and $q(x_j)$ are (dual) probabilities of scenario j occurring; so we can write them as p(j) and q(j) (although we need the information $x_1, x_2, \ldots, x_n$ to link p and q). If we use the risk-adjusted q(j) in conjunction with the EC of Section 3, how much more capital would be required?

Let us examine two polar cases. First consider the situation in which the variance of the distribution is small. This is indicative of a well balanced company not taking on much risk. By definition, $x_j - E[x]$ is small. Also, G is close to E[x], so $a(x_j)$ is close to $x_j - E[x]$. Very small risk adjustments to $p(x_j)$ would take place, and so there would be little if no change from EC based on unadjusted probabilities. EC would tend to be minimal (even negative) in this situation.

In the second polar case there is a positive probability of ruin (including surplus assets).[13] Here, the answer to how much capital is needed is completely focused on the ruin scenario. There would be maximal weighting on the ruin scenario and EC would be the amount needed to stave off ruin. This particular risk-adjusting probability transformation will not countenance a firm's acceptance of foreseeable ruin without requiring that considerable effort be made to redress that possibility. The fallibility of projection models today is undoubtedly in their *underestimation* of the chance and severity of economic events (although if modeling became ultra-conservative the opposite might become the case). So (presently, at least) if the model is measuring a positive probability of ruin, one would be justified in acting to eliminate that chance. EC based on the above Geometric Mean Transform would compel that action.

The appropriate responsiveness of this transformation to the bad case scenarios (in conjunction with the EC concept) also ameliorates the problem of picking the tail size (over which the CTE is taken) arbitrarily. The choice of α will become less important (and in the extreme polar case it would be unimportant).

Taking this a step further provides partial guidance with regard to the skeptical stance we adopt in viewing the model risk (discussed earlier in Section 1.2). This transformation contours the enterprise's stance toward the issue of hedging these key risks by providing a measurement of just how key they are for a particular enterprise. If the impact of a given source of risk is not very large, it will not attract much attention. If, on the other hand, the impact is quite significant for the enterprise, then we will want to pay close attention to the possibility that we have underestimated its probability of occurrence ($p(x_j) = 1/n$) and simply hedge against it *as if* the probability were the larger $q(x_j)$. This guidance is only partial because the probability

[13] Instead of the term "ruin," some proponents of Option Pricing Theory may prefer the expression "exercising one's put option to default."

distribution may ignore even worse possibilities by assigning to them a probability of zero. There will be some residual skepticism with regard to possibilities completely missed by the model.

To illustrate, suppose you can identify four different events that would each significantly hurt your firm. Say you also judged the probability of each to be 10 percent; each has a 0 percent chance of occurring more than once; and they were independent of each other. This implies that there is a 10^{-4} chance of all four occurring in the same scenario. Now, if under such a scenario your firm (A) would survive, then it is likely that you work for a firm with a very strong capital structure. Consequently, you have less reason to second guess your "best guess" model than would one working for a firm (B) that would not survive the four horsemen scenario. How much less? While for both the chance of the scenario occurring is believed to be 10^{-4}, for firm B the transformed probability, q(x), is 1 and it is less than 1 for firm A. How much less than 1 depends on A's (whole) distribution. From a solvency perspective, the further from 1 it is, the more the company can safely choose to ignore this compound risk; it has sufficient capital. It is freer to consider the upside of its prospect. Firm B needs to concentrate on the four horsemen scenario and act remedially.

4.4 Some Clarification of the Policy for Addressing Key Risks

In Section 1.2 we prescribed the development of a policy for addressing key risks. At this point, we can suggest some specification of this.

First, let us settle on the definition of economic capital (EC) as the one described in Section 3.2 and based on the conditional tail expectation (CTE) of the additional assets required when not including surplus assets. Additionally, in calculating EC, we will not employ the real-world scenario probabilities, but rather we will employ the transformed probabilities obtained by applying the Geometric Mean Transform. Second, recall that the key sensitivities take the form of a sudden, immediate shock after which a new scenario set is generated as usual but from that point onward. A key sensitivity is a scenario set derived from the real-world scenario set.

One way to specify the policy is to require remedial action if EC, recalculated with a given key sensitivity, would come too close to the accounting value of actual surplus. An alternative policy would require remedial action if the key sensitivity would lower ELUV by some determinable amount. Either policy form would properly focus action toward enterprise-specific risk.

Another way to reflect the constraints the key sensitivities would impose is to bring them inside the model. Because the Geometric Mean Transform weights the scenario results in a way that is highly meaningful for ERM, one could proceed as follows:

1. Denote the set of baseline stochastic scenarios as $SCEN_0$.
2. Enumerate the key sensitivities scenario sets $SENS_j$ for j from 1 through m. Recall that each scenario in a sensitivity-set will correspond to a scenario in $SCEN_0$; they will differ by the impact of the immediate occurrence of the sensitivity event.

3. For the first sensitivity event, *hazard* a guess as to its probability of occurring; call it P_1.
4. Form a new set of scenarios $SCEN_1$ by multiplying the probability of each scenario in $SCEN_0$ by $(1-P_1)$, multiplying the probability of each scenario in $SENS_1$ by P_1, and then forming the union of the two sets. One will now have 2n scenarios.
5. Proceed with other sensitivity sets in order repeating steps 3 and 4. SCENj is formed by multiplying probability of each scenario in $SCEN_{j-1}$ by $1-P_j$, multiplying the probability of each scenario in $SENS_j$ by P_j, and then forming the union. For m separate sensitivities we will end up with $SCEN_m$ consisting of $2^m n$ scenarios.
6. Evaluate the Geometric Mean Transform for $SCEN_m$ and then evaluate EC as the $CTE_{1-\alpha}$ with respect to the *transformed* probabilities.

Suppose that the worst case entails the occurrence of one or more of the key risk events. It will get the most weight in the EC. If the worst case result is bad enough or even solvency threatening, then that scenario will dominate the EC measure and the cost of capital will be quite high. Action taken to reduce the capital requirement will be profitable under the cost-of-capital-adjusted performance measure. The fact that we used a fairly wild guess will not be so important. On the other hand, if the worst case is not awful, measures to improve will be welcome provided they do not cost too much, but again the fact we made a wild guess is not harmful to us (it may convince us to toss out that particular sensitivity).

Actually, within the ELUV framework, EC is not really needed at all. Indeed it is silly to evaluate the Geometric Mean Transformed EC when all that is needed is G, the geometric mean which happens to equal ELUV. ELUV is a robustly risk-adjusted performance measure that subsumes the concept of economic capital. Within the framework, any decision would be made in light of the attempt to increase ELUV. The decision to, for example, purchase derivatives to mitigate the largest key risk will be like any other decision to purchase an asset: it will need to improve ELUV the most among the identified alternatives. Under traditional performance measures, risk is thought of as a side-constraint (with capital requirements being the constraint). Under the ELUV framework, risk is an important dimension of value. As a conceptual framework, ELUV is as objective as any other (because the choice of utility function is prescribed)[14], consummately coherent and normatively powerful.[15]

4.5 Relationship to Fair Value and Option Pricing Theory

ELUV is an ideal basis for an accounting system or performance measurement. Today, a great deal of attention is paid to a similar concept: that of fair value, which has come to mean for many, market consistent value. Indeed, at first blush, our account of ELUV may seem to be merely the market consistent embedded value without the market consistency. Under the fair value concept, anything that has a well established market is to be valued at the current market

[14] The forming of degrees of partial belief about the future (i.e., the scenarios and behavioral assumptions) is surely quite subjective and fallible, but that is a problem not peculiar to the ELUV; rather it is a universal issue.

