Myers, Stewart C; Read, James A, Jr Journal of Risk and Insurance; Dec 2001; 68, 4; ProQuest Central pg. 545

©The Journal of Risk and Insurance, 2001, Vol. 68, No. 4, 545-580

CAPITAL ALLOCATION FOR INSURANCE COMPANