© *The Journal of Risk and Insurance*, 2005, Vol. 72, No. 1, 1-21

Real and Illusory Value Creation by Insurance Companies

David F. Babbel
Craig Merrill

Introduction

One of the fundamental tenets of financial economics is that insurance companies, just as other financial and non-financial firms, have a very strong incentive to maximize current shareholder value.[1] This seemingly simple observation leads to a wide variety of managerial behavior.[2] In this article we will review a simple decomposition of the value of an insurance company. The decomposition illuminates the motivation for a variety of strategies that can be observed in practice. These strategies run the gamut from accounting manipulation through risk transfer schemes to positive net present value (NPV) "project" selection. We will review these in some detail below. Then we will illustrate these techniques from our experience with every major U.S. life insurer insolvency since Baldwin United in 1984.

Components of Firm Value

The market value of insurance company owners' equity is defined as the difference between the market value of assets and the market value of liabilities. For purposes of valuation, it is helpful to partition more finely the components of equity value. In this section, we will partition the value of insurance company owners' equity, or stock in the case of a stock company, into its four major components: franchise value, market value of tangible assets, present value of liabilities, and put option value (see Figure 1). The first two of these components are clearly assets. The third component

David F. Babbel is at the Wharton School, University of Pennsylvania. Craig Merrill is at the Marriott School, Brigham Young University.

[1] In the case of a company with publicly traded common shares, this amounts to maximizing the price of the stock. While this firm goal is widely accepted for a stock corporation, its applicability to a mutual insurer differs because a policyholder of a mutual is both a creditor and an owner. In this case, maximizing the combined value of the policy claims and equity interests to policyholders would be analogous.

[2] There is a vast literature on agency theory that stems from articles such as Jensen and Meckling (1976) or Fama and Jensen (1983). Applications in the insurance markets begin with articles such as Mayers and Smith (1981) or Mayers and Smith (1994). The behavior of management will depend in part on the incentives provided by the compensation scheme.

FIGURE 1
Market Value of Insurance Equity

Market Value of Equity	=	Market Value of Assets	−	Market Value of Liabilities

Market Value of Equity	=	Franchise Value	+	Market Value of Tangible Assets	−	Present Value of Liabilities	+	Put Option

is related to liability value. The put option value can be treated as part of the liability value as a contra-liability or as an asset. We will discuss each component in turn.

The franchise value stems from what economists call "economic quasi-rents." It is the present value of the "quasi-rents" that an insurer is expected to garner because it has scarce resources, scarce capital, charter value, licenses, a distribution network, personnel, reputation, and so forth. It includes renewal business.[3] Franchise value is dependent on firm insolvency risk. The less insolvency risk there is, the more likely the firm is to remain solvent long enough to capture all the available economic rents arising from its renewal business, its distribution network, its reputation, and so forth.[4]

The market value of liabilities includes pricing of all contingencies within the insurance liabilities. In our decomposition we have elected to separate out the risk of the insurance company defaulting on its obligations. The risk of such a default is captured by the put option. Other contingencies, including such risks as interest rate contingencies, mortality contingencies, or equity market contingencies, are included in the present value of liabilities. Thus, the present value of liabilities is the present value of all promised liability cash flows discounted at Treasury rates rather than discounting at rates that would reflect the possibility of default. In the case of interest sensitive cash flows, it is the value of the Treasury security portfolio (including derivatives as needed) that fully defeases the liabilities with all non-default contingencies fully hedged. This quantity will be larger than the market value of liabilities where the promised liability cash flows are discounted at rates that reflect the risk of failure to make the promised cash flows. Merton (1974) initially observed that the value of a debt obligation could be expressed as the value of a default-risk free bond minus a put option. The put option represents the value to the issuer of the possibility of defaulting on the obligation. The same holds true for insurance liability obligations. The market value of liabilities can be expressed as the present value of liabilities minus

[3] Quasi-rents are revenue minus the variable costs but not the capital costs. A firm with zero net present value would result from the present value of quasi-rents equals the present value of its capital costs. Quasi-rents would be reduced (usually slightly) by the present value of payments that the solvent insurer is expected to pay to state insurance insolvency guaranty programs, to the extent that these payments are not fully offset by state and federal tax reductions.

[4] Often, it is protection of the franchise value that motivates risk management. Smith and Stulz (1985) discuss the theory of hedging by firms. Santomero and Babbel (1997) discuss the practice of financial risk management by insurers and Colquitt and Hoyt (1997) discuss corporate hedging in the life insurance industry.

the put option that represents the value to the issuer of potentially defaulting on the obligations.[5]

The market value of tangible assets and the present value of liabilities can be netted together, producing what we will call "net tangible value." This value is simply the market value of tangible assets, less the present value of liabilities.[6] This net tangible value is independent of what *kind* of assets an insurer has, but does depend on the *amount* of assets it holds. This is because the quality of assets will affect the put option value and franchise value but not the present value of liabilities or the market value of tangible assets. For instance, if the firm swaps Treasury securities worth $5 billion for junk bonds worth the same (but with higher coupons and/or face values), or swaps them for $5 billion worth of pork belly futures, the market value of tangible assets is the same. Similarly, the present value of the *promised* cash flows to insurance consumers remains the same (although the quality of the promise has changed—more will be said about this later). Therefore, the net tangible value of the firm at any given moment is unaltered by the kind of assets the firm holds. Moreover, it is completely independent of firm insolvency risk, although how this value evolves over time will depend on risk.

Put option value arises from the limited liability enjoyed by equity holders when their firm issues debt (i.e., insurance policies, the major debt of an insurance company). It is the value to equity holders of capturing the upside earnings while not incurring all the downside costs of default. The insolvency put option increases in value as the insurer takes on more risk. If the insurer faces minimal insolvency risk, there may be little benefit inuring to it from this component of value; but if it is a risky firm, the implicit insolvency put option may be of considerable value.

In what follows, we will use the more compact notation to represent the above equations:

$$V(E) = V(A) - V(L), \tag{1}$$

$$V(E) = F + V(A_T) - PV(L) + O, \tag{2}$$

where $V(\,)$ is the market value, $PV(\,)$ is the present value, and E is the owner's equity; A, the assets; L, the liabilities; F, the franchise value; A_T, the tangible assets; O, is the default put option.

[5] An equivalent interpretation of the put option value is that it is the value to the insurer of being able to issue debt (i.e., insurance policies) with embedded interest rates lower than would otherwise be required were it not for the guarantee programs. The decomposition of liability value into a default risk-free component and a default put option is discussed more completely in Babbel, Gold, and Merrill (2002). Phillips, Cummins, and Allen (1998) provide a discussion of financial pricing as applied to insurance liabilities.

[6] Elsewhere this has been referred to as "liquidation value." However, because that expression is laden with connotations that are unhelpful in this context, we prefer to use the expression "net tangible value" here. In an actual liquidation, assets are sometimes unloaded at "fire sale" prices. This reduction in value may or may not be offset if the liquidator is able to extract favorable terms from assumption reinsurers, because the book of business being transferred may still have some renewal value.

At this point it is useful to pause and consider these juxtaposed equations. Because the left-hand side of both equations is the market value of owner's equity the right-hand sides of both equations must be equal to each other. That is,

$$V(A) - V(L) = F + V(A_T) - PV(L) + O. \tag{3}$$

It then follows that the market value of liabilities differs from the present value of liabilities unless the market value of assets is defined as the sum of the franchise value, market value of tangible assets, and put option. That is,

$$V(A) \equiv F + V(A_T) + O. \tag{4}$$

But, this definition of $V(A)$ and $V(L)$ would have a regulatory implication. This broad definition of the market value of assets is not universally used and, therefore, neither is the associated equivalence of market value of liabilities with present value of liabilities. For instance, oftentimes people will think of the market value of liabilities as being the same thing as the market value of debt from an investor's point of view. In which case the value of the insurance liability would be $V(L) \equiv PV(L) - O$. Thus, if a bond that is subject to credit risk has been issued, its market value is reduced below that of an otherwise comparable default-free bond. Accordingly, the market value of liabilities is reduced when the debt becomes riskier. This is ironic from an insurance regulatory perspective. If the insurer is likely to default on its insurance policies, it would suggest that the liabilities are worth less than they would be if they were secure. The riskier a firm is the lower would be the market value of its liabilities. Therefore, were the insurer to report this market value of liabilities to regulators, the lower market value could suggest the insurer is in better financial health than if the insurer were to report a present value of liabilities that is not reduced by the prospect of insolvency.[7]

From a financial economics perspective, a firm issuing a default-prone bond has issued a combination of a default-free bond along with a default put option. The value of the bond is reduced by the value of the put. The investor is long a default-free bond and short a put.[8] The issuer is short a default-free bond and long a put. This put option would normally be on the asset side of an "economic balance sheet," although it appears nowhere on the typical accounting statement; however, it is implicit in the choice to state liabilities on a market value basis.

In Figure 2, each of these components of value is displayed separately as a function of firm insolvency risk.[9] Because we are focusing on insolvency, or default, risk franchise

[7] Executive Life Insurance Company of California, a $16 billion insurer which bet heavily on junk bonds until its demise in 1991, is an example of a company whose market value of liabilities (under this definition) would have been reduced far below the present value of its promised cash flows to policyholders.

[8] This notion was introduced by Merton (1974, 1977).

[9] We deliberately leave the method of measuring firm insolvency risk general at this point because the general shapes of curves in the Figures depicting the theoretical relationships are robust to the method of risk measurement.

FIGURE 2
Insolvency Risk and the Components of Equity Value

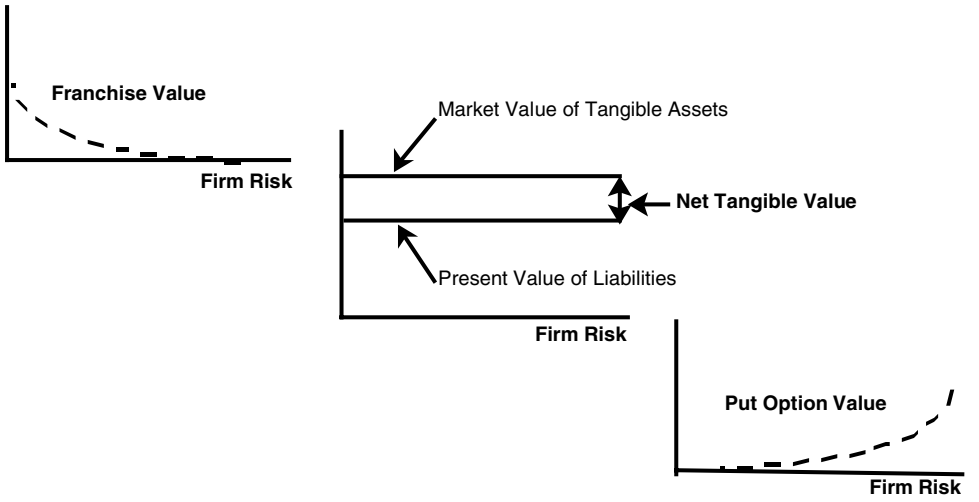
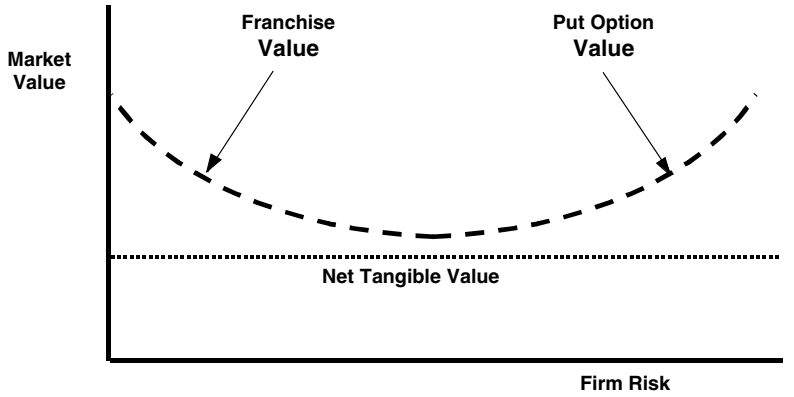


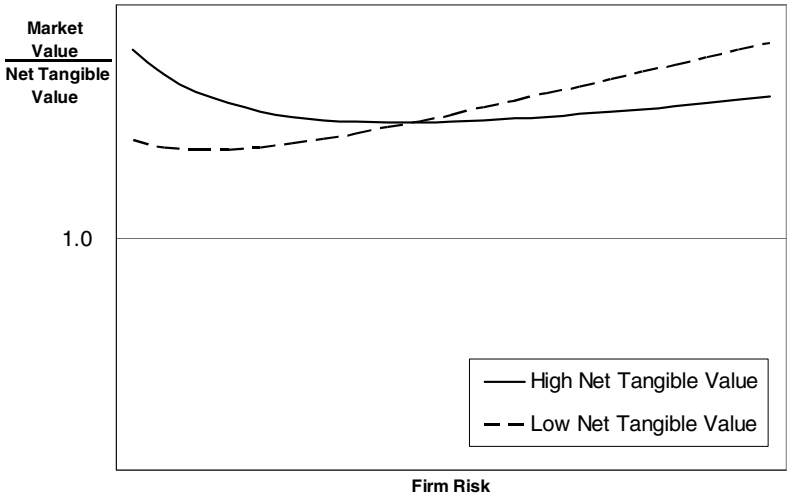
FIGURE 3
Components of Equity Value Combined



value declines while put option value increases with increases in risk. By definition, tangible asset value and the present value of liabilities will be invariant to increases in insolvency risk.

When these value elements are displayed together, as in Figure 3, we can see how the overall market value of the firm is related to its risk exposure. As the firm increases in insolvency risk beyond some moderate level, firm market value increases, and as it decreases in risk below that same moderate level, again there is an increase in firm value. This equity market value premium over net tangible value stems either from franchise value or from put option value, or from some combination of the two. This is not merely a theoretical construct. It has been accepted wisdom in the

FIGURE 4
Effect of a Decline in Net Tangible Value



financial institutions literature for decades. Empirical research suggests the effect is pronounced in the case of insurance firms.[10]

Figure 4 shows the effect of a decline in net tangible value. The figure was constructed assuming a fixed market value of tangible assets and two levels of present value of liabilities. The line labeled "High Net Tangible Value" is associated with a lower present value of liabilities and the line labeled "Low Net Tangible Value" is associated with a higher present value of liabilities. The vertical axis is scaled so that both hypothetical firms' values are stated relative to net tangible value. Figure 4 shows how a decline in net tangible value is associated with lower franchise value and increased value for the put option as risk increases. A firm seeking to maximize its value will have a greater incentive to take risk, and increase the put option value, as net tangible value declines.

While theory suggests that maximizing corporate value yields social efficiency in perfect markets. Real markets are far from financial economic perfection. Bankruptcy is attended by tremendous deadweight costs. Regulators are charged with the protection of the individual policyholder. Thus, a goal of regulatory policy should be to provide an incentive structure that is consistent with maximizing the value of owners' equity by having an insurer adopt a low-risk profile. Otherwise, the insolvency guarantee programs become mechanisms for rewarding risky firms and subsidizing their behavior by low-risk firms. Any regulatory action that would reduce franchise value is likely to have the effect of enticing some insurers to adopt riskier profiles in order to maximize their stock prices. Insurers most likely to respond in such a manner

[10] Babbel and Staking (1989, 1991) and Staking and Babbel (1995, 1998).

include those who have little to lose—whose franchise value is minimal and whose net tangible value is also low as can be seen in Figure 4.[11]

While the incentive to adopt a risky profile in an effort to maximize owners' equity value is strong for many insurers, particularly for those with low franchise value, there is still a downside to being risky, even when the franchise value is low. A portion of the equity value is comprised of liquidation value, and this will be lost to the equity holders if the insurer makes some risky bets that turn out to be wrong.

Thus, the name of the game for insurers with low franchise value is to carry as little liquidation value as possible, so that the equity holders are faced only with upside gains potential.[12] Then if the insurer makes risky bets and wins, the equity holders reap the rewards. If the insurer loses its bets, others (i.e., the policyholders and the various state insurance insolvency guarantee programs) pick up the tab.[13] The incentive to "bet the bank" becomes overwhelming when franchise value and liquidation value are both low, often overcoming any restraints or inhibitions provided by statutes aimed at promoting insurer solvency.

MANIPULATIVE APPROACHES TO VALUE MAXIMIZATION

Numerous techniques have been used by insurers to get into a position where they are motivated more heavily by the put option value. These techniques involve trading away franchise value and minimizing net tangible value. Note that it is the true net tangible value which would be forfeited in a default, and not the statutory surplus, that mitigates to some extent the incentive to exploit the put option component of equity value. The net tangible value is determined, in part, by the true market value of assets. Regulators, however, generally focus on statutory surplus, which is determined, in part, by the book values of admitted assets. Thus, if an insurer sells assets containing capital gains, while holding onto those assets exhibiting losses, it can inflate statutory surplus, while leaving liquidation value unimproved. Indeed, an insurer that makes a habit of doing this can show a positive statutory surplus but a zero or negative net tangible value. Changes in insurance accounting render this tack less effective, but there remains the possibility of doing this, at least to some degree.

---

[11] Consider the U.S. savings and loan industry during the 1970s and 1980s. Two of the key ingredients of the industry's franchise value were its ability to attract deposits by offering rates merely 1/4 of 1 percent higher than those of commercial banks and its important role in raising money for residential mortgages. With the deregulation of financial institutions, S&Ls lost their statutory advantage in attracting deposits, and with the advent of mortgage-backed securities, they lost their pivotal role in the mortgage industry. With much of their franchise value thus eroded, it became attractive to S&Ls to "bet the bank" on risky investment strategies, a bet which many of them (as well as taxpayers) lost.

[12] It is not always a deliberate strategic plan for an insurer to carry a low liquidation value. Sometimes a firm happens into a streak of bad luck and loses much of its liquidation value. Other times it is a corporate strategy to operate with as little liquidation value as is permissible. But in either case, once the liquidation value falls sufficiently low, the incentive to adopt a risky investment posture can become very strong.

[13] Ultimately, taxpayers end up bearing much of the burden of insolvency to the extent that insurer payments to remedy insolvencies offset the taxes they pay, or are tax deductible.

Critical to the scheme is the insurer's ability to maintain enough statutory surplus so that regulators will allow the insurer to continue playing the game. This may be accomplished by getting regulators to allow the economic value of possible future profits to substitute for surplus through the insurer engaging in such dubious practices as surplus relief reinsurance and commission financing.

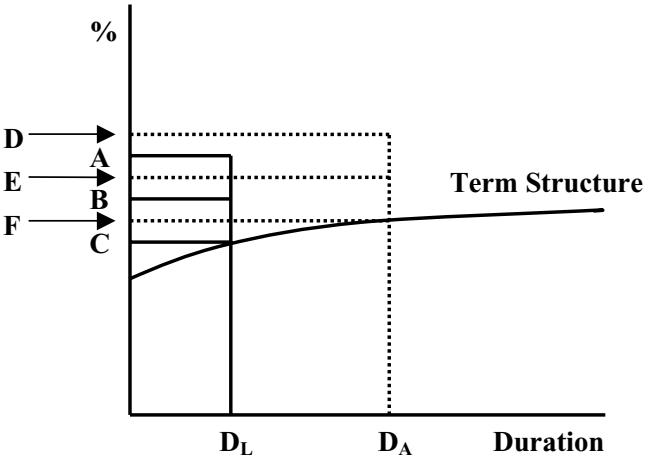
Surplus relief reinsurance is a portfolio reinsurance contract where an admitted reinsurer assumes a ceding company's business to relieve stress on the cedent's policyholder surplus. The portfolio transfer allows the ceding insurer to recapture the equity in unearned premium reserves on the business ceded. Surplus relief reinsurance has been a way used by some insurers to transfer off of their books the value of some of their liabilities, in exchange for a portion of expected future profits. This improves the surplus ratios relative to liabilities and assets, as forgone future profits do not generally appear on the balance sheet. However, some surplus relief reinsurance has even been designed in such a way as to avoid transferring any commercial risk whatsoever, yet still producing the positive book effects on surplus of reducing reported liabilities in exchange for a relatively small fee.

With commission financing the insurance firm avoids statutory surplus drain by setting up a separate entity that undertakes debt to finance the payment of sales commissions. This is a way of "front-loading" earnings and is done through a separate company to keep the true economic substance of the transaction off the books. Effectively, The resulting amortization of sales commissions produces favorable short-term earnings and reduces surplus drain, as described in greater detail later in this article.

Other techniques can be used to "front-load" earnings. For example, assets whose book yields are higher than their expected returns may be purchased. Investments such as mortgage-backed securities, corporate bonds laden with call features, and junk bonds will accomplish this. Certain tranches of Collateralized Bond Obligations (CBOs) and Interest Only mortgage strips (IOs) can serve a similar purpose. All of these investment classes can also serve a legitimate purpose. The point is that when net tangible value and franchise value are low, it is more likely that the front-loaded earnings characteristic of these investments is an important motivation to acquire them in order for the insurer to show book results that will enable it to remain in business longer than it might otherwise. The insurer might hope for improvement in the economic condition of the firm before problems appear in statutory filings because statutory valuation principles can limit a regulator's authority or mandate to shut down an insurer that is within the strict limits of the statutes.

Another technique to front-load earnings is to engage in a duration mismatch strategy, where higher yields can be booked on assets than what would be the case if the assets were better matched to the liabilities. Figure 5 depicts one possible result of a duration mismatch strategy. This figure is based on a life insurance company issuing single premium deferred annuities (hereafter SPDAs). Points A, B, and C relate to liability interest rates and points D, E, and F relate to asset yields. In this case the insurer is trying to enhance results by choosing assets with longer duration than liabilities. This strategy has the potential to enhance book results and statutory surplus. It may even increase net tangible value over time if interest rates are not too volatile. It can, however, actually decrease the economic value of the firm if franchise value is more

FIGURE 5
Basis Risk in an Attempt to Front-Load Earnings



Liabilities:
- A = Effective (option adjusted) yield on liabilities
- B = Stated crediting rate on liabilities
- C = Treasury rate with liability duration

Assets:
- D = Stated yield on asset with duration D_A
- E = Effective yield on asset with duration D_A
- F = Treasury rate with asset duration D_A

Spreads:
- D – B = Stated spread (book spread)
- E – A = Effective spread (< 0)
- $D_A - D_L$ = Duration mismatch (basis risk)

than commensurately reduced by the risk that the economic rents may not be fully captured due to the increased exposure to interest rate risk.

Consider the liabilities. The stated crediting rate on the SPDAs is B. The insurer is offering a spread over Treasury securities with similar duration. The effective crediting rate includes the stated crediting rate as well as the cost of options inherent in the liability. For example, one option that SPDA holders may exercise is the right to surrender their policies when interest rates change. Other options that SPDA holders may benefit from are minimum value guarantees or death benefits associated with the contract. Both of these option-like features have value for the policyholder and increase the effective yield (or cost) of the liability. Thus, while the stated crediting rate, or cost, of the policy is B, the actual cost, with the cost of options included, may be A.

On the asset side, imagine that the insurer depicted in Figure 5 invests in low-grade bonds. Low-grade bonds offer a higher yield than would bonds with identical features if issued by the Treasury. The difference in yield is due to default risk. While the stated yield at time of purchase may have been D, offering a potential spread over similar Treasury securities equal to D – F, the effective yield declines as defaults occur. The effective yield may wind up at E. In other words, the effective (realized) yield on default-prone corporate securities is lower than the promised yield because of losses when defaults occur.

The insurer depicted in Figure 5 wrote liabilities and invested in assets to generate a stated spread of $D - B > 0$ yet generates an effective spread of $E - A < 0$. The negative effective spread is an outgrowth of failure to properly model option-like features on both the asset and liability side of the balance sheet.

There is strong evidence that in the property/liability insurance industry, the market rewards firms that adopt a high exposure to interest rate risk via duration mismatch, as well as rewards those that minimize the risk of loss in franchise value by matching assets to liabilities. Firms in the middle, with moderate risk levels, do not command as high a stock price as those in other two groups. This evidence is consistent with the relationships depicted in Figures 3 and 4.[14]

There are also numerous actuarial techniques available that can accomplish similar objectives (i.e., showing enough statutory surplus to keep the insurer licensed while having little or no liquidation value and trading away franchise value). Using the most recent mortality and morbidity tables available, trending the mortality data aggressively if it tends to produce lower required reserves, applying the highest discount rates permitted, and using less conservative reserving methods, to name only a few, are techniques that can produce lower reserves and bump up the statutory surplus.

Suffice it to say that when the motivations are strong to exploit the put option value via added risk, some insurers are going to avail themselves of the opportunities, creditor interests notwithstanding. The trend toward reporting market values both of assets and liabilities and of imposing risk-based capital requirements, while far from perfect, are important steps in a generally positive direction to reduce the ability of an insurer to exploit the put option value component of equity value.

POSITIVE NPV APPROACHES TO VALUE MAXIMIZATION

Rather than using manipulations based on accounting treatment, actuarial assumptions, basis risk bets, or acceleration of earnings, insurers can focus on fundamental approaches to creating value. Value is created for shareholders when the firm engages in positive NPV projects. In a competitive market, positive NPV is not easy to find, but economic theory does suggest a couple of possibilities for insurers.

There are two possible places for an insurer to go looking for positive NPV projects: the asset side and the liability side of the business. Insurers tend to hold a lot of publicly traded securities as assets. The extensive finance literature on market efficiency suggests that it is very, very difficult to create sustainable positive NPV through investing in publicly traded securities. While it is true that an investor might be lucky

[14] See footnote 8.

because of surprise positive returns or extra yield through unrealized credit risk, the market for investment in publicly traded securities is simply too competitive for a reasonable chance of sustainable excess performance in this area.[15]

Insurers also invest in illiquid assets such as real estate and private placement debt. Both of these are areas that have the possibility of rewarding sound analysis and shrewd investment. It is important to note, however, that there are many players in both of these markets and competition tends to drive excess profits (NPV) to zero.

Insurers raise capital through the issuance of liabilities. In place of issuing bonds, that are likely to have zero NPV, insurers issue policies that offer the opportunity to yield a positive NPV. It is well established in the economic literature that insurance demand is driven by preferences that allow insurance companies to charge more than the expected loss cost for insurance coverage.[16] Of course competition tends to drive the NPV of providing insurance toward zero, but it is at least theoretically possible to create insurance products that have a positive NPV for the issuer.

As compared to the publicly traded debt markets, insurance liabilities have several features that allow for the possibility of product differentiation consistent with profits that might exceed those available in a perfectly competitive market. In a perfectly competitive market products are homogeneous, allowing for easy comparison and switching by consumers. Insurance liabilities tend to be complex. Individual insurers may offer products that differ from one state to another. In addition, policy provisions vary from one company to another. Beyond varying policy provisions, insurer quality varies in both customer service and in probability of default.

Another assumption underlying perfect competition is that full information is freely available. Information about insurers and policies tends to be opaque from the consumer's perspective, and many sales agents are only marginally better equipped to fully analyze the value of policy nuances. There are dozens of different methods for determining cost. If the industry hasn't determined the "correct method" to measure cost then consumers and sales agents cannot be expected to be able to measure and compare costs completely. When there is asymmetric information, the better informed party ought to be able to profit.

Asset/Liability Management as an Approach to Preserve Value

An insurance policy is a contingent claim to the policyholder. The contingency is triggered by a loss or a claim or a policyholder decision. An insurer takes the opposite side of that contingent claim. To the extent that the losses or claims are uncorrelated, an insurer can use the risk pooling principle and satisfy the claims through a broadly diversified investment portfolio. Much of the existing and proposed insurance investment regulations have implicitly operated on that assumption. However, many demands on insurers arise in clusters, based on loss events and policyholder claims which are highly correlated.

---

[15] Our analysis does not consider the effect of taxes. Clearly, with taxes, there are ways to allocate assets that will reduce the dead-weight losses due to inefficient portfolio structure in the face of taxes.

[16] See Dionne and Harrington (1992) for a good review of the literature.

Indeed, in recent years, more and more of the policy claims to insurers exhibit these characteristics. Three examples of these, two of which are taken from the property/liability branch of the insurance business, are SPDAs, workers compensation, and earthquake coverage. With SPDAs, policyholder withdrawals and surrenders tend to occur when interest rates on competing products rise above contract rates sufficiently to make it economically attractive for the policyholders to do so. With workers compensation, there seems to be an increase of claims (supposedly based on workplace injuries) when recessions occur and unemployment rises. With earthquake coverage, clustered claims arise from properties similarly affected.

In any of these cases, and in a host of others, having a broadly diversified portfolio of high-grade securities does not ensure that the policyholder demands and claims will be met. We will illustrate this with an example and discuss other examples in the next section of the article.

Consider an insurer that has 100 percent of its portfolio invested in mortgage-backed securities issued by government-sponsored agencies. Presumably, the reason for this is due to the low perceived credit risk in such securities, as well as the higher stated yields they offer relative to other securities of similar default risk. This might make some sense if the investment risk is viewed in isolation from the liabilities they are designed to hedge.

The problem of funding life insurance and annuities with mortgage-backed securities arises in the prepayment risk of mortgages and the lapse risk of the insurance products.[17] At the same time that mortgages are being prepaid in a declining interest rate environment, thereby shortening the maturity profile of the asset portfolio, the insurance policies are "lengthening" due to the favorable portfolio yields that are being credited on life policies and annuities. Some of the contributing factors to the lengthening duration are lower lapse and lower surrender rates, exercise of over-depositing privileges, the basic effect of lengthening durations occasioned simply by lower discount rates, and minimum crediting rate guarantees that are more likely to be in the money. During protracted periods of declining interest rates, the market value of callable and prepayable assets cannot keep pace with the increasing present value of liabilities. Conversely, when interest rates rise, mortgages lengthen as the prepayment rate slows, existing residential mortgages at lower interest rates are assumed (where legal) as part of the house sale, while life insurance and annuity products shorten, as lapse rates increase. In this case, asset market values decline more rapidly than the present value of liabilities. In either case, surplus is squeezed as interest rates change.

A position where value declines regardless of the direction of interest rate movements is called a "short straddle" by options traders—a very risky posture to assume when interest rates are volatile.[18] In a short straddle, an insurer may be profitable as long as interest rates do not move appreciably in either direction, but becomes damaged if interest rates change in either direction by more than a modest amount. While it is

[17] Prepayment risk is the financial risk that a mortgage of fixed term will be prepaid prior to its maturity date, with its associated loss in value. Lapse risk is the risk that a policyholder will cease making periodic payments as required by the policy terms, with its associated loss in value to the insurer.

[18] See Babbel and Stricker (1989).

possible to combine mortgage-backed securities with other assets, such as interest rate caps and floors, or to actively manage them in a manner that avoids the predicament of a short straddle, institutional and regulatory barriers to derivatives use may hinder such strategies or limit the extent to which they may be employed.

Even when duration and convexity are correctly measured, there is "basis risk" that is often overlooked, yet very important. Most duration measures are designed to indicate the percentage change in market value (or present value) of an asset or a liability if interest rates move by 1 percent. The problem is that the interest rates of focus are generally the yields or discount rates of the asset or liability in question. Because of basis risk, the yields and discount rates on assets and liabilities of differing qualities do not change in lockstep. For example, two assets may both exhibit durations of 10, yet they may respond in differing degrees to movements in market interest rates. The one of lower quality may change less in value than the one of higher quality, because the yields on lower quality bonds may move less than those on higher quality bonds.[19] Moreover, even if the credit quality is the same across more than one bond, as is the case for Treasuries, there remains basis risk. The yield-to-maturity of one Treasury bond is unlikely to move by the same amount as the yield-to-maturity of another Treasury bond of different maturity.[20]

Beyond accounting for non-parallel shifts in yields, duration calculations should also incorporate interest rate contingencies in the cash flows. This is called effective, or option-adjusted, duration. Just as call provisions in a bond shorten duration, interest rate contingencies in insurance liabilities change the interest rate sensitivity of those liabilities. Examples of the consequences of ignoring interest rate contingencies for duration matching will be given in the next section.

There is a more fundamental challenge to proper asset/liability management. Even the best duration calculation, when performed on book values, is meaningless. A necessary prerequisite to good duration and convexity calculations is good valuation models. There is a growing literature on financial valuation of insurance liabilities that is forming the basis for better valuation tools in practice.[21] In addition, the literature on valuation of interest sensitive financial claims is also expanding our understanding of how to value disparate instruments in a consistent fashion. A consistent valuation approach will facilitate duration calculations that are robust to less than perfectly correlated yield shifts.

The Financial Accounting Standards Board (FASB) and the American Academy of Actuaries, as well as the International Accounting Board (IAB) and other international accounting and actuarial bodies, are working on guidelines for fair valuation of insurance liabilities. In this context the fair value accounting rules are intended to provide something like a market valuation methodology for liabilities that are illiquid. As better tools emerge in this area, insurers will be able to do a better job of modeling interest rate contingencies in their liabilities in order to price them and manage the associated risks more effectively.

[19] See Babbel, Merrill, and Panning (1997).
[20] This relationship is depicted in Babbel (1983).
[21] Babbel, Gold, and Merrill (2002) and Babbel and Merrill (1998).

LESSONS FROM THEORY ILLUSTRATED IN PRACTICE

The concepts presented above seem fairly simple to understand. And yet, the true importance of these concepts is really seen in how often they have been at the heart of insurance company failures. In this section we offer a list of anecdotes that illustrate how important the concepts in this article are for management, directors, regulators, and investors.

As mentioned earlier, we have been involved in some way in every major U.S. life insurer insolvency since Baldwin United in 1984. Our roles ranged from insurance commissioner appointments to study and report on various insurance issues, court appointments to advise, plaintiff and defense expert witnessing, company advisors or directors, federal government advisors, rehabilitation/conservatorship consultants, portfolio valuations, and liquidation/merger advisors. Among the companies involved were Baldwin United, Executive Life of California, Executive Life of New York, E. F. Hutton Life, First Capital Life, Fidelity Bankers Life, Mutual Benefit Life, Confederation Life, American General, Integrity Life, National Integrity Life, Mutual Fidelity, and a host of other insolvencies and large "near misses."

During these engagements, we were often provided with private or court-protected information. A number of these companies continue under litigation and we cannot reveal certain information. We are at liberty to disclose identities and information from some of the companies, but were we to do so, it would be fairly easy to identify some of the other companies through the process of elimination. Therefore, we will illustrate a number of actual problems we studied, but without providing any identifying information. Most of the illustrations provided below relate to more than one company, and most of the companies we studied were involved in more than one of these problems.

Bogus Surplus Relief Reinsurance

In the section of this article discussing manipulative approaches to value maximization we discussed surplus relief reinsurance as an approach to front-end load earnings. Under certain circumstances, an insurer that is surplus constrained can "rent" the surplus of another carrier who is better situated, or that of a reinsurer. Generally, these arrangements call for a healthy rental fee and/or the forfeiture of future profits from the reinsured block of business in exchange for the surplus relief. Regulators recognize these treaties and allow this "admitted reinsurance" to offset a portion of the insurer's liabilities, thereby strengthening its financial position.

However, we found a number of these contracts to be completely bogus. They were crafted in such a way that there was no real risk transferred, or if a risk seemed to be transferred, the treaties required the primary carrier to reimburse the reinsurer completely for any and all losses. These contracts were rather burdensome to decipher—typically 90 pages long or more—and it would require someone pretty skilled in algebra to sort through all of the commingled terms before it could be proved that, indeed, there was no risk transferred at all, yet surplus relief was granted.

The insurer and its shareholders may potentially benefit from this behavior because it allows the insurer to pursue risky strategies whose costs are borne by others. It allows

the insurer to circumvent the regulatory surveillance that is designed (poorly in many cases) to prevent this behavior.

Commission Financing Schemes

Statutory accounting principles (SAP), in some ways, have a conservative bias in order to foster solvency. An example of this bias is how sales commissions are expressed when paid, rather than amortized over time like under generally accepted accounting principles (GAAP). The SAP treatment creates a surplus drain that can limit the growth rate of underwriting new liabilities. Some insurers have found a clever way around this "speed bump." They get a bank to pay the sales agents, perhaps through a special purpose entity (SPE). The insurer then writes a note to a subsidiary, which in turn writes a note to the bank or SPE. The loan is then amortized over a period of several years and, voilá, a GAAP-type expensing pattern has effectively been substituted for the up-front commission hit that would have occurred under SAP. The net effect of this scheme is to allow the insurance company to leverage its surplus farther than intended under the SAP rules.

Effective vs. Ineffective Duration

Some companies reported regularly to their shareholders that their policy was to avoid interest rate risk by carefully matching assets to liabilities through a disciplined duration management regime. The ultimate goal of this practice was to constrain the duration of net tangible value (or, in some cases, the duration of surplus) to less than two years. This practice is a variation on the ideas presented in Figure 5 above. In Figure 5 the insurance company was attempting to take duration mismatch risk in order to enhance yield. Unfortunately, while the stated yields seemed to create positive value, the effective yields were actually consistent with earning a negative spread. Just as when effective yield is measured (or stated) incorrectly, the same sort of error can arise when duration is measured (or stated) incorrectly.

Recall that traditional duration measures, such as Macaulay or modified duration, measure interest sensitivity due to changes in discount rates. For more complex securities that have interest rate contingencies in their cash flows, traditional duration measures can significantly misstate actual interest rate sensitivity. Measures of effective duration take into account how changes in interest rates cause both changes in cash flows, through interest rate contingencies, and the effect on discounting.[22]

What we often found in practice was that companies were not measuring their interest rate exposure properly. They were using unsuitable traditional duration measures for their assets, liabilities, or both, when interest rate contingencies existed and were significant. Because traditional measures of duration misstate effective duration, traditional measures must not be applied to life insurer liabilities such as annuities and life insurance policies, nor should they be applied to the vast majority of insurer assets, because these measures do not capture the interest-sensitive nature of the cash flows.

[22] See Babbel (1999) or Babbel, Merrill, and Panning (1997) for discussion of effective duration.

On many assets, effective duration measures, that are correctly option-adjusted, will be within 20–30 percent of the traditional duration measures, but when it comes to insurance liabilities, the difference between effective and traditional measures can differ by several hundred percent. Clearly, a prudent and informed company should not attempt to match assets and liabilities by use of traditional measures or the apparent match can result in an effective duration mismatch of gargantuan proportions.

One company we studied purported to have a very close (traditionally measured) duration match, yet we found it had an effective duration for net tangible value (NTV) of over 80 years, a huge mismatch. Another company, also purporting to have eliminated exposure to interest rate risk, had an effective NTV duration of 235 years. Both of these companies were so highly levered that even a modest mismatch put their NTV at considerable risk. While they showed only a tiny mismatch using traditional measures, they were mismatched considerably using effective measures, and their high leverage exacerbated the problem.

As an alternative to a duration management strategy, some companies would attempt to control their interest rate exposure through cash flow matching. Other companies used a combination of duration management along with cash flow matching.

We noticed problems with the pure cash flow matching approach, however. The asset side was almost always done very poorly. Rather than adjust for the market risk in projecting cash flows, these firms would either simply project promised coupon payments, or adjust the promised coupons only for historical default rates and, in some cases, test for the effect of higher default rates. But in no case did they check for the market value of the bonds, which incorporates the full effect of market risk. In other words, these firms were guilty of exactly what was presented in the asset portion of Figure 5. While they did discount expected yields for historical default rates, they did not fully discount their projections to take into account the effective yield that would be consistent with contagion risk.[23]

Some of the companies would check for sustainability of cash flows in the presence of higher default and surrender rates, but by not modeling asset prices, they never checked for solvency. They were depending on the statutory accounting provisions of maintaining book values on assets, thereby perpetuating the fiction of solvency. But there is clearly a difference between statutory solvency and economic solvency, and insurance regulators have the right to take over or shut down a company even when it is technically solvent from a statutory point of view. If regulators do not act quickly to conserve funds when a company is economically insolvent, the massive policyholder withdrawals that typically ensue will shortly transform economic insolvency into statutory insolvency, as more and more embedded losses on the company assets become realized upon sale to meet policyholder withdrawal demands.

In one prominent example of this problem the downgrade of a large junk bond portfolio triggered a "run on the bank." Even though the insurance company could have

[23] Contagion risk is the risk that an entire class of securities or policies will change in value in a highly correlated fashion. For example, following the Russian debt collapse, low rated bonds declined in value as a category.

ultimately met all of its cash flow needs, absent any policy lapses, the temporary decline in value of its junk bond portfolio led to massive lapses. These lapses, in turn, created a liquidity crunch and the insolvency of the insurance company.

In two other companies we studied the companies were exposed to significant interest rate risk despite the fact that they claimed to be cash flow matching to manage interest rate risk. Even a small change in the level of interest rates (125 basis points in the first case, and 43 basis points in the second case) could either double or entirely wipe out all of their true equity value, depending on the direction of the interest rate move. The cash flow projections did not take into account the optionality inherent in insurance assets and liabilities. Matching cash flows is equivalent to matching dollar duration but effective duration will not be matched if there is optionality, including default risk.

Covering All the Basis

Some companies were very careful to limit their interest rate risk to low levels, following a disciplined approach toward effective duration management. We found, however, that the claim that effective duration is being managed can still be misleading. Effective duration captures interest rate risk. There are other forms of risk that may not be adequately captured by effective duration measures. One example of basis risk that might exist beyond what is measured by effective duration is the liquidity spread associated with securities that are rarely traded.

For example, in an effort to achieve greater statutory earnings, while ignoring the true underlying economics of their strategy, two life insurers offered short-term annuities and Guaranteed Investment Contracts (GICS) with crediting rates that reset every 90 days to 6 months. These liabilities also allowed for book value surrenders upon 7 days notice. Because crediting rates were tied tightly to short-term market rates of interest, there was little incentive to exercise the 7-day book value surrender option.

Assets were of low credit risk, and typically featured floating interest rates with 6-month resets that provided a satisfactory spread over liability crediting rates. Even if there were massive withdrawal requests, so the thinking went, because the asset durations were only a few months long, there would be only small losses on their sale. It seemed that the short-term, low credit-risk, nature of the assets made them a good match to the issued liabilities. The problem arose because a large portion of the assets were very illiquid. So, while the interest rate sensitivities of the assets and liabilities were well matched, the illiquidity of the assets was a poor match to the short term nature of liabilities with a 7 day notice put option embedded.

While it was true that the effective duration of assets with respect to changes in the general level of interest rates was only a few months long, the effective duration with respect to a change in basis, or spread above Treasury rates, that the assets provided was very long—in some cases 16 years or more. In other words, when the assets' credit quality changed slightly, the credit quality fell below the level required by the liability holders. Thus, when the spread widened appreciably a few years ago, the assets plunged in value, even through their coupons were indexed to general interest rate levels and their credit ratings were unchanged. Massive surrenders ensued and the companies became insolvent.

Realize the Gains, Retain the Losses

Under SAP, assets are often shown at book value. Prior to recent reforms, this was the practice with almost all assets; now, it applies to a portion of assets determined by company intention. However, in the past when an asset was liquidated, a capital gain or loss used to flow directly and instantly to surplus. Now, the flow has slowed down considerably, because gains or losses are spread over many years.

While there is less gaming going on today due to the changed SAP treatment, a lot of struggling companies engaged in the practice of systematically selling only assets with capital gains, while holding onto the losses, in an effort to prop up the statutory surplus position. Of course, this does not improve the economic position of the insurer engaged in such practices one iota (and may even accelerate taxes), but it does allow the insurer to "remain at the gambling table" and continue to "bet the company."

Several of the companies mentioned above held a sizable portion of their portfolio in very liquid securities of short duration so that demands for cash could be met without a major tax or surplus effect. However, when the companies got into trouble, this most liquid portion of the asset portfolio was sold, leaving the companies with only depreciated and illiquid assets. Any sale of these less liquid assets at their embedded substantial capital losses would trigger a surplus event and result in conservatorship or insolvency.

Chasing Hot Money

Some of the aforementioned companies pursued high growth goals at high economic costs. The growth came because costs associated with optionality were underestimated and the products were under priced relative to true economic values. The easiest way to grow rapidly, however, is to attract "hot money." Such money is readily available in plentiful supply and easy to attract by offering a combination of above-market yields and cheap exit provisions. Or, in other words, hot money is easy to attract by changing the assumptions used to price a policy in such a way that the policy is priced below competing products in the market.

The reason it is called "hot money" is that it chases yield and is easy come, easy go. Some companies would chase hot money by having their products distributed through the "hottest" channels—Wall Street brokers or managing general agents, instead of the stodgier channels such as direct sale or exclusive agents. With these "hotter" channels, the investors would be kept abreast continually regarding alternative products and available yields, and the business was subject to much higher potential surrender rates.

As discussed in connection with Figure 5, promising above-market yields makes it expensive to attract hot money, and the cheap exit provisions make the effective, option-adjusted yield even costlier. An insurer can attract the hot money easily, but it cannot later lower crediting rates and expect to retain it, for it will depart in search of higher yields elsewhere. It will also depart at the first signs of financial distress. Key to its ability to do this is low or no surrender charges. We found several companies offering deferred annuities with 0 percent surrender charges in each year, or with scheduled surrender charges above 0 percent, but with free surrender with a 30-day window surrounding policy anniversaries or when rates were reset.

These generous surrender provisions forced the insurers to maintain crediting rates that were relatively high, or else the hot money would leave in search of other venues. The companies often supported their hot money liabilities, as well as their other liabilities, with junk bonds and high-yielding mortgages. But when a dip in junk bond prices occurred, the hot money clients removed the full amount of their funds at par value, leaving in greater jeopardy the remaining policyholders and precipitating a "run on the bank." It is difficult to say whether the hot money strategy necessitated risky asset strategies or if the desire to pursue risky investment strategies caused firms to issue hot money liabilities. But, in either case the firm was placed in a riskier position.

Some of the companies we reviewed were basing their rates and profitability projections on unrealistically favorable assumptions regarding mortality and morbidity rates. While these could not be justified based on past history, and were more favorable than realizations proved to be, they seemed to be selected in some cases based on what kind of assumptions would be required to produce favorable projections to regulators and reduced book values of liabilities. These assumptions also made products appear less expensive to consumers and led to increased sales.

As with mortality, morbidity, or investment yields, the cost of liabilities to an insurer is also closely related to surrender and lapse rates. The effective cost of the options an insurer offers implicitly in its policies is higher when the options are more valuable to policyholders. These higher costs do not show up on a statutory basis until the options are exercised. However, in projections of profitability and surplus shown to regulators, the business can be made to look better by selecting surrender and lapse assumptions that are favorable to the insurer. Insurers can exercise a wide degree of latitude in these matters, and those who are inclined to manipulate this latitude to cast their firms in the best possible lights face few restrictions.