[15] A deeper discussion of this valuation paradigm can be found in “Value and Actuation” (2006 ERM Symposium) by C. Perrin.

price. Anything that doesn't have a market is to be valued in a manner that is in the spirit of a market price; even though that is a vague notion. The fair value of any composite is the net sum of the fair values of its parts. It too, is objective and coherent. One might see fair value as the ELUV in which we replace the logarithmic-utility function with the utility function of the "Market" and replace the model with the market's beliefs regarding the scenarios, their probabilities, and the resulting behavior in reaction to them. Being a composite of many rational (enough) traders, the market is treated as some kind of rational being (and this treatment might exemplify the fallacy of composition). Under the theory, this rational being has a (composite) utility function. Ascertaining this is a daunting challenge. Fortunately, the market's utility function need not be observed; (risk-neutral) Asset Pricing Theory is the finesse. But were it observed, one would find that it lacks the normative rationality of logarithmic utility. In addition, one would realize that the coherence of the market is not stable over time; rather, the market's unobserved utility function is likely to be a capricious one.

This is not to say that Asset Pricing Theory is irrelevant; it appears to be a reliable model for market prices which, of course, represent important facts. But price is not the same thing as value. The Asset Pricing Theory may be put to good use in the modeling of future prices of securities that one may decide to buy or sell; indeed, a decrease in something's price (depending on the reasons for the reduction) ought to *enhance* the value of (the decision to buy) it relative to alternatives. There are better over-arching frameworks of measuring the worth of the enterprise (and decisions made to enhance the value of the enterprise) than MCEV.

It is also not to say that the collective wisdom of the "market" ought to be ignored, in general (though that is possibly true). Much has been made of the wisdom of crowds lately[16]; that the diversity of experience and approach to solving a problem, upon the proper aggregation of those solutions, will provide more accurate answers than the word of one elite expert. Besides the framework (maximization of expected utility or risk neutral valuation; take your pick), the other aspect of 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. This description is the true heart of the matter; the most critical for success and the most difficult to achieve. The "market" may be seen as having degrees of belief that are superior to what we could come up with on our own; and if we could discern them (i.e., separate them from the market's supposed utility function), this might be the best way to establish our own. It might be that this method would be a good adaptive approach. Fair value is then, according to its proponents, the best answer to the problem of valuation.

There are two arguments to be made. The first is that this is all academic. Unfortunately, we cannot separate the market's belief from its utility function (even if we assume the form of the utility function). Even if its beliefs are an improvement over the traditional methods we use, it will not benefit us because the market's beliefs are inscrutable. Indeed, when we are told to *"extend"* the market value concept to those objects that lack a market, the only thing transported or preserved is the risk-neutral valuation technique which is nothing more than a trick that allows one to view the expectation as being the price. The very thing that would be profoundly useful to have (the crowd's wisdom regarding the probabilities of possible future events and how policyholders will react in those various possible worlds) is not brought over. The hull has been

[16] See *The Wisdom of Crowds* (2004) by James Surowiecki.

taken as valuable, but the kernel has been left behind. What we then have is still a subjective judgment call made by the enterprise; a single modeler. So, the result of extending market valuation is really as subjective as any other.

Secondly, the wisdom of crowds emerges only when certain conditions are met. Diversity of opinion is one (without it everyone's opinion would be the same as the crowd's). Independence is another. As individual agents in the economy, we fix these beliefs through inductive methods where possible and otherwise substitute judgment. Although they are completely fallible, we have no choice but to use the most reliable methods available to us (this is the key role of the actuarial profession). If wisdom is an emergent property of the market, it is due to the hard work of its participants developing their respective beliefs *independently* of the other market participants. Looking to the market as a short-cut for determining something's value might work as long as no one else also took the same short-cut. Like risk-free arbitrage opportunities, that would not last and they'd find themselves aboard a ship of fools. As much as we'd like to move above the fray, we are inextricably in it.

In times of great market stress, it is said that "all the correlations go toward 100 percent"; that is when market price (qua measurement of true value) is most unreliable. Market price becomes unreliable as a measure of value precisely when market participants cease working independently and attempt to use market price as a short-cut.

Fair or market value is best regarded as providing information that is relevant to a decision. But the decision requires more; namely that the information be processed by an intermediating agent with a goal to increase his or her valuation. Fair value, as a framework for valuing, confuses information for value. The appeal of Asset Pricing Theory as the valuation framework is its objectivity. But an entity's value necessarily depends on a context. The fundamental mistakes made in market consistent valuation are twofold. First, it assumes that the object of valuation (for which there is no real market) is, at base, the same as something else (for which there is a market); otherwise "extending" makes no sense. But the lack of a market is more a sign of dissimilarity than of similarity. Second, it assumes buyers are alike (not merely that they all have complete information); it ignores the fact that value is value *for* the potential buyer. One potential buyer's situation may be that the addition of the object would constitute an improvement, whereas for another, its addition may constitute an injury.

The prudent course is to proceed with valuation in a rational, disciplined way that accounts for *wholes*; with discipline, the subjective process approaches objectivity. That discipline is what the utility framework enables without trying to side-step the hard work that needs to be done. Contrarily, using the market valuation technique in this way only short-circuits the process and begs the question of value.

4.6 Operational Issues for ELUV

There are a few operational issues to mention briefly. We will finish our discussion of the ELUV framework with them.

4.6.1 Attribution to Lines of Business

An important point to be made regarding the use of utility functions in general is that they are only meaningfully applied to wholes. The expected utility method described here does not lend itself to 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. Embedded value based on "market-consistent assumptions," on the other hand, is linearly decomposable into parts because of the reductive nature of the asset (or option) pricing model. In order to attribute performance under ELUV, one would re-evaluate the whole enterprise assuming that a particular part was not there. The resulting shortfall would then be the (marginal) value added of that part. The increase in this (sub-) measure would indicate the period's contribution made by the line of business to the whole company; it should be possible in this way to maintain a sense of competitive sharpness within the various lines of business.

4.6.2 Planning

Pricing with ELUV requires the stochastic modeling of hypothetical new product sales; this is nothing new. What is new is that we also need the background information that is the modeling of all old business as a context for the decision to write the new business. How new sales of all the various and seemingly unrelated products fit with each other (as well as with the old) will influence the decision. Further, alternative price levels and investment strategies should be considered. One aspect of this is some approximation (provided most likely by the marketing division) of how a change in price would affect sales; that is, a demand curve for each product. From all this information there would be an attempt to set all prices at once that together will (roughly) maximize ELUV. The product of this would be not only an enterprise-wide pricing matrix, but also the marketing and financial plan coming together in an ultimately integrated way.

Now in comparing actual results to planned, there is a consequence that might seem controversial although it is not unintended. Meeting a sales goal is still a very good result, but substantially beating one without having adjusted the demand schedule (i.e. without having adjusted prices to maximize ELUV in light of new information) may not be as welcome. Undoubtedly, determining demand curves is a matter of guess-work; it should be viewed as a learning process that improves with experience. In an empirical trial and error approach, the more information produced the better. One needs to consider the impact on the vitality of the distribution channel of what may seem to be frequent pricing changes in certain markets. If that is an issue, the risk of a false move could prevent experimentation, but the resulting lack of information would diminish value unless the initial guess happened to be perfect.

4.6.3 Optimal Capital

In Section 3 we described economic capital as the (smallest) collection of additional assets needed by a collection of liabilities (and assets "backing" them) to ensure solvency; this is after any hedging induced by the sensitivity tests. EC is what we *must* hold to meet the guarantees to the stakeholder group that own policies. Having established EC for the sake of the policyholder group, we can now turn to the other stakeholder group; the equity holders.

For the same collection of liabilities and assets backing them (including hedges of the key risks identified by the sensitivity tests and the additional assets required for EC), we can ask what additional assets *should* we hold. The answer is that for a given collection of liabilities, we ought to hold the additional assets that maximize the ELUV. This optimal capital (OC) will be borrowed in concept whether it is provided by existing surplus or not. The cost of borrowing is thus a limiting factor for OC; i.e., there is such a thing as too much capital. Any surplus beyond EC and OC is for developing new business or paying equity-holder dividends.

But one must be careful here. OC has the hallmarks of a dangerous idea. Maximization relies on the Monte Carlo model which will be fallible, and so, this reliance will be impossible to completely justify. At the very least, measures to survive disaster scenarios (the key sensitivities) discussed previously in Sections 1.2 or 4.4 should be taken prior to the establishment of ELUV or EC. EC, in turn, needs to be established prior to OC. Sound ERM should make the notion of OC more concrete.

4.6.4 Hedging ELUV

Sudden changes in the economic environment will change ELUV. The sensitivity tests could be used to set an objective such as "if a 20 percent drop in the S&P500 Index occurs, we do not want more than a 5 percent drop in ELUV." If the sensitivity test reveals that a 10 percent drop in ELUV would result, the company could determine a hedge that would reduce the drop to the desired 5 percent. Of course, putting on such a hedge may itself reduce ELUV significantly, so a judgment as to its suitability would need to be made outside of utility theory. If a suitable hedge cannot be found, this may indicate a change in marketing strategy due to a capacity constraint. Now, what if a hedge could be found that mitigated the effect on ELUV of a sudden drop in the stock market, but reduced ELUV only moderately. Would not the ELUV comparison lead us to decide against such a hedge?

The ELUV framework is inspired by the Principle of Maximizing Expected Utility (PMEU). Ideally, PMEU is (intended to be) the last word. Once the utility function is optimized, there is no more to be done but react as the future unfolds; that is, re-optimize at points in the future. If two alternative probability distributions presently have the same ELUV, we are supposed to be indifferent; to not prefer one over the other. Buying a financial derivative to hedge, for example, will be judged ideally like any other potential decision: if it increases ELUV, then we would buy; otherwise we would not.

But PMEU is dependant upon the underlying probability distributions as though they were known. A probability distribution is a system of partial degrees of belief that are internally consistent or coherent.[17] Coherence assures that we cannot be victimized by the "money pump" (a series of bets in which we are guaranteed to lose). Once we come to have full belief in the distribution, we will act on it and PMEU is an absolutely sound guide for ensuring the internal consistency. It is one thing to fully believe a coin is fair and to act on the premise that, when flipped, it will come up heads with a 50 percent chance. But it is something else altogether to believe fully that a particular probability distribution governs changes in the stock market.

---

[17] See F. P. Ramsey's "Truth and Probability."

We need to have a rather high degree of belief in our model; otherwise we need to work on it until we do. But what should we do if we cannot (being honest with ourselves) *completely* specify (and trust) the requisite degrees of partial belief? There is no sound foundation for us to *fully* believe *any* probability distribution no matter how scientifically that system of degrees of partial belief was formed.

Holism requires that we look "at the big picture." This applies not only for viewing the parts of the enterprise in relation to the whole, but also for *how* we fix our beliefs and the stance we take toward the residual doubt we have about them (how we face Knightian uncertainty). It is entirely consistent with holism to consider our limitations in this regard as an aspect of the whole world whose value we seek to increase; that is, to make room for a skeptical perspective. There is an epistemological dimension to value.[18]

Using inductive methods augmented with common sense judgment have shown themselves to be good (though fallible) habits of fixing belief. At the same time, identifying key sensitivities without trying to specify their probability is a way to acknowledge our limited abilities to know (particularly in very complicated settings). But understand that this is also fallible since we can easily have 1) left something off our list of key risks or 2) paid too much attention to those items on the list.

In using ELUV as a performance measure, we will necessarily settle for serviceable degrees of belief and attempt to hedge against events that we acknowledge would hurt us, but hedge only to the extent they would hurt us. So, in allowing for the skeptical viewpoint, we may well judge two possible distributions with the same ELUV differently. That is, we might have second-order preferences for the one with less volatility. ELUV is an organizing principle; an ideal. But as with any ideal, it is subject to practical considerations.

