

RISK MANAGEMENT, CAPITAL BUDGETING, AND CAPITAL STRUCTURE POLICY FOR INSURERS AND REINSURERS

Kenneth A. Froot

ABSTRACT

This article builds on Froot and Stein in developing a framework for analyzing the risk allocation, capital budgeting, and capital structure decisions facing insurers and reinsurers. The model incorporates three key features: (i) value-maximizing insurers and reinsurers face product-market as well as capital-market imperfections that give rise to well-founded concerns with risk management and capital allocation; (ii) some, but not all, of the risks they face can be frictionlessly hedged in the capital market; and (iii) the distribution of their cash flows may be asymmetric, which alters the demand for underwriting and hedging. We show these features result in a three-factor model that determines the optimal pricing and allocation of risk and capital structure of the firm. This approach allows us to integrate these features into: (i) the pricing of risky investment, underwriting, reinsurance, and hedging; and (ii) the allocation of risk across all of these opportunities, and the optimal amount of surplus capital held by the firm.

INTRODUCTION

The cost of bearing risk is a crucial concept for any corporation. Most financial policy decisions, whether they concern capital structure, dividends, capital allocation, capital budgeting, or investment and hedging policies, revolve around the benefits and costs of a corporation holding risk. The costs are particularly important for financial service firms, where the origination and warehousing of risk constitutes the core of value added. And because financial service firms turn their assets more frequently than nonfinancial firms, risk pricing and repricing are needed more frequently.

Insurers and reinsurers are an important class of financial firm. They encounter financial risk in their underwriting and reinsurance portfolios, as well as in their investment

Kenneth A. Froot is the André R. Jakurski Professor of Business Administration, Harvard University and NBER, Harvard Business School, Boston, MA 02163, USA. The author can be contacted via e-mail: kfroot@hbs.edu. The article was presented as the Geneva Risk Lecture at the 30th Seminar of the European Group of Risk and Insurance Economists, Zurich. I thank two anonymous referees and the editor, Rob Bentley, John Major, David Moss, Ryan Ogaard, Gary Venter, and especially Paul Embrechts and other participants of the EGRIE seminar for their comments and suggestions. Naturally, all mistakes are mine.

and hedge portfolios. Often they have a large number of clients who view their insurance contracts as financially large and important claims. Two basic features make insurers and reinsurers especially sensitive to the costs of holding risk.

The first feature is that customers—especially retail policyholders—face contractual performance risks. Premiums are paid, and thereafter policyholders worry whether their future policy claims will be honored swiftly and fully. Customers are thought to be more risk averse to these product performance issues than are bondholders. There are several mechanisms for this greater risk sensitivity, some behavioral and others rational.

The behaviorist story is formalized in Wakker, Thaler, and Tversky (1997). They argue that “probabilistic insurance”—the name given by Kahneman and Tversky (1979) to an insurance contract that sometimes fails to pay the contractually legitimate claims of the insured—is deeply discounted by an expected utility maximizer who pays an actuarially fair premium. They also provide survey evidence that the discount is striking in size: in comparison to a contract with no default risk, a contract with 1 percent independent default risk is priced 20–30 percent lower by survey participants.

The rational argument owes most to Merton (1993, 1995a,b). He argues that customers of financial firms value risk reduction more highly than do investors because customer costs of diversifying are higher. Insurance contracts are informationally complex documents and claims payments often require customer involvement. Buying from a single insurer reduces costs. Furthermore, insurance pays off when the marginal utility of customer wealth is high. If absolute risk aversion is declining in wealth, as many suspect it is, then a customer will be more averse to an insurer’s failure to perform than to a debtor’s failure to perform, even if the performance failure is of equivalent size. Higher costs of diversification and greater impact on utility therefore makes risk aversion higher for customers than investors.

This motivates our assumption that when an insurer’s financial circumstances decline, customer demand falls. Indeed it falls more than investor demand. There is evidence in the literature to support this view of “excess” sensitivity. Phillips, Cummins, and Allen (1998), for example, estimate directly price discounting for probability of insurer default. They find discounting to be 10 times the economic value of the default probability for long-tailed lines and 20 times for short-tailed lines. These numbers are too large to be consistent with capital markets pricing.¹ In this article, we take the excess sensitivity hypothesis at face value and trace out the implications for firm value-maximization.

There is a second feature that makes insurers and reinsurers especially sensitive to the costs of holding risk: they often face negatively asymmetric or skewed distributions of outcomes. Many insurance and reinsurance portfolios contain important exposures

¹ Several papers rely on customer preferences to drive insurance pricing, some interacting with the capital markets. See Zanjani (2002), Cummins and Danzon (1997), Cummins and Sommer (1996), Taylor (1994), and Hoerger, Sloan, and Hassan (1990). In addition, Sommer (1996), Grace, Klein, and Kleindorfer (2001), Grace et al. (2003), and Epermanis and Harrington (2001), contain different types of evidence consistent with the hypothesis that insurers suffer from reduced demand when credit ratings fall.

to catastrophic risks—natural and man-made perils of sufficient size and scope to generate correlated losses across large numbers of contracts and policies. Such events can be particularly damaging to insurers and reinsurers and can be expected to have first-order effects on risk allocation, pricing, and capital structure decisions.

Surprisingly, there has been relatively little work that demonstrates how the shape of the payoff distribution affects key decisions.² Perhaps one reason is that an infinite number of moments is required to fully describe a distribution's shape. Another is that higher-order moments appear only rarely in the capital markets literature. This is because discrete-time asymmetric distributions can be derived from continuous-time normally distributed innovations. We adopt here a functional approach to describing asymmetric distributions, treating them as functional transformations of symmetrically distributed normals. This allows us to generate general pricing and allocation results without having to specify moments. In this way, this article takes a step toward explicitly incorporating asymmetrically distributed risks in formal corporate pricing and allocation metrics.

But neither the sensitivity of customer demand nor the asymmetric nature of underwriting exposures can invalidate the Modigliani–Miller irrelevance theorems.³ A firm with perfect access to capital could fund investment opportunities using any combination of instruments. It could raise external finance at the same cost regardless of the strength of its capital base. It could not add to its market value by managing its risk, holding different amounts of capital, or pricing customer risk exposures differently from the outside capital market. Thus, if M&M is to fail, some form of capital-market imperfection is needed.⁴

The approach we take here follows Froot and Stein (1998). It employs two realistic capital-market imperfections. The first is that internal capital is assessed a “carry” charge. Carry costs imply that, all else equal, an additional dollar of equity capital raises firm market value by less than a dollar. The most straightforward source of carry costs would be corporate income taxation. Taxes on additional interest are zero if the dollar is instead deposited in a mutual fund or other pass-through savings vehicle. Another source is agency costs: perhaps management will not use the dollar in shareholders' best interest. Both sources suggest that the market should place limits on how much equity capital a firm can feasibly raise. Indeed, the market does this.⁵

² Kraus and Litzenberger (1976) and Kozik and Larson (2001) add higher-order moments to the CAPM directly, the latter specifically in the case of insurance exposures. However, even if we were to use this for the capital market's pricing of liquid risks, it does not adjust pricing for that which is relevant inside the firm. See the discussion in the section “Commentary on the Pricing of Asymmetric Distributions” below.

³ See Doherty and Tinic (1982) for an application of the Modigliani and Miller logic to insurance firms.

⁴ There are a large number of articles that help motivate the failure of Modigliani–Miller, appealing to imperfections of various sorts. On the pricing of insurance and reinsurance claims, see, for example, Cummins and Phillips (2000), Garven and Lamm-Tennant (2003), Cagle and Harrington (1995), and Winter (1994). On capital (i.e., risk) allocation see Merton and Perold (1993), Froot and Stein (1998), and Myers and Read (2001).

⁵ See Jaffee and Russell (1997) on this point.

The second imperfection is an “adjustment” cost of capital. This is a standard approach, used in models of financing under asymmetric information (e.g., Myers and Majluf, 1984), where raising external funds becomes expensive when existing capital is low. It provides another motivation—in addition to the product–market and asymmetry concerns above—for the firm to get around paying carry costs by holding *too little* capital. Following Froot, Scharfstein, and Stein (1993), if the firm allows internal funds to run down, it will increasingly have to choose between cutting highly rewarding investments or incurring the high costs of external finance. In the insurance and reinsurance industries, adjustment costs of capital appear most clearly in the aftermath of catastrophic events, when depleted industry capital results in high prices and reduced availability of insurance and reinsurance.⁶

With capital-market imperfections, product-market sensitivity to risk, and exposure asymmetry, the firm exhibits strong valuation effects from risk and capital management. As one might expect, these latter factors tend to make the insurer or reinsurer more conservative in accepting risk, more eager to diversify investment and underwriting exposures, more aggressive in hedging, and more willing to carry costly equity capital. As in Merton and Perold (1993) and Myers and Read (2001), our framework jointly and endogenously determines optimal hedging, capital budgeting, and capital structure policies.⁷

We find that product-market considerations contribute additively to capital-market distortions in reducing the desire to hold risk and to price risk at fair market levels. Product-market considerations also tilt the optimal level of capital and surplus toward higher levels. In addition, we prove the firm will price distributional asymmetries in payoffs which are unpriced in the capital market. Negatively skewed exposures impose costs that in general are higher than those of positively skewed exposures, and this causes firms to seek more aggressive reinsurance and hedging and less aggressive and more diversified underwriting and investment.