Discounting the Reserves

The discount rates applicable to insurance liabilities typically face restrictions by the regulators. However, it is generally the case that regulators allow a single discount to be used against a block of liabilities. While this is administratively convenient, the problem with this practice is that it is simply incorrect—that is, the value of a liability cannot be properly assessed using such methods.[24] Rather, an interest-sensitive method that can take into account the value of options granted to policyholders must be used. An insurer can "game" this system by offering policies replete with generous options to the policyholders, yet have the economic value of its liabilities understated by discounting them with a single rate that does not capture the value of those options. Among the aforementioned companies that went insolvent, we saw several whose liability values were grossly understated because of this discounting practice. It may be argued that statutory valuation methods are conservatively biased in that they assume worst case behavior—e.g., the higher of surrender value or the present value of future cash accumulations—but the degree of conservatism is different when options are present, and hence, the amount of the buffer can be inadequate.

[24] See Babbel and Merrill (1998, 2000) and Babbel, Gold, and Merrill (2002).

Turning Lead Into Gold

Because of the way the asset valuation reserves (AVR, and formerly the Mandatory Securities Valuation Reserves, or MSVR) are calculated, where the reserving penalty is assessed against individual securities using a fairly arbitrary penalty function and proxy for default risk (i.e., agency ratings), rather than on a portfolio based on total risk, the system is open to gaming by the insurers. For instance, a penalty of 20 percent is typically assessed against junk bond holdings based on the Standard & Poor's, Moody's, or Securities Valuation Office rating of each bond held. A bond rated "CCC" will be assessed this penalty and an insurer must hold 20 percent of its face value as a buffer against possible default.

Insurers have learned, with the help of their Wall Street financial engineering friends, how to avoid much of the valuation reserve penalty and the buffer against default that it provided. Some of the firms we mentioned were taking their junk bond portfolios and putting them into a new structured note vehicle whereby they would assign tranches to the cash flows, much like is done with collateralized mortgage obligations (CMO) or their bond counterparts (CBO). Clearly, the highest tranches have first rights to any cash flows until their quota is satisfied, so they are least likely to default. The highest tranches then receive a triple-A rating. Tranches below that may receive a double-A or single-A rating, and are therefore assessed only minimum reserve penalties. The lowest tranches are assessed the maximum penalty of 20 percent. Thus, an insurer, maintaining possession of virtually the entire junk bond portfolio by owning all of the tranches, is able to get by with an asset valuation reserve of only 4–5 percent instead of the 20 percent that would have been assessed against it for holding the exactly the same bond portfolio, but without the magic of "structuring." Of course, nothing about the portfolio structuring has changed its underlying risk, yet the insurer has managed to subvert the aim of the AVR safety program. This practice will inflate the insurer's statutory surplus and allow it to grow more rapidly than otherwise, or stay operating longer than may be safe.

CONCLUSION

In this article we have described situations in which insurance companies have incentives to take on more risk than might be desirable to shareholders or regulators. In addition, we have identified methods that might be employed to create illusions of value while actually taking on more risk or even destroying value. We then provided a series of anecdotes illustrating the serious consequences of failing to maximize true economic value. These insights should be of interest to management, directors, regulators, and investors.

REFERENCES

Babbel, D. F., 1983, Duration and the Term Structure of Interest Rate Volatility, in: *Innovations in Bond Portfolio Management: Duration Analysis and Immunization* (New York: JAI Press).

Babbel, D. F., 1999, Effective and Ineffective Duration Measures for Life Insurance, *Investment Management for Insurers* (New Hope, PA: Frank J. Fabozzi Associates).

Babbel, D. F., and C. Merrill, 1998, Economic Valuation Models for Insurers, *North American Actuarial Journal*, 2(3): 97-144.

Babbel, D. F., and C. Merrill, 2000, *Valuation of Interest-Sensitive Financial Instruments* (revised ed.) (New Hope, PA: Society of Actuaries, Frank J. Fabozzi Associates).

Babbel, D. F., J. Gold, and C. Merrill, 2002, Fair Value of Liabilities: The Financial Economics Perspective, *North American Actuarial Journal*, 6(1).

Babbel, D. F., C. Merrill, and W. Panning, 1997, Default Risk and the Effective Duration of Bonds, *Financial Analysts Journal*, 33(1): 35-44.

Babbel, D. F., and K. Staking, 1989, The Market Reward for Insurers That Practice Asset/Liability Management, *Financial Institutions Research* (New York: Goldman Sachs).

Babbel, D. F., and K. Staking, 1991, It Pays to Practice A/L M, *Best's Review, Property/Casualty Edition*, May.

Babbel, D. F., and R. Stricker, 1989, Asset/Liability Management for Insurers, *Insurance Perspectives* (New York: Goldman Sachs).

Colquitt, L. L., and R. E. Hoyt, 1997, Determinants of Corporate Hedging Behavior: Evidence From the Life Insurance Industry. *Journal of Risk and Insurance*, 64(4): 649-671.

Dionne, G., and S. E. Harrington, 1992, An Introduction to Insurance Economics, in: *Foundations of Insurance Economics: Readings in Economics and Finance* (Kluwer).

Fama, E. F., and M. C. Jensen, 1983, Separation of Ownership and Control, *Journal of Law and Economics*, 26: 301-326.

Jensen, M. C., and W. H. Meckling, 1976, Theory of the Firm: Managerial Behavior Agency Costs and Ownership Structure, *Journal of Financial Economics*, 3(4): 305-360.

Mayers, D., and C. W. Smith, Jr., 1981, Contractual Provisions, Organizational Structure, and Conflict Control in Insurance Markets, *Journal of Business*, 54(3): 407-434.

Mayers, D., and C. W. Smith, Jr., 1994, Managerial Discretion, Regulation, and Stock Insurer Ownership Structure, *Journal of Risk and Insurance*, 64(4): 638-655.

Merton, R. C., 1974, On the Pricing of Corporate Debt: The Risk Structure of Interest Rates, *Journal of Finance*, 29(2): 449-470.

Merton, R. C., 1977, An Analytic Derivation of the Cost of Deposit Insurance and Loan Guarantees: An Application of Modern Option Pricing Theory, *Journal of Banking and Finance*, 1(1): 3-11.

Phillips, R. D., J. D. Cummins, and F. Allen, 1998, Financial Pricing of Insurance in the Multiple-Line Insurance Company, *Journal of Risk and Insurance*, 65(4): 597-636.

Santomero, A. M., and D. F. Babbel, 1997, Financial Risk Management by Insurers: An Analysis of the Process, *Journal of Risk and Insurance*, 64(2): 231-270.

Smith, C. W., and R. M. Stulz, 1985, The Determinants of Firm's Hedging Policies, *Journal of Financial and Quantitative Analysis*, 20(4): 391-405.

Staking, K., and D. F. Babbel, 1995, The Relation Between Capital Structure, Interest Rate Sensitivity, and Market Value in the Property-Liability Insurance Industry, *Journal of Risk and Insurance*, 62(4): 690-718.

Staking, K., and D. F. Babbel, 1998, Insurer Surplus Duration and Market Value Revisited, *Journal of Risk and Insurance*, 65(4): 739-743.