5. Conclusion

We have described two ideas that try to achieve the same goal: to contour decision making in a way that addresses the most relevant risks facing the enterprise. EC provides a way of explicitly adjusting for such risk within an existing performance measurement framework. ELUV adjusts for risk implicitly through the logarithmic utility function. ELUV is an idealized extension of the EC concept. It more purely embodies the concept of holism because it seamlessly relates the parts of the enterprise to the whole. The EC approach (with its performance adjustments) gets closer to holism, but retains some reductionistic aspects of traditional management approaches. The ELUV approach is actually a replacement for an historical accounting method whereas EC is an incremental change to the method. Even if it constitutes an improvement, ELUV is perhaps too abstract, too difficult to implement at this time, and it may require more information processing capacity than is available today.

[18] See the discussion of Newcomb's Problem in Perrin (2006).

By being incremental, the EC framework is perhaps more practical than one focusing on ELUV. The EC framework is being made more and more concrete each day. It is informing the regulation of the insurance industry through its instantiation as a risk based capital requirement for variable annuities and is being fully implemented in some form by some insurance company management teams. The use of a transformed probability measure in conjunction with EC represents a further refinement; and we have shown that one particular transformation (the Geometric Mean Transform) gets to the heart of the matter of uncertainty in ERM by identifying enterprise-specific risk. Perhaps when the realization of the EC concept is complete, the jump to ELUV will not seem to be such a large one.

References and Further Reading

Gerber, H., and Shiu, E. 1994. "Option Pricing by Esscher Transforms." Transactions of the Society of Actuaries XLVI: 99–140.

Mango, D. 2005. "Insurance Capital as a Shared Asset. ASTIN Bulletin 35(2): 471–476.

Perrin, C. 2006. "Value and Actuation." ERM Symposium.

Peirce, C. S. 1991. *Peirce on Signs: Writings on Semiotic by Charles Sanders Peirce,* edited by James Hoopes. Chapel Hill: The University of North Carolina Press.

Poundstone, W. 2005. *Fortune's Formula: The Untold Story of the Scientific Betting System that Beat the Casinos and Wall Street.* New York: Hill and Wang.

Ramsey, F.P. 1931. Truth and Probability (1926). Published posthumously in *Foundations of Mathematics and other Logical* Essays, edited by R.B. Braithwaite, 156–198. London: Routledge and Kegan Paul. Or in *Philosophical Papers*. 1990, edited by D. H. Mellor, 52–94. Cambridge: Cambridge University Press.

Taleb, N.N. 2004. *Fooled by Randomness: The Hidden Role of Chance in Life and in the Markets,* 2nd edition. New York: Random House.

________. 2007. *The Black Swan: The Impact of the Highly Improbable*. New York: Random House.

Surowiecki, J. 2004. *The Wisdom of Crowds.* New York: Doubleday.

Venter, G.G. 2006. "Adapting Banking Models to Insurer ERM." ERM Symposium.

Appendix 1

Logarithmic Utility and Expected Long-Run Growth[19]

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 payoff 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.

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 $r_{max} > 0$. α can vary from $\frac{-1}{r_{max}}$ *to* $\frac{-1}{r_{min}}$.

What is the proper choice of 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:

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

$$\text{where } k_m = n - \sum_{j=1}^{m-1} k_j \text{ and } K_i = n - \sum_{j=1}^{i-1} k_j \qquad \text{[A1.1]}$$

$$E[w_n] = w_0 * \left(1 + \alpha * \sum r_j * p_j \right)^n$$

$$= w_0 * \left(1 + \alpha * E[r] \right)^n \qquad \text{[A1.2]}$$

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; p_j = -1 for

[19] Appendix 1 is an excerpt from Perrin (2006).

some j. The distribution of terminal wealth is such that one ends with \$0 with probability $1 - (1 - \sum_{i \neq j} p_i)^n$.