We also demonstrate that the capital- and product-market imperfections bias firms toward removing risks. Firms maximize value by removing a risk source completely unless: (i) illiquidity makes the risk costly to trade or (ii) the firm has expertise in that risk source that allows it to outperform. Thus, even though the firm’s hurdle for bearing negatively asymmetric insurance exposures is higher than that of the capital market, insurers nevertheless derive their value by earning returns on insurance exposures that, after acquisition costs, exceed required capital market hurdles. This has the normative implication that financial intermediaries should shed all liquid risks in which they have no ability to outperform and devote their entire risk budgets toward an optimally diversified portfolio in exposures where they have an edge. For insurers specifically, this means warehousing insurance risks, where they arguably have informational advantages, and shedding all others. Their warehoused insurance risks will generally be those that are also illiquid—after all, if any were frictionlessly available, shareholders could directly provide their own capital to back them, and average returns would be competed down to rates required by the capital market. Thus, insurers

⁶ See, for example, Gron (1994a,b), Cummins and Danzon (1997), Froot and O’Connell (1997), and Froot (1999, 2001).

⁷ See Venter and Major (2002) for a survey of techniques as applied to insurance.

also generally should not hold perfectly liquid insurance risks unless they have private information that allows them to cherry pick among them. This is all the more true because of the negative skew in “long”-position insurance returns.

In practice, of course, insurance and reinsurance companies, do not seem to eliminate all liquid exposures. While beyond the scope of this article, the reasons for this might be several. One is that our assumed capital- and product-market imperfections do not exist, or have limited financial impact. Most articles do not dispute the existence of at least some of these imperfections, though their exact specifications are a matter of debate.⁸ A second potential reason is that as financial investors, insurers and reinsurers have a real or perceived ability to outperform capital market hurdles. Realized insurer returns on their investment portfolios probably do not provide evidence that this ability is real, Berkshire Hathaway notwithstanding. This leaves corporate overconfidence concerning capital market investment opportunities as a possible explanation. This is an area with interesting new evidence.⁹

THE MODEL: TIMING AND ASSUMPTIONS

The model, which follows Froot and Stein (1998), has three time periods, 0, 1, and 2. In the first two periods, time 0 and 1, the insurer chooses its capital structure and then makes underwriting and hedging decisions. These two periods highlight the fact that consumers will pay less for an insurance contract written by a firm with low surplus with given risk or by a firm with more negatively asymmetric risk. The last period closes the model by allowing the insurer to raise additional capital after paying (or defaulting on) its losses. At this stage additional frictional costs of raising capital are added.

Time 0: Insurer Capital Structure Decision

At time 0, the insurer chooses how much equity capital, K , to hold. The capital goes toward financing various risks, in both investments and underwriting. There are, however, distortions that make the use of capital expensive. First are the deadweight costs of carrying capital, which can be thought of as arising from corporate income taxes or from agency costs associated with shareholders' imperfect controls over management. These costs make it expensive for a firm to carry large amounts of capital. We summarize these deadweight costs as τK , where τ is the effective “tax” rate. If $\tau = 0$, the accumulation of equity capital would be costless, and the firm would not conserve on its use of equity capital.

Also at time 0, we assume that the insurer inherits a given portfolio of risk exposures. This portfolio results in a time 2 random payoff $Z_P = \mu_P + \varepsilon_P$, where μ_P is a mean and ε_P is a mean-zero disturbance term. The risks arise from the insurer's preexisting portfolio of investments in securities, derivatives, and underwriting risks.¹⁰ The total payoff from the insurer's internal funds in place as of time 0 is $Z_P + (1 - \tau)K$. As discussed below, payoffs are realized in the model at time 2.

⁸ See Kaplan and Zingales (2000).

⁹ See Malmendier and Tate (2005) for a discussion of CEO overconfidence and corporate investment. For a broader survey of behavioral corporate finance see Baker, Ruback, and Wurgler (2006).

¹⁰ In practice, these risks may come in the form of assets, liabilities, or neither.

The payoff Z_p has components from the capital markets as well as the insurance and reinsurance market activities of the insurer. We define these components in such a way so that they are independent. To allow for explicit solutions above, we will also assume that the capital markets portion is normally distributed. We denote the capital market payoffs by $\mu_p^C + \varepsilon_p^C$, distributed normally with mean μ_p^C and variance $(\sigma_p^C)^2$.

The normality assumption is convenient for deriving explicit solutions below. But it is not very realistic for the skewed payoffs typical of insurance and reinsurance positions. Consequently, the insurance disturbance term can be nonnormally distributed. The insurance payoff is given by $\mu_p^I + f(\varepsilon_p^I)$, where $f(\varepsilon_p^I)$ is a functional transformation of ε_p^I , a normally distributed random variable with mean zero and variance $(\sigma_p^I)^2$. We assume that f is at least twice differentiable and that it satisfies $E[f(\varepsilon_p^I)] = 0$ and $E[f'(\varepsilon_p^I)] = 1$, i.e., that the transformation of the underlying normal alters neither the mean nor the average exposure to ε_p^I .

This transformation is flexible for describing option-like and other asymmetric payoffs based on a symmetrically distributed underlying variable. For negatively skewed payoffs, such as those on the underwriting book, it is sufficient to require $f' > 1$ for all $\varepsilon_p^I < 0$. For positively skewed payoffs, such as those on a reinsurance portfolio, a sufficient condition would be that $f' < 1$ for all $\varepsilon_p^I < 0$.^{11,12}

The mean payoff from the insurance market component, μ_p^I , differs from the capital markets component, in that μ_p^I contains an endogenous component while μ_p^C is for convenience fixed. Different time 0 choices of risk affect the product-market opportunities through μ_p^I . Thus, it represents the expected return on an ongoing business, rather than the return on a particular portfolio of policies in force at time 0. We assume that increases in firm riskiness reduce μ_p^I . This can be interpreted as due to some combination of behavioral and rational factors mentioned above. We represent the sensitivity of the preexisting product-market opportunities by writing $\mu_p^I = \mu_p^I(s^2)$, where $s^2 = (\frac{\sigma}{K})^2$ measures the squared standard deviation of company-wide risk per unit of capital. An increase in risk/capital—"effective leverage"—reduces the expected opportunity set between time 0 and time 2, i.e., $\frac{d\mu_p^I}{d(s^2)} = \mu_p^I' < 0$.

¹¹ The function f can be thought of as mapping the payoff of an underlying normally distributed asset into the value of an appropriate number of derivative contracts. For example, if the assets of an unlevered company provide payoffs ε_p^I , then the equity of that company is represented by $f(\varepsilon_p^I) = \varepsilon_p^I$, i.e., f is an affine transform with slope of one, so $f(\varepsilon_p^I)$ is distributed normally. Alternatively, ε_p^I could represent the payoffs on a share of the company's equity, and $f(\varepsilon_p^I)$ might represent an amount of cash plus an appropriate number of, say, call options written on the company's stock $f(\varepsilon_p^I) = \Omega \max(\varepsilon_p^I - S, 0)$, where S is the strike price of the option and $\Omega = \Omega(S)$ is the number of options that ensures the average exposure to the stock is unity, $E[f'(\varepsilon_p^I)] = 1$. The cash amount is to ensure that the cash plus options have equal value to the shares, $E[f(\varepsilon_p^I)] = E[\varepsilon_p^I]$. Clearly, the payoff distribution of the cash plus options resembles a standard call option payoff, is positively asymmetric, and therefore not normal.

¹² There is no exact correspondence between the f function and the third moment of its distribution, skewness. The distribution of f defines all higher-order moments, not just the third. Thus, in the text below, we often refer to the distribution of f as negatively (or positively) *asymmetric*, rather than negatively (or positively) skewed, to be clear that we are not merely referring to the behavior of the third moment.

The existing book of insurance underwriting should be interpreted as a multiyear portfolio of insurance, even though each contract requires annual renewal. As the risk of the insurer changes, premiums previously paid for coverage on this book cannot change, of course. However, our assumption is that the premium rate on renewals declines with company-wide increases in leverage. That is, we interpret the underwriting book in place to be an ongoing business.

Naturally, higher leverage increases the risk of default. By increasing both risk and the probability of default, managers can transfer value from policyholders to shareholders. This generates risk seeking behavior on the part of management and shareholders. Following Froot and Stein (1998), we rule this out by assuming that risk levels are not so great as to eliminate completely shareholder payoffs. Policyholders therefore dislike risk because it suggests that the insurer will scrimp on investments in policyholder services, increase the hassles of receiving claims, and make policyholders more likely to incur the deadweight costs of switching insurers. For simplicity then, policyholders lose value to deadweight costs, not to shareholders.¹³

To summarize the set-up at time 0: The insurer chooses an amount of cash capital, K , which is entirely equity-financed. Cash capital results in a deadweight cost of $K\tau$. This cost forms a wedge between the value of the assets inside the firm and the value of the firm in the marketplace. The insurer also holds a portfolio of exposures with net payoffs of $Z_P = \mu_P + \varepsilon_P = \mu_P^C + \mu_P^I(s^2) + \varepsilon_P^C + f(\varepsilon_P^I)$, where the two components of ε_P , ε_P^C and ε_P^I , are defined so as to be independent, mean-zero normals. Because the premiums available in customer underwriting decrease with increases in firm-wide risk, $\mu_P^I < 0$, the mean return on Z_P is endogenous to the choice of K .

Notice that the deadweight costs of capital thus far come from the stock of capital, K , and not from any costs of adjusting the capital structure. This is different from most standard models of financing under asymmetric information. In those models, costs are generally incurred when raising incremental external funds, not by having equity capital on the balance sheet *per se*. Indeed, we assume below that there are exactly such flow costs of new external finance at time 2.

The absence of any cost to raising capital at time 0 is clearly a shortcut. Nevertheless, it allows us to focus better on the appropriate long-run "target" level of capital for the insurer. Over long periods of time (over which adjustment costs can best be amortized), what kind of ratings standard is best for the insurer or reinsurer? How should it position itself in that regard? We recognize that, if at any point in time, the insurer is far away from its ideal target, it may face costs of adjustment in getting to the target quickly, but it is nonetheless interesting to ask the question of what the target should be.

Time 1: Insurer Underwriting and Hedging Decisions

At time 1, the insurer makes additional decisions. It decides how much new investment and underwriting to undertake and how much hedging or reinsurance to invoke. We treat investment, underwriting, hedging, and reinsurance in a single

¹³ See Hall (1999) for evidence on the magnitude of apparent risk seeking financial decisions in U.S. insurers.

new-opportunity specification. As above, the j th new opportunity has an expected and unexpected component, given by $Z_{N,j} = \mu_{N,j} + \varepsilon_{N,j}$, where $\mu_{N,j}$ is the mean payoff, and $\varepsilon_{N,j}$ is a mean-zero normally disturbance term. Each opportunity can represent a new underwriting opportunity, a reinsurance contract, a capital markets hedging vehicle, or any other self-financing new investment.

As above, it is useful to think of each opportunity as being comprised of two independently distributed parts, a capital market and an insurance-market component. In what follows, we treat the capital market component as fully liquid and costlessly traded. As a result, it will always trade at "fair" prices. We also assume that the insurance market exposures are illiquid, and allow for the possibility that they are mispriced, offering returns that are not fair based on their capital market risks. This latter group of exposures could easily include not only traditional insurance and reinsurance risks, but also credit exposures. Indeed, any risk with a component orthogonal to the major capital markets risks could be considered as an insurance market exposure rather than a capital market exposure. All that matters for our analysis is that the market-wide factors priced by the capital markets trade costlessly and at fair prices.

To see all this more precisely, we additively decompose the total payoff from the j th new opportunity into the independent capital and insurance market exposures:

$$Z_{N,j} = \mu_{N,j} + \varepsilon_{N,j} = (\mu_{N,j}^I + \mu_{N,j}^C) + (\varepsilon_{N,j}^I + \varepsilon_{N,j}^C),$$

where the superscripts I and C denote, respectively, insurance and capital market exposures. We assume that the j th new-opportunity disturbances, $\varepsilon_{N,j}^I$ and $\varepsilon_{N,j}^C$, are independently distributed mean-zero normals.¹⁴ The combined disturbance, $\varepsilon_{N,j}$, may contain any combination of insurance and capital market components.

The magnitude of the insurer's exposure to the new opportunity is a choice variable, given by n_j . Examples of n_j would be the number of policies written, or the number of futures or reinsurance contracts used to hedge, or the number of credit default swaps purchased or written, etc.; even a new business opportunity can be described in this way. In other words, n_j is meant to cover essentially any new risk decision that the firm contemplates. Thus, the total payoff from the new underwriting, hedging, and investment opportunities is $nZ_N = n(\mu_N + \varepsilon_N) = \sum_j n_j(\mu_{N,j}^I + \mu_{N,j}^C + \varepsilon_{N,j}^I + \varepsilon_{N,j}^C)$.

The insurer's realized internal wealth at time 2, w , is therefore given by:

$$w = Z_P + nZ_N + K(1 - \tau). \quad (1)$$

In words, the amount of cash the insurer has on hand at time 2 to pay claims will depend on the realizations on its preexisting capital market and insurance exposures,

¹⁴ Normality is used to derive explicit solutions for the choice variables. However, our approach allows for more generality in the distributions of the disturbances and we comment on this below.

Z_P , on the extent of its new underwriting, investment, and hedging outcomes, nZ_N , and on the amount of capital, K , raised at time 0.¹⁵

The sensitivity of underwriting opportunities to firm-wide risk levels creates a meaningful need for risk management for the value-maximizing insurer. To see this, suppose the insurer finds an underwriting opportunity where, loosely speaking, premiums are high relative to risk. In this circumstance, the insurer will choose to underwrite, setting $n_j > 0$. However, as n_j grows, internal risk increases, raising s^2 and reducing premiums on *all* existing underwriting by $\mu_P^I d(s^2)$. All else equal, this product-market externality lowers expected internal funds, $E[w]$, and raises internal hurdle rates. The underwriting opportunity itself and other risky investments appear less attractive, while costly hedging opportunities and balance-sheet capital appear more attractive.

Lastly, we note that this specification of $\mu_P^I = \mu_P^I(s^2)$ need not be thought of as restricted to just the existing underwriting business. It is likely that higher s^2 would degrade future product market opportunities, and not simply underwriting already in place. In a sense, s^2 and its impact on expected underwriting performance can be viewed as partly coming through a reputational channel. Thus, if these effects—both on business in place and on future business through reputation—are important, then the magnitude of $\mu_P^I d(s^2)$ may be large if measured as a fraction of business that is already in place.

Time 2: The Realization of Cash Flow, and the Insurer Response

Next we specify how the insurer uses internal funds to create value. The simplest possibility is to assume that the insurer terminates its operations at time 2, dividending back all realized internal funds, w , to investors. In that case, the *ex post* value of the firm, denoted by P , equals the realization of w , so that $P = w$.

Insurers and reinsurers are likely to perceive risk as costly because it degrades capital market opportunities as well as product market opportunities. That is, if internal capital becomes riskier, the perceived cost of raising external funds increases. This is the theme in much of the literature on costly external capital markets (see, e.g., Greenwald, Levinson, and Stiglitz, 1991). Our formulation follows Froot, Scharfstein, and Stein (1993). We assume that after w is realized, the company has a further investment opportunity—e.g., it might be able to build a new interface with its customers, open an entirely new line of business, buy an existing business, etc. This investment requires a cash commitment of I , and yields a gross return of $H(I)$, where $H(I)$ is an increasing, concave function of I . The investment can either be funded out of internal sources, or external sources in an amount e . Thus, $I = w + e$. The critical point is that there are convex costs to raising external finance, given by $C(e)$. This means that it

¹⁵ Note that in our specification, the insurer's risk level impacts expected internal funds only through $\mu_P(s^2)$; μ_N is not a similar function of squared risk per unit of capital. One might expect that additional risk would affect new opportunities as well as the existing opportunity set. Including this in the specification complicates the expressions, but doesn't change importantly any of the results. We therefore leave out the sensitivity of the new opportunity to firm-wide risk.

becomes more costly to raise funds the larger is the amount that must be financed externally.^{16,17}

Denote by $P(w)$ the solution to the insurer's time-2 problem:

$$P(w) = \max_I H(I) - I - C(e), \text{ subject to } I = w + e. \quad (2)$$

Froot, Scharfstein, and Stein (1993) demonstrate that $P(w)$ can be rigorously derived in the context of one standard optimal contracting models as an increasing concave function. Thus, $P_w > 1$ and $P_{ww} < 0$.

The concavity of the $P(w)$ function generates a capital markets rationale for insurer risk management. This concavity in turn arises from two sources. First is the convexity of $C(e)$, which matters to the extent that fluctuations in internal cash result in fluctuations in additional costs of raising external funds. Second is the concavity of $H(I)$, which matters to the extent that fluctuations in internal cash result in fluctuations in investment, lowering the average return on investment. Loosely speaking, the more difficult it is for the insurer to raise external funds on short notice at time 2, the more averse it will be to fluctuations in its time 2 internal wealth w . Thus, fund raising opportunities in the capital markets become less appealing as internal funds become increasingly unpredictable.

The derivative P_w summarizes the *ex post* value of having an additional dollar of internal funds. The value $E[P_w]$ summarizes the *ex ante* value of this dollar. Naturally, a property of the model solution must be that $E[P_w] \geq 1$: an additional dollar of internal wealth has expected *ex ante* value of at least one dollar. To see this, note that even if there were no costs of carrying capital ($\tau = 0$), a value-maximizing firm would never raise a dollar of external funds to increase firm value by less than a dollar. In the presence of costs of carrying capital ($\tau > 0$), the inequality will be strict, i.e., $E[P_w] > 1$.

ANALYSIS

To solve the model, we work backwards. We have already seen that any given realization of w at time 2 can be mapped into a payoff, $P(w)$. Next we ask from the perspective of time 1, when w is still uncertain, what amount of new investment, underwriting, hedging, and reinsuring will maximize expected market value? Then we move back

¹⁶ Froot, Scharfstein, and Stein (1993) give a number of microeconomic rationales—based on agency and/or information problems—to justify this sort of specification for the $C(E)$ function. They show how this convex functional form arises in a specific optimal contracting setting, a variant of the costly state verification model due to Townsend (1979) and Gale and Hellwig (1985). Stein (1996) generates a similar formulation in a banking model where nondeposit liabilities are subject to adverse selection problems.

¹⁷ It is straightforward to allow the insurer's gross return on investment to be sensitive to product-market perceptions about time 2 risk and time 0 capital. For example, we could specify that $H = H(I, s^2)$, where $H_s < 0$, $H_{ss} > 0$. This would make more explicit the reputational channel through which past risk and capital behavior affect future product market opportunities.

to time 0 to solve for the right amount of capital given the costs of carrying it, and the costs of potentially having too little capital at time 2, thereby necessitating cutting profitable investment or raising costly external financing.

Shareholder Valuation of the Insurer at Time 1

From the perspective of time 1, the shareholder payoff, $P(w)$, is a random variable. In order to value the payoff, we need a pricing model. We assume that the company's equity trades costlessly and is priced fairly. Specifically, we assume that the fair pricing rule has required returns as a linear, decreasing function of covariance with a "market" factor, M . It is straightforward to generalize this to a multifactor setting, where covariances with many factors determine fair-market required returns. With just a single factor, the present value of the insurer's shares V , will be:

$$V = \{E[P(w)] - \gamma \text{cov}(P(w), M)\} \quad (3)$$

where, for simplicity, the riskless rate of interest is zero.

First, we value the pure capital market exposures that are contained in both Z_P and Z_N . Expected returns for these exposures, μ_P^C and μ_N^C , will be determined by the pricing model in Equation (3), because liquid exposures are priced fairly and have no transaction costs. Fair pricing implies simply that $\mu_P^C = \gamma \text{cov}(\varepsilon_P^C, M)$ and $\mu_N^C = \gamma \text{cov}(\varepsilon_N^C, M)$.

Pricing these liquid exposures helps clarify our assumptions about the expected return on underwriting opportunities, $\mu_P^I = \mu_P^I(s^2)$. If fair market pricing held for μ_P^I , then it would be a function solely of market covariance as in Equation (3), and not a function of firm-wide risk, s^2 . This would occur if capital markets investors had direct, costless access to insurance exposures, since competition among them would drive insurance expected returns to fair pricing levels. They do not, however, have such access. These exposures are expensive to underwrite because of the costs of rate evaluation, monitoring, claims evaluation and payment, regulation, etc.

Substituting the definition of fair pricing into Equation (1), we can rewrite time 1 internal funds as:

$$\begin{aligned} w = & \gamma \text{cov} \left(\varepsilon_P^C + \sum_j n_j \varepsilon_{N,j}^C, M \right) + \mu_P^I(s^2) + \sum_j n_j \mu_{N,j}^I + \varepsilon_P^C + f(\varepsilon_P^I) \\ & + \sum_j n_j \varepsilon_{N,j} + K(1 - \tau), \end{aligned} \quad (4)$$

where μ_P^I and $\mu_{N,j}^I$ can be interpreted as expected excess returns, above the fair market, earned on insurance market exposures. Note that we have used the fact that the insurance market shocks are independent of M , thereby setting their covariances to zero. To collapse terms, it is useful to rewrite (4) as:

$$w = \mu + \varepsilon + f(\varepsilon_P^I), \quad (5)$$

where $\mu \equiv \gamma \text{cov}(\varepsilon_P^C + \sum_j n_j \varepsilon_{N,j}^C, M) + \mu_P^I (s^2) + \sum_j n_j \mu_{N,j}^I + K(1 - \tau)$ and $\varepsilon \equiv \varepsilon_P^C + \sum_j n_j (\varepsilon_{N,j})$. Equation (5) has the virtue of separating w into a mean plus a normal disturbance and nonnormal disturbance.

Optimal Hedging Policy at Time 1 for Capital Market Exposures

The insurer designs its risk management policy so as to maximize shareholder value, V . Using Equation (3), this means it will choose n_j at time 1 to satisfy:

$$\frac{dV}{dn_j} = \frac{dE[P(w)]}{dn_j} - \gamma \frac{d\text{cov}(P(w), M)}{dn_j} = 0, \tag{6}$$

where all expectations are taken with respect to the ε 's, and where dn_j represents a change in the quantity of the j th new investment. Taking derivatives and using the definition of covariance, we can write:

$$\frac{dV}{dn_j} = E[P_w]E\left(\frac{dw}{dn_j}\right) + \text{cov}\left(P_w, \frac{dw}{dn_j}\right) - \gamma \frac{d\text{cov}(P(w), M)}{dn_j}. \tag{7}$$

Some algebra, Equation (4), and the fact that $\frac{d(s^2)}{dn_j} = \frac{d\text{var}(\varepsilon + f(\varepsilon_P^I))}{K^2 dn_j} = \frac{2}{K^2} \text{cov}(w, \varepsilon_{N,j})$ implies:

$$\frac{dV}{dn_j} = E[P_w]E\left(\frac{dw}{dn_j}\right) + \text{cov}(P_w, \varepsilon_{N,j}) - \gamma \frac{d\text{cov}(P(w), M)}{dn_j}, \tag{8}$$

where

$$\frac{dw}{dn_j} = \mu_P^I \frac{2}{K^2} \text{cov}(w, \varepsilon_{N,j}) + \mu_{N,j}^I + \gamma \text{cov}(\varepsilon_{N,j}^C, M) + \varepsilon_{N,j}. \tag{9}$$

To simplify this, we need to express the covariances in terms of the underlying random variables. Note first that, by assumption, both $\varepsilon_{N,j}$ and M are normally distributed. This allows us to use a generalization (see Stein, 1981) for normal random variables x, y , and z :

$$\text{cov}(g(x, y), z) = E[g_x] \text{cov}(x, z) + E[g_y] \text{cov}(y, z), \tag{10}$$

where g is a continuous function (subject to some mild regularity conditions) and the expectations are taken with respect to the joint distribution of x and y .

Next, assume that there are in total $J + 2$ new opportunities. The first J describes opportunities which contain combinations of insurance and capital market exposures. The last 2 are reserved for capital market exposures that span all the capital market disturbances.¹⁸ With this in mind, suppose that the $J + 2$ nd new product market opportunity is a costlessly traded instrument with payoffs identical to the market return,

¹⁸ We can add arbitrarily many of such factors, but with no change in generality or important results.

M . Since capital market opportunities are fairly priced and there is no insurance market component, we have that $\varepsilon_{N,J+2} = \varepsilon_{N,J+2}^C = M$ and $\mu_{N,J+2}^I = 0$. The independence of insurance and capital market components means that $\text{cov}(\varepsilon_p^I, \varepsilon_{N,J+2}) = 0$. Using this fact, and Equations (8) and (9), we can write:

$$\frac{dV}{dn_j} = E[P_w] \left(\mu_p^I \frac{2}{K^2} \text{cov}(w, M) \right) + E[P_{ww}] \text{cov}(w, M) - \gamma E \left[P_{ww} \frac{dw}{dn_j} \right] \text{cov}(w, M) = 0, \quad (11)$$

where we substituted M for $\varepsilon_{N,J+2}$, and $\text{cov}(w, M)$ for $\text{cov}(\varepsilon, \varepsilon_{N,J+2})$.

We can then state our first proposition, taking as given the first $J + 1$ new opportunity decisions, n_j :

Proposition 1: *It is optimal for the firm to choose a market-risk hedge such that $\text{cov}(w, M) = 0$. Thus, the optimal hedge is the minimum variance hedge ratio, i.e., $n_{J+2}^* = \frac{-\text{cov}(\varepsilon_p^C + \sum_{j=1}^{J+1} n_j \varepsilon_{N,j}^C, M)}{\text{var}(M)}$.*

To prove this, one can see readily that Equation (11) is satisfied by setting n_j such that $\text{cov}(w, M) = 0$.¹⁹ In other words, the insurer will hedge out all of the market exposure in internal funds, w , minimizing its variance with respect to M . The firm's market value rises because the reduced variance of internal funds improves product market opportunities and lowers the costs of future external finance. The optimal hedge ratio then solves $\text{cov}(w, M) = 0$ for n_{J+2} , with n_{J+2}^* as above. In words, it is optimal to strip market risk entirely from internal funds. Any exposure to M contained in the preexisting portfolio, or in other new investment opportunities, is offset with the $J + 2$ nd hedge instrument, with returns equal to M .

Our next result is analogous: like the market risk, M , other capital markets exposures will be fully hedged, provided that they are fairly priced. To see this, let the $J + 1$ st new opportunity be a hedge of all other capital market exposures that are independent of M . There is no insurance component to the $J + 1$ st new opportunity, so that $\varepsilon_{N,J+1} = \varepsilon_{N,J+1}^C$, $\mu_{N,J+1}^I = 0$, and $\text{cov}(\varepsilon_p^I, \varepsilon_{N,J+1}) = 0$. This and the previous hedge of M implies that the last two terms of Equation (11) disappear.

It is then easy to show:

Proposition 2: *The optimal hedge ratio for capital market exposures sets internal funds to be uncorrelated with the hedge of capital market exposures, $\varepsilon_{N,J+1}^C$, i.e., $\text{cov}(w, \varepsilon_{N,J+1}) = \text{cov}(\varepsilon, \varepsilon_{N,J+1}) = 0$. The implied solution is given by the minimum-variance hedge, $n_{J+1}^* = \frac{-\text{cov}(\varepsilon_p^C + \sum_{j=1}^J n_j \varepsilon_{N,j}, \varepsilon_{N,J+1})}{\text{var}(\varepsilon_{N,J+1})}$.*²⁰

¹⁹ We assume that the second-order conditions are satisfied throughout.

²⁰ Note that since these other capital markets exposures are assumed to be orthogonal to M , the solution for the $J + 1$ st hedge does not depend on n_{J+2}^* . In general, if the capital market exposures are not independent, then the relevant n_j first-order conditions must be solved simultaneously rather than sequentially as above. See the discussion below.

Overall, the hedging of capital market risks strips out those exposures entirely from internal funds. Thus, the expression for internal wealth with capital market hedges in place becomes:

$$w_H = \mu_p^I(s^2) + \sum_j n_j \mu_{N,j}^I + f(\varepsilon_p^I) + \sum_j n_j \varepsilon_{N,j}^I + K(1 - \tau). \tag{12}$$

These results echo those in Froot and Stein (1998): fairly priced exposures will be hedged fully by the firm. Because they raise the variability of internal funds, the costs that such risks impose on the firm will not be fully compensated by the fair market expected return.

Optimal Hedging Policy at Time 1 for Illiquid Exposures

With the capital market risks fully hedged at fair market prices, we are ready to add in new opportunities that contain insurance market exposures, i.e., those portions of risk that may be illiquid. There are two things that distinguish the required return and hedging decisions of these illiquid components from the capital market counterparts that we examined above. The first is that the insurance exposures interact with the nonnormal payoffs already in the portfolio. This means that insurance instruments may be able to alleviate the costs, or enhance the benefits, of the asymmetrically distributed risks already in the firms' portfolio. The second is that these exposures are illiquid, and therefore may have required returns that differ from fair market returns.

We explore first the optimal amount of investment in the new opportunity. We then look to the implied hurdle rates, to understand better the firm's reservation price for new opportunities. We do this below for the *j*th new opportunity, which, unlike in the previous subsection, may contain both capital market and insurance market exposures.

To see the implications of these, we rearrange Equation (11), where the last term has disappeared due to the full hedging of capital market exposures.²¹ We then have one of our two main propositions in the article:

Proposition 3: *The optimal amount of the new opportunity is given by three factors: the excess risk-adjusted return relative to its own variance, adjusted by the firm's tacit risk aversion, $(\frac{1}{F+G})(\frac{\mu_{N,j} - \gamma \text{cov}(\varepsilon_{N,j}^C, M)}{\text{var}(\varepsilon_{N,j}^I)})$; a kind of asymmetry-adjusted covariance of the opportunity relative to own variance, $(\frac{\tilde{G}}{F+G})(\frac{\text{cov}(\varepsilon_p^I, \varepsilon_{N,j}^I)}{\text{var}(\varepsilon_{N,j}^I)})$; and the minimum-variance amount of the opportunity, $\frac{\text{cov}(w_j, \varepsilon_{N,j}^I)}{\text{var}(\varepsilon_{N,j}^I)}$:*

$$n_j^* = \left(\frac{1}{F+G} \right) \left(\frac{\mu_{N,j} - \gamma \text{cov}(\varepsilon_{N,j}^C, M)}{\text{var}(\varepsilon_{N,j}^I)} \right) - \left(\frac{\tilde{G}}{F+G} \right) \left(\frac{\text{cov}(\varepsilon_p^I, \varepsilon_{N,j}^I)}{\text{var}(\varepsilon_{N,j}^I)} \right) - \left(\frac{\text{cov}(w_j, \varepsilon_{N,j}^I)}{\text{var}(\varepsilon_{N,j}^I)} \right),$$

²¹ We also use the fact that with $E[f'] = 1$, $\text{cov}(f(\varepsilon_p^I), \varepsilon_{N,j}^I) = \text{cov}(\varepsilon_p^I, \varepsilon_{N,j}^I)$.

where $w_j = w - n_j \varepsilon_{N,j}^I$ is total internal funds less the disturbance from the j th new opportunity.

Before we interpret the proposition in its entirety, several terms in the above expression require definition. First, $F = -\mu_p' \frac{2}{K^2} > 0$ describes the sensitivity of internal funds to an increase in firm-wide risk that comes through the product-market channel. Essentially, F measures the magnitude of the negative demand externality created by customers who are more sensitive to diversifiable risk than are capital providers. This externality is greatest when leverage is high, that is, when the variance of internal funds is large relative to external capital. If the capital base grows, but firm-wide risk remains constant, μ_p' (and therefore F) moves toward zero. The hypothesis is that customers are concerned about risk not absolutely, but relative to the size of the capital base.²²

The second term, $G = \frac{-E[P_{ww}]}{E[P_w]} > 0$, measures the concavity of firm-wide valuation in the capital markets (see Froot and Stein, 1998). It is similar to absolute risk aversion based on an investor's utility function and is the result of variability in expected future costs of raising external finance. The variability of internal funds not only negatively impacts product-market opportunities through F ; it also degrades future growth opportunities by making external funds on average more expensive. G is a declining function of the realization of w (F is actually a function only of the *ex ante* properties of w). G therefore converges toward zero as w becomes large relative to K .

The terms F and G appear in an additive way because they are complementary channels through which changes in firm-wide risk impacts value. To form a measure of the effective risk aversion of the insurer, both F and G are required. If $F = G = 0$, then the insurer acts as though it is risk neutral, being willing to undertake an infinite amount of a new investment with positive risk-adjusted excess return, $\mu_{N,j} - \gamma \text{cov}(\varepsilon_{N,j}^C, M)$. When $F + G > 0$, the insurer's response to a positive NPV new opportunity is more muted, as long as there is some unhedgeable risk imposed on firm-wide capital as a result (i.e., $\text{var}(\varepsilon_{N,j}^I) > 0$).

The analogy with risk aversion is not exactly correct for two reasons. First, F and G are endogenous to the properties and realization of w ; in many cases investor risk aversion is treated as an exogenous parameter, determined entirely by the form of the utility function. Second, and more importantly, F and G measure the willingness of the insurer to pursue risky returns in excess of *fair value*. By contrast, investor risk aversion measures the willingness to pursue risky returns in excess of the risk free rate. An investor will always want to hold at least a small amount of any risk that has positive expected returns above the risk free rate. A corporation will not wish to hold any risk unless it at least pays a return greater than fair value as determined in the capital markets.

Third, $\tilde{G} = \frac{(E[P_{ww}] - E[P_{ww} f'])}{E[P_w]}$ measures the impact of the payoff asymmetry of preexisting exposures. To see this, suppose that $f(\varepsilon_p^I)$ describes the negatively asymmetric

²² We abstract from any absolute size effects, which would have customers be more sensitive to a given risk/capital ratio for smaller firms. There is probably a strong argument that such size effects can be important, in that there are at least some fixed costs of producing insurance.

payoffs of a typical insurance contract—occasional large negative outcomes offset by far more frequent small positive outcomes. In that case, $E[f'] > 1$ for $\varepsilon_p^I < 0$ and $E[f'] < 1$ for $\varepsilon_p^I > 0$. Given the concavity of $P(w)$, this creates the presumption that $\tilde{G} > 0$. Loosely speaking, the more negatively asymmetric is the preexisting portfolio, the more concave is the value function in expectation; i.e., as the distribution of $f(\cdot)$ becomes more negatively asymmetric, $E[P(w)]$ falls. A positively asymmetric risk will provide the firm the additional benefit of reducing the concavity of the value function, so that $\tilde{G} < 0$. In this way, the asymmetry of preexisting payoffs enhances or reduces the potency of individual new opportunities, depending how they interact with the underlying risk, ε_p^I . A firm will want to hedge to a greater (lesser) extent to reduce the negative (positive) asymmetry of its underlying payoffs.

Indeed, Proposition 3 shows that the firm will hedge a negatively asymmetric exposure *beyond* the minimum variance amount. To see this, consider a new hedging opportunity that pays no return above fair market value, $\mu_{N,j} - \gamma \text{cov}(\varepsilon_{N,j}^C, M) = 0$, so that the first term in Proposition 3 disappears. If the preexisting positions all have normal distributions, the firm will of course want to hedge by an amount that eliminates the preexisting exposure to the risk, $\frac{\text{cov}(w, \varepsilon_{N,j}^I)}{\text{var}(\varepsilon_{N,j}^I)}$. However, if its preexisting payoffs are nonnormal, it will wish to go further, hedging more (less), provided that by doing so it can reduce the negative (increase the positive) asymmetry in its preexisting exposures. The ability to use the hedge to reduce the negative asymmetry in payoffs is given by $\tilde{G} \text{cov}(\varepsilon_p^I, \varepsilon_{N,j}^I)$.

We can take this one step further by looking back to the first-order condition in Equation (8) above. It shows that the firm is actually trying to reduce covariance of each exposure with the shadow value of internal funds, P_w . A perfect set of hedges would completely eliminate uncompensated covariation in P_w . For normally distributed risks, this is equivalent to eliminating the exposure with internal funds, w , itself. However, if the distribution of preexisting exposures is nonnormal, the asymmetry generally interacts with the concavity of $P(w)$. In that case, minimizing the variance of P_w is no longer the same as minimizing the variance of w . The optimal n_j must take into account the new investment's impact on the covariation of P_w by altering the asymmetry of w .

Finally, note that, like F and G , \tilde{G} falls to zero as internal funds increase relative to initial capital. Asymmetry in the preexisting payoff distribution is less costly when there is plenty of internal financial slack. Another way to say this is that, when w is large, the shadow value of internal funds, P_w , converges toward 1, and there is little need to manage internal funds. As a result, F , G , and \tilde{G} are all near zero.

Now that we have dissected the optimal new opportunity decision rule, it is straightforward to rearrange the expression to derive new opportunity hurdle rates:

Proposition 4: *The required incremental rate of return on an incremental amount of the j th new opportunity is given by*

$$\mu_{N,j} = \gamma \text{cov}(M, \varepsilon_{N,j}^C) + (F + G) \text{cov}(w, \varepsilon_{N,j}^I) + \tilde{G} \text{cov}(\varepsilon_p^I, \varepsilon_{N,j}^I).$$

Proposition 4 shows that pricing internal risks requires a three-factor model. The first factor is standard—the quantity of market risk, $\text{cov}(M, \varepsilon_{N,j}^C)$, multiplied by the price of market risk, γ .

The second factor is the quantity of firm-wide risk, $\text{cov}(w, \varepsilon_{N,j}^I)$, multiplied by the price of firm-wide risk. Firm-wide risk is also the second factor in the Froot and Stein (1998) two-factor model. This source of risk is also key in related studies, such as that of Myers and Read (2001) and Merton and Perold (1993). However, in the present context, the price of firm-wide risk, $F + G$, is greater than in previous studies. That is because of the additional product-market channel through which firm-wide risk impacts firm value. Notice also that the quantity of firm-wide risk is measured with respect only to the insurance market component of the j th new opportunity. Because the capital market risks are fully and costlessly hedged, there is no need to penalize a new opportunity for containing them.