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 (plus the constant 1) 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 $K_i = n - \sum_{j=1}^{i-1} k_j$. [A1.3]

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 of the existing stake is wagered successively. If α=0, $g_n(\alpha) = 1$. If $\alpha = \frac{-1}{r_{min}}$ or $\alpha = \frac{-1}{r_{max}}$, then $g_n(\alpha) \xrightarrow[n \to \infty]{} 0$. On the domain in which we are interested, namely $\left[\frac{-1}{r_{max}} , \frac{-1}{r_{min}} \right]$ note that $g_n(\alpha) > 0$. Note, in particular, 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 $\left(\frac{-1}{r_{\max}} \quad , \quad \frac{-1}{r_{\min}} \right)$ because it approaches zero at the endpoints of the interval and is much bigger at α=0.

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

$$g_n'(\alpha) = 0 \quad \text{and} \quad g_n''(\alpha) < 0$$

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:

$$h_n(\alpha) = \sum_{k_1=0,}^{n-1} \sum_{k_2=0,}^{K_2} \sum_{k_3=0}^{K_3} \cdots \sum_{k_{m-1}=0}^{K_{m-1}} \left(\frac{(n-1)!}{k_1! * k_2! * \ldots * k_m!} \right) * \prod_{j=1}^{m} p_j^{k_j} * \prod_{j=1}^{m} (1+\alpha * r_j)^{k_j/n} \qquad \text{[A1.4]}$$

And:

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

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

$$0 < (g_{n-1} - h_n)(\alpha) \leq g_{n-1}(\alpha) * \left[1 - \left(1 + \alpha * r_{\max} \right)^{1/n} \right] \rightarrow 0 \text{ as n gets large.} \qquad \text{[A1.5]}$$

Thus, $g_n'(\alpha)$ converges and it converges to $g'(\alpha)$.

We then have $g'(\alpha) = g(\alpha) * \sum_{i=1}^{m} \left(\frac{r_i * p_i}{1 + \alpha * r_i} \right)$. [A1.6]

The solution to this differential equation is:

$$g(\alpha) \quad = \quad \exp\{\sum_{i=1}^{m} p_i * \log(1+\alpha * r_i)\} \qquad \text{[A1.7]}$$

$$g''(\alpha) \quad = \quad g'(\alpha) * \sum\left(\frac{r_i * p_i}{1 \quad + \quad \alpha * r_i}\right) \quad - \quad \left(g(\alpha) \quad * \quad \sum_{i=1}^{m}\left(\frac{r_i^{\;2} \quad * \quad p_i}{\left(1+\alpha * r_i\right)^2}\right)\right) \qquad \text{[A1.8]}$$

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

$$\sum_{j=1}^{m}\left(\frac{r_j * p_j}{1 \quad + \quad \alpha * r_j}\right) \quad = \quad 0. \qquad \text{[A1.9]}$$

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

$$f(\alpha) \quad = \quad \sum_{j=1}^{m} p_j \quad * \quad u[w * (1+\alpha * r_j)] \;, \qquad \text{[A1.10]}$$

The expected utility of a single play, risking α*w when the current stake is w.

$$f'(\alpha) \quad = \quad \sum_{j=1}^{m} w * r_j * p_j * u'[w * (1+\alpha * r_j)] \;\; \text{and} \qquad \text{[A1.11]}$$

$$f''(\alpha) \quad = \quad \sum_{j=1}^{m}\left(w*r_j\right)^2 \quad * p_j * u''[w * (1+\alpha * r_j)] \qquad \text{[A1.12]}$$

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:

$$\sum_{j=1}^{m}\left(\frac{r_j * p_j}{1 \quad + \quad \alpha * r_j}\right) \quad = \quad 0 \qquad \text{[A1.13]}$$

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 logarithmic 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)).

There is a rational reason for selecting the logarithmic function as our utility function, particularly in financial decision-making situations. Its selection is consistent with and responsive to the notion of long-term success.

Appendix 2

Transformed Probability for Logarithmic Utility

Consider a positive random variable x with a (nontrivial) finite probability distribution, $p(x_j)$ where $x_j>0$ for j=1 to n. The sure-thing equivalent for the logarithmic utility function is given by:

$$G \;=\; u^{-1}(E[u(x)]) \;=\; e^{\sum_j \ln(x_j) \;*\; p(x_j)} \;=\; \prod_j x_j^{p(x_j)} \qquad \{A2.1]$$

which is the geometric mean of the distribution.

Can we describe a transformation $q(x_j) = f(p(x_j))$ such that $q(x_j)$ is a probability distribution and the expected value of x (with respect to q(x)) is equal to G?

Let $q(x) \;=\; p(x) \;*\; \frac{(G \;+\; a(x))}{x}$ where

$$a(x) \;=\; \alpha \;*\; (x \;-\; E[x]) \text{ and}$$

$$\alpha \;=\; \frac{(G * E[1/x] \;-\; 1)}{(E[x] * E[1/x] \;-\; 1)}. \qquad [A2.2]$$

Note that for positive distributions E[1/x] is well defined and that:

$$E[x] * E[1/x] \;>\; 1. \qquad [A2.3]$$

This follows from Jensen's Inequality. Note that equality implies a trivial (i.e., "sure thing") probability distribution. Also note that:

$$G \;\leq\; E[x] \qquad [A2.4]$$

(again, this is an instance of Jensen's Inequality) and that:

$$\frac{1}{G} \;\leq\; E[1/x] \qquad [A2.5]$$

(Jensen's Inequality applied to the transformation y=1/x).

These together imply that α is well defined and that it is positive.

Now, we have:

$$\sum_j (x_j \;*\; q(x_j)) \;=\; G \;+\; E[a(x)] \;=\; G \qquad [A2.6]$$

So, we have remaining the need to show a) that $\sum_j q(x_j) = 1$ and b) that $q(x_j) \geq 0$. It suffices to demonstrate these in the discrete case, but the same argument applies to the continuous as well.

a. $\sum_j q(x_j) = G*E[1/x] + E[a(x)/x] = G*E[1/x] + \alpha - \alpha*E[1/x]*E[x]$

$= \big((G*E[x]*E[1/x]*E[1/x] - G*E[1/x]) + (G*E[1/x] - 1)$

$- (G*E[x]*E[1/x]*E[1/x] - E[x]*E[1/x])\big) \div (E[x]*E[1/x] - 1)$

$= (E[x]*E[1/x] - 1)/(E[x]*E[1/x] - 1) = 1$ [A2.7]

b. for p(x) > 0, q(x) ≤ 0 if and only if $\frac{(G + a(x))}{x} \leq 0$. Since x is assumed to be >0 and because α>0, q(x) ≤ 0 if and only if G +a(x) ≤0. Now if this is true of any x_j for which $p(x_j) > 0$, it will be true for the smallest such x_j. Thus, it will suffice to show that $G + a[\min(x_j)] > 0$. Let x_S be the minimum value for the distribution. By assumption it is > 0. Also, $G*E[1/x] \geq 1$ and $E[x] - G \geq 0$ (from [A2.5] and [A2.4] respectively).

Thus, $x_s * (G*E[1/x] - 1) + (E[x] - G) \geq 0$. [A2.8]

Adding to and subtracting from the left-hand side the term G*E[x]*E[1/x], then rearranging we get:

$(x_s - E[x]) * (G*E[1/x] - 1) + G*E[x]*E[1/x] - G \geq 0.$ [A2.9]

Equation [A2.3] above implies that $E[x]*E[1/x] - 1 \geq 0$, so if we divide [A2.9] by $E[x]*E[1/x] - 1$ we get:

$$a(x_s) + G \geq 0. \quad \text{[A2.10]}$$

Q.E.D.

Now as x_s approaches zero (and the other values are unchanged), the transformed probability $q(x_s)$ will approach unity and the other $q(x_j)$ will go to zero.