The third factor is the asymmetry of firm-wide payoffs. It is the product of the quantity of asymmetry risk, $\text{cov}(\varepsilon_P^I, \varepsilon_{N,j}^I)$ —i.e., the covariance of the j th opportunity with the asymmetrically distributed components of internal funds—with the price of asymmetry risk, \tilde{G} . It is worth reemphasizing that our definition of asymmetry does not correspond precisely with skewness, or any particular group of moments of the asymmetric payoff distribution. The price of asymmetry risk comes not from asymmetry *per se*, but from the interaction of the asymmetric distribution with variation in the marginal value of internal funds, P_w .

Of course, for both empirical and theoretical reasons, it is sensible to assume that an insurer's or reinsurer's preexisting portfolio will be negatively asymmetric. This creates the presumption that \tilde{G} will be positive for these firms, and probably for financial intermediaries in general.

Commentary on the Pricing of Asymmetric Distributions

It is often argued that much of the standard valuation models in finance are poorly suited for incorporating the pricing of asymmetric payoffs. For example, because investors (and firms) dislike negative skewness, capital market pricing formulae might include a factor based on the skewness or co-skewness of a position. Others in finance rebut this criticism, however, by allowing for continuous-time liquidity and trading with changing first and second moments. In such a case, negatively skewed distributions over discrete time intervals can often be replicated by Brownian motions (distributed normally), with conditional expected returns and variances that evolve over time. Basically, the effects of asymmetries on pricing can often be reproduced by more frequent trading.²³

In our setup, we give the benefit of doubt to the continuous trading argument for a company's equity. (Although for insurance companies, in particular, the possibility of a large, sudden event that discretely diminishes value, like that of an earthquake or terrorist attack, is very real.) Our valuation function, Equation (3), therefore ignores distributional asymmetries in determining the value of an asset based on expected cash flow and its covariance with the market.

²³ This assumes that over the discrete interval, the value path is continuous, so that every value path can be drawn without "lifting the pencil." Gaps in the value path may indeed require additional factors to be priced accurately. Notice, however, that while individual insurance contracts or derivatives may indeed have such value gaps, if each of these is small and reasonably independent, then, at the level of the portfolio, the central limit theorem applies, and internal funds are approximately normally distributed.

However, even after ignoring return asymmetries for the pricing of cash flows in Equation (3), asymmetries matter for determining the deadweight costs facing the firm. Low realized levels of internal funds make future investment opportunities more costly to finance because of deadweight financing costs, $C(e)$. Thus, negatively skewed payoffs will tend to increase the expected deadweight costs of external finance, thereby directly reducing value. This leads to a positive \tilde{G} and a reduced level of $P(w)$, all else equal.²⁴ Asymmetries here affect expected net cash flows, and thereby affect value.

While the continuous trading argument above might plausibly apply to a company's equity, it certainly wouldn't apply to the illiquid insurance exposures that comprise w . Asymmetries in the illiquid insurance exposures can reduce $P(w)$ in our model because, for these risks, a continuous trading argument is not tenable. There is essentially no trading in individual illiquid risks. This is what allows financial intermediaries and, indeed, most nonfinancial firms as well, to add value. They originate and warehouse illiquid assets and exposures, allowing them in the aggregate to be liquidly traded in the form of company stock.

Proposition 4 allows for distributional asymmetries in valuation, but not by including additional factors driven by skewness, or other higher-order moments, *per se*. The problem with a moment-based approach is that there are, of course, an infinite number of moments and the criteria for choosing among them can be rather arbitrary. The advantage of our functional approach is that all of the properties of the asymmetric distribution—in effect all of its moments—are contained in our single third factor.

However, there are also disadvantages to this analytic approach. Chief among them is that we cannot easily allow for nonnormally distributed new opportunities. Our method for deriving explicit solutions—the generalization of Stein's Lemma—can be applied to at most one nonlinear function of normals at a time. Essentially, there is no general analytic solution for the covariance of two nonnormal random variables. Our response is to assume that the new opportunities are distributed normally (and therefore symmetrically). This works, but leads to pricing errors when new opportunities are asymmetrically distributed.

Consider, for example, an insurer exposed to asymmetrically distributed property-casualty losses. An excess-of-loss reinsurance contract (essentially a call spread written on the insurer's underlying event losses) would commonly be used to cede the risk. Such a contract would, of course, have a highly asymmetric payoff distribution. The contract would be *more* valuable to the insurer than a contract whose payoffs were symmetric (and normal) because the excess-of-loss contract would completely cede the insurer's event losses without adding any other risks, something that a contract with symmetric payoffs could not do. By approximating the new opportunity payoffs as normal, our approach therefore understates value if a positively asymmetric new

²⁴ The distributional asymmetry could impact $P(w)$ through the product-market channel as well. For example, if the negative demand shocks associated with greater risk-to-capital ratios negatively affected the return on future investments, $F(I) = F(I; s^2)$, \tilde{G} would operate through the product-market as well as the capital-market channel. Indeed, an earlier version of this article included such argument of F . It was dropped because, while probably accurate, it added more to the model's complexity than to its intuition.

opportunity perfectly hedges preexisting exposures, and overstates value if there are no preexisting exposures. Nevertheless, our contribution is to provide at least a basis for pricing new opportunities that alter the profile of preexisting asymmetric exposures.

Simultaneous Multiple New Opportunities

Thus far, in solving for the optimal quantity and hurdle rate for the j th new opportunity, we have taken as given all the other J quantities and hurdle rates. It is as though the insurer is considering a single new opportunity, having already fixed decisions on the $J - 1$ others. Clearly, the optimal thing to do is to solve the J first-order equations simultaneously. This is straightforward to do. To express the results, we use notation in which n^* is a $J \times 1$ vector of optimal amounts of new opportunities; Σ is the $J \times J$ covariance matrix of the shocks of the unhedgeable insurance components, $\varepsilon_{N,j}^I$; μ_N is the $J \times 1$ vector of total expected returns on the new opportunities; C_{NM} is the $J \times 1$ vector of covariances of the new opportunities with the market, $\text{cov}(\varepsilon_{N,j}, M)$; and C_{NP} is the $J \times 1$ vector of covariances of the new opportunities with the preexisting exposures, $\text{cov}(\varepsilon_{N,j}, \varepsilon_P)$. The vector of optimal investments in new opportunities is given by:

$$n^* = \left(\frac{1}{F + G} \right) \Sigma^{-1} (\mu_N - \gamma C_{NM}) - \left(1 + \frac{\tilde{G}}{F + G} \right) \Sigma^{-1} C_{NP}, \quad (13)$$

whereas the vector of incremental required returns for an incremental opportunity is:

$$\mu_N = \gamma C_{NM} + (F + G)(C_{NP} + \Sigma n) + \tilde{G} C_{NP}. \quad (14)$$

Both of these expressions are solutions to the J first-order conditions in Propositions 3 and 4. However, simultaneous solution makes the coordination problem among the J individual decisions clear. The decision to hold more or less the j th opportunity impacts the desirability of the other $J - 1$ opportunities. If the decisions must be made simultaneously, a central planner would need to implement Equation (13). The planner would need to collect all information on the J opportunities and then mandate the actions. Of course, in practice, such microcoordination is rarely even a goal, let alone an actual achievement.

One reason that organizations do not attempt such coordination is that the individual decisions are typically small relative to the firm-wide capital. In that case, coordination is less important. We can approximate this above by treating each of the opportunities as small. In that case, the hurdle rate in Proposition 4 can be applied independently to each opportunity. That would imply that the incremental hurdle rate for an infinitesimal amount of the J new opportunities is:²⁵

²⁵ Equation (16) has two distinct factors, rather than the three in Proposition 4. However, that turns out to be a special feature of our simplifying assumptions and will not obtain generally. If we were to divide additively the asymmetrically distributed disturbance, ε_P^I , into asymmetric and normally distributed components, then three factors would be required to express the hurdle rate.

$$\mu_N = \gamma C_{NM} + (F + G + \tilde{G})C_{NP}. \quad (15)$$

This expression further highlights the reason that we refer to the hurdle rates as “incremental.” Usually, hurdle rates are independent both of the amount being invested and of other investments. That is not the case here. The j th hurdle rate depends not only on how much investment takes place in the other $J - 1$ opportunities; it depends also on n_j itself. The hurdle rate in Proposition 4 reflects this directly through the second priced factor, $\text{cov}(w, \varepsilon_{N,j}^I) = \text{cov}(\varepsilon_P^I + \sum_i n_i \varepsilon_{N,i}^I, \varepsilon_{N,j}^I)$, which is a linear function of all J of the n 's and an increasing linear function of n_j .

There are also indirect effects of n_j on the hurdle rate. As n_j increases, the factor loadings, F , G , and \tilde{G} will generally change as well. Thus, for example, if the insurer takes a large, very risky position in one opportunity—even if this position is independently distributed of all others—this will generally increase all the factor loadings, and thereby make the firm less willing to take on other risks. This is what Froot and Stein (1998) call a “firm-wide” risk effect: the distribution of internal funds, w , and therefore the factor loadings, are affected by changes in n_j .

Because $\mu_{N,j}$ is dependent on n_j , the hurdle rate is applicable only to an incremental new opportunity, where the contemplated opportunity is small and where n_j is the amount already in place. That is what we mean by an “incremental hurdle rate for an incremental new opportunity.” For a discrete opportunity of size \bar{n}_j , one would need to integrate the expression in Proposition 4 over n_j to get the hurdle appropriate to a discrete-sized opportunity.

OPTIMAL CAPITAL STRUCTURE AT TIME 0

We can now step back to time 0 and solve for the optimal capital level, K . A simple tradeoff is at work: on the one hand, as noted above, higher K moves the components of the insurer's effective risk aversion, F , G , and \tilde{G} , toward zero. Propositions 3 and 4 show that, *ex ante*, the insurer can invest more aggressively in new opportunities that promise an above-market return at time 1, hedge less with products that appear very costly at time 1, and cut back on overpaying for and/or overhedging with products whose chief purpose is to improve distributional asymmetries. All of these benefits, of course, must be balanced against the higher deadweight costs, τK , that come with higher levels of capital.

To illustrate the first part of the tradeoff most transparently, suppose that the new opportunity in question is a small one and that the other $J - 1$ decisions are fixed. In this setting, a natural question to ask is how the insurer's hurdle rate—as given by Proposition 4—changes with K :

$$\frac{d\mu_{N,j}}{dK} = (F_K + G_K)\text{cov}(w, \varepsilon_{N,j}^I) + \tilde{G}_K \text{cov}(\varepsilon_P^I, \varepsilon_{N,j}^I). \quad (16)$$

The first two factor loadings, F_K and G_K , are unambiguously negative. If the covariance of the new opportunity with the preexisting portfolio is positive, then this factor will push the hurdle rate smoothly toward fair market value as the amount of capital K is increased. It is a bit more complex for \tilde{G}_K , which will be negative (positive) for negatively (positively) asymmetric risks. Generally, \tilde{G} will move toward zero from

whatever side of zero it is on. Thus, any difference between firms with positive and negative \tilde{G} 's will be mitigated by more capital.

As of time 0, the insurer's objective function is to pick K so as to maximize $V - K$, recognizing that $V = V(w(\mu_P(n, K), n(K), K))$. In words, K affects w directly through the amount of financial slack that will be available at time 2, as well as indirectly through its influence on the optimal choice of new opportunities, n , and through its influence on the product market expected opportunity set, μ_P . Fortunately, one can use the envelope theorem to show that the solution to this problem can be written simply as:

$$\tau = F \frac{\sigma^2}{K} + 1 - \left(\frac{1}{E[P_w]} \right). \quad (17)$$

Equation (17) has an intuitive interpretation. The insurer can hold another dollar of slack capital at time 0 and pay the costs, τ , of carrying that capital. By spending the additional deadweight cost, the firm earns two benefits. First, it benefits from the improved product-market terms of trade. This is summarized by $F \frac{\sigma^2}{K} > 0$ —the marginal impact on product market terms of trade of additional capital, holding risk constant. Second, it benefits from the reduction in the expected costs of external funds. This is summarized by $\frac{E[P_w]-1}{E[P_w]} > 0$, which is literally the percentage gain from substituting for an expensive dollar of external capital the (now cheap) future value of an additional dollar of internal capital. The optimal level of K equates marginal costs against marginal benefits.

In the limiting case where $\tau = 0$, there are no deadweight costs of holding capital, so the insurer holds an arbitrarily large amount. As K becomes large, the expected shadow value of external funds, $E[P_w]$, falls to one. This in turn implies that both G and F converge to zero as well. The insurer behaves in a classical manner, making pricing and allocation decisions according to a purely market-based model of risk and return. In contrast, as τ increases above 0, the insurer holds less capital, thereby raising its effective risk aversion, and amplifying the deviations from textbook capital budgeting principles.²⁶

Notice also that the existence of product-market imperfections encourages the insurer to hold more capital than would be the case if there were only capital-market imperfections. These imperfections make the insurer more risk averse, increasing its demands for hedging and capital, and reducing its appetite to bear risk through asset holdings or underwriting.

SOME IMPLICATIONS OF THE THREE-FACTOR MODEL

Adding the third factor—asymmetry—provides another channel through which cross-sectional variation in firms' appetite for risk management might be explained.

²⁶ While G and \tilde{G} do not appear explicitly in Equation (18), both the concavity of P and the degree of negative asymmetry of distribution of w have, all else equal, negative impacts on $E[P(w)]$.

Consider, for example, financial firms with strong reputations and solid access to capital markets. With a lot to lose, from a regulatory or legal scandal or a large loss, they have negatively asymmetric distributions. We would expect such firms to be conservative in a number of ways, including in their internal pricing. Such conservatism is predicted by the asymmetry factor.

Small "growth" firms, by contrast, behave in just the opposite way. They are most likely to have positively asymmetric distributions of their assets, which include growth options. These firms may appear more aggressive in their capital budgeting and internal pricing, with surprisingly low hurdle rates, especially given their less durable access to capital markets. More aggressive hurdles for firms with powerful growth options would be predicted by the third factor.

The third factor also has implication for the pricing of risk in excess-of-loss insurance and reinsurance contracts. These contracts commit the underwriter to paying for a slice of losses, those between predetermined upper and lower bounds. The further out are the upper and lower bounds, the more extreme is the interaction between the negative asymmetry and other portfolio risks. Thus, the model predicts that insurers and reinsurers, who warehouse catastrophic exposures, find it very expensive to hedge them, at least among one another. In fact, there is evidence that realized prices of reinsurance tend to become much more expensive relative to fair value as contract retentions (i.e., deductibles) increase.²⁷ Taken literally, this suggests that more negatively asymmetric exposures are most costly for writers to take on and most valuable for cedents to shed.

ASYMMETRICALLY DISTRIBUTED NEW OPPORTUNITIES

This last section modifies an approach, discussed in Froot and Stein (1998), for valuing new opportunities that are asymmetrically distributed. This approach gives some useful insight; however, it can be applied in the case of insurance exposures only in rather limited circumstances, which we clarify below. To fix ideas, imagine that a reinsurer has a new opportunity to write a very small amount of excess-of-loss cover for the catastrophe losses of an insurer's underwriting book. (The excess-of-loss contract is similar to a call spread using standard options.) Imagine also that the insurer has issued a capital market instrument with payoffs tied directly linearly to its time 2 catastrophe results.²⁸

Let us assume that the price of the underlying instrument, e , follows a simple geometric Brownian motion with drift θ and instantaneous standard deviation, v , between times 1 and 2:

$$de = \theta e + ve dz. \quad (18)$$

²⁷ See Froot (1995).

²⁸ Such an instrument is not really the same as what are usually called "cat bonds," in that it would not have a maximum or promised payment of interest and principal, but rather payments that are directly linked to the catastrophe component of the firm's underwriting. It is closer to a market traded quota-share contract than to a catastrophe bond.

From Proposition 4, the excess return that the insurer requires on the underlying position is²⁹

$$\mu_{N,e} = \gamma \text{cov} \left(M, \frac{de}{e} \right) + (F + G) \text{cov} \left(w, \frac{de}{e} \right) + \tilde{G} \text{cov} \left(\varepsilon_p^I, \frac{de}{e} \right). \quad (19)$$

By Ito's lemma, the instantaneous change in the value of the derivative, L , is given by:

$$dL = \left(\theta e L_e + L_t + \frac{v^2 e^2}{2} L_{ee} \right) dt + v e L_e dz. \quad (20)$$

That is, the derivative and the underlying have perfectly correlated instantaneously normal innovations, implying that $\text{cov}(dL, x) = \text{cov}(de L_e, x) = L_e \text{cov}(de, x)$. Applying Proposition 4 and using the linearity of the covariance operator then yields the reinsurer's required excess return on the derivative:

$$\begin{aligned} \mu_{N,L} &= \frac{e L_e}{L} \left(\gamma \text{cov} \left(M, \frac{de}{e} \right) + (F + G) \text{cov} \left(w, \frac{de}{e} \right) + \tilde{G} \text{cov} \left(\varepsilon_p^I, \frac{de}{e} \right) \right) \\ &= \frac{e L_e}{L} (\mu_{N,e}). \end{aligned} \quad (21)$$

The continuous-time formulation implies that instantaneous innovations in the price of the derivative, dL , are normal, even though over discrete intervals, the distribution of ΔL is not normal because the moments dL in Equation (20) are changing over time.

Equation (21) says simply that the reinsurer's required return on the derivative is equal to the return required on the underlying times the elasticity of the derivative price with respect to the underlying price, $\frac{e L_e}{L}$. In the case of standard call options, for example, this elasticity is always greater than one. Moreover, the elasticity is increasing in the strike price of the call; e.g., more out-of-the-money options have greater price elasticities with respect to the price of the underlying. The firm will require a higher return for a call option with a higher strike price. In the context of our example, this implies that the reinsurer will require a higher return on the excess-of-loss contract than on the underlying. In addition, this required return is more sensitive with respect to F , G , and $|\tilde{G}|$.

The advantages of this approach are several. First, it allows us to price the reinsurance contract or, for that matter, any derivative written on an underlying illiquid instrument. The continuous time formulation allows for us to write the new investment opportunity as continuously normal, and therefore apply our pricing formulas. Second, it makes it clear that the reinsurance contract will have payoffs that are more negatively asymmetric than those of the underlying. Loosely speaking, the greater is the negative asymmetry of the new opportunity, the greater is the reinsurer's required return.

²⁹ In this continuous-time setting, ε_p^I should now be interpreted as the instantaneous innovation in the rate of return on the insurance component of the preexisting portfolio.

Proper caveats include the assumption that the underlying cat instrument is continuously priced in the marketplace and is liquidly tradable. Also, as before, the reinsurance contract must be very small. The assumption that expected underlying returns are constant would fail if the position were large.

CONCLUSIONS

This article attempts to provide a detailed framework for the pricing and allocation of risk by insurers and reinsurers. We find that internal pricing for these firms differs from external pricing of risk in the capital markets because of imperfections. We build on previous work by Froot and Stein (1998), where capital-market imperfections are paramount, by adding imperfections that come from the product-market sensitivity of customers to risk, and by adding features that allow for the pricing of asymmetric risk distributions, a key risk feature facing insurers and reinsurers.

It is unclear whether the model results should be interpreted as positive or normative. Certainly, they are positive in the sense that insurers and reinsurers *are*, in practice, concerned with risk management and capital allocation. Managers and practitioners do not need to be told that the Modigliani–Miller irrelevance theorems fail, and that risk management can raise value. The results are also positive in the sense that they provide a possible explanation for why reinsurance prices are so high (relative to expected loss); why reinsurance prices for particularly large-scale catastrophic perils, like Florida wind, are among the highest; why prices rise and quantities of reinsurance supplied fall in the aftermath of large event losses; and why the highest reinsurance layers often appear the most expensively priced vis-à-vis expected loss.

However, there are clearly normative aspects to the theory as well. Practitioners may use cost of capital formulations that are derived in *ad hoc* ways and that are not necessarily rooted in corporate value maximization. The theory may, therefore, have prescriptive content. Yet the precise results are also sensitive to the nature of the capital and product market distortions. So the approach may point in the right direction even if the exact results vary based on the details of the distortions.

It would be desirable to extend the results to allow for asymmetries in the distribution of new opportunity outcomes. Unfortunately, we cannot solve explicitly for nonnormal new opportunity disturbances. However, we do not need restrictive assumptions about *any* distributions in order to reach Equation (8), prior to the application of the generalization of Stein's Lemma.³⁰ Thus, while we cannot solve explicitly for n_j or $\mu_{N,j}$ if the j th new opportunity is nonnormally distributed, the first-order condition in Equation (8) always applies.

REFERENCES

- Baker, M., R. Ruback, and J. Wurgler, 2006, Behavioral Corporate Finance: A Survey, in: Espen Eckbo, ed., *The Handbook of Corporate Finance: Empirical Corporate Finance* (New York: Elsevier/North Holland).
- Cagle, J., and S. Harrington, 1995, Endogenous Supply with Capacity Constraints and Endogenous Insolvency Risk, *Journal of Risk and Uncertainty*, 11: 219-232.

³⁰ Except that we do need to assume that second own-moment and cross-moments exist.

-
- Cummins, J. D., and P. M. Danzon, 1997, Price, Financial Quality, and Capital Flows in Insurance Markets, *Journal of Financial Intermediation*, 6: 3-38.
- Cummins, J. D., and R. D. Phillips, 2000, Applications of Financial Pricing Models in Property Liability Insurance, in: G. Dionnes, ed., *The Handbook of Insurance* (Boston: Kluwer Academic Publishers).
- Cummins, J. D., and D. W. Sommer, 1996, Capital and Risk in Property-Liability Insurance Markets, *Journal of Banking and Finance*, 20: 1069-1092.
- Doherty, N. A., and S. M. Tinic, 1982, A Note on Reinsurance Under Conditions of Capital Market Equilibrium, *Journal of Finance*, 36: 949-953.
- Epermanis, K., and S. E. Harrington, 2001, Market Discipline and Reaction to Rating Changes in U.S. Property-Liability Insurance Markets, Presented at NBER Insurance Project Conference, February.
- Froot, K. A., 1995, Incentive Problems in Financial Contracting: Impacts on Corporate Financing, Investment, and Risk Management Policies, in: K. Froot, D. Crane, S. Mason, A. Perold, R. Merton, Z. Bodie, E. Sirri, and P. Tufano, eds., *The Global Financial System: A Functional Perspective* (Boston: Harvard Business School Press), 225-261.
- Froot, K. A., 1999, The Limited Financing of Catastrophe Risk: An Overview, in: K. Froot, ed., *The Financing of Catastrophe Risk* (Chicago and London: University of Chicago Press), 1-22.
- Froot, K. A., 2001, The Market for Catastrophe Risk: A Clinical Examination, *Journal of Financial Economics*, 60: 529-571.
- Froot, K. A. and P. G. J. O'Connell, 1997, On the Pricing of Intermediated Risks: Theory and Application to Catastrophe Reinsurance, Harvard Business School Working Paper No. 98-024, forthcoming in the special issue on insurance of the *Journal of Banking and Finance*, 2007.
- Froot, K. A., D. S. Scharfstein, and J. C. Stein, 1993, Risk Management: Coordinating Corporate Investment and Financing Policies, *Journal of Finance*, 48: 1629-1658.
- Froot, K. A., and J. C. Stein, 1998, Risk Management, Capital Budgeting and Capital Structure Policy for Financial Institutions: An Integrated Approach, *Journal of Financial Economics*, 47: 55-82.
- Gale, D., and M. Hellwig, 1985, Incentive-Compatible Debt Contracts I: The One-Period Problem, *Review of Economic Studies*, 52: 647-664.
- Garven, J. R., and J. Lamm-Tennant, 2003, The Demand For Reinsurance: Theory and Empirical Tests, *Insurance and Risk Management*, 7(3): 217-238.
- Grace, M. F., R. W. Klein, and P. R. Kleindorfer, 2001, The Demand for Homeowners Insurance with Bundled Catastrophe Coverage, Presented at NBER Insurance Project Conference, February.
- Grace, M. F., R. W. Klein, Paul R. Kleindorfer, and M. R. Murray, 2003, *Catastrophe Insurance: Consumer Demand, Markets, and Regulation* (Boston: Kluwer Academic Publishers).
- Greenwald, B. C., A. Levinson, and J. E. Stiglitz, 1991, Capital Market Imperfections and Regional Economic Development, in: Giovannini, ed., *Finance and Development: Lessons and Experience* (Washington DC: CEPR), 65-93.

- Gron, A., 1994a, Capacity Constraints and Cycles in Property-Casualty Insurance Markets, *Rand Journal of Economics*, 25: 110-127.
- Gron, A., 1994b, Evidence of Capacity Constraints in Insurance Markets, *Journal of Law and Economics*, 25: 110-127.
- Hall, B. 1999, Property and Casualty Solvency Funds as a Tax and Social Insurance System, in: K. Froot, ed., *The Financing of Catastrophe Risk* (Chicago and London: University of Chicago Press).
- Hoerger, T. J., F. A. Sloan, and M. Hassan, 1990, Loss Volatility, Bankruptcy, and the Demand for Reinsurance, *Journal of Risk and Uncertainty*, 3: 221-245.
- Jaffe, D., and T. Russell, 1997, Catastrophe Insurance, Capital Markets, and Uninsurable Risks, *Journal of Risk and Insurance*, 64: 205-230.
- Kahneman, D., and A. Tversky, 1979, Prospect Theory: An Analysis of Decision Under Risk, *Econometrica*, 47(2): 263-291.
- Kaplan, S., and L. Zingales, 2000, Investment-Cash Flow Sensitivities Are Not Valid Measures of Financing Constraints, *Quarterly Journal of Economics*, 115(2): 707-712.
- Kozik, T. J., and A. M. Larson, 2001, *The n-Moment Insurance CAPM*, *Proceedings of the Casualty Actuarial Society*, 88: 168-169.
- Kraus, A., and R. Litzenberger, 1976, Skewness Preference and the Valuation of Risky Assets, *Journal of Finance*, 31: 1085-1100.
- Malmendier and Tate, 2005, CEO Overconfidence and Corporate Investment, *Journal of Finance*, 60: 2661-2700.
- Merton, R. C., 1993, Operation and Regulation in Financial Intermediation: A Functional Perspective, in: P. Englund, ed., *Operation and Regulation of Financial Markets* (Stockholm: The Economic Council), 17-68.
- Merton, R. C., 1995a, A Model of Contract Guarantees for Credit-Sensitive Opaque Financial Intermediaries, Harvard Business School Working Paper.
- Merton, R. C., 1995b, A Functional Perspective on Financial Intermediation, *Financial Management*, 24: 23-41.
- Merton, R. C., and A. F. Perold, 1993, The Theory of Risk Capital in Financial Firms, *Journal of Applied Corporate Finance*, 5: 16-32.
- Myers, S. C., and N. Majluf, 1984, Corporate Financing and Investment Decisions When Firms Have Information that Investors Do Not Have, *Journal of Financial Economics*, 13: 187-221.
- Myers, S. C., and J. A. Read, 2001, Capital Allocation for Insurance Companies, *Journal of Risk and Insurance*, 68: 545-580.
- Phillips, R. D., J. D. Cummins, and F. Allen, 1998, Financial Pricing of Insurance in Multiple Line Insurance Companies, *Journal of Risk and Insurance*, 65: 597-636.
- Sommer, D. W., 1996, The Impact of Firm Risk on Property-Liability Insurance Prices, *Journal of Risk and Insurance*, 63(September): 501-514.
- Stein, C., 1981, Estimation of the Mean of a Multivariate Normal Distribution, *The Annals of Statistics*, 9: 1135-1151.
- Stein, C., 1996, Rational Capital Budgeting in an Irrational World, *Journal of Business*, 69: 429-455.

- Taylor, G., 1994, An Equilibrium Model of Insurance Pricing and Capitalization, *Journal of Risk and Insurance*, 62: 409-446.
- Townsend, R. M., 1979, Optimal Contracts and Competitive Markets and Costly State Verification, *Journal of Economic Theory*, 21: 265-293.
- Venter, G. G., and J. A. Major, 2002, *Allocating Capital by Risk Measures: A Systematic Survey*, Guy Carpenter.
- Wakker, P. P., R. H. Thaler, and A. Tversky, 1997, Probabilistic Insurance, *Journal of Risk and Uncertainty*, 15: 7-28.
- Winter, R. A., 1994, The Dynamics of Competitive Insurance Markets, *Journal of Financial Intermediation*, 3: 379-415.
- Zanjani, G., 2002, Pricing and Capital Allocation in Catastrophe Insurance, *Journal of Financial Economics*, 65(2): 283-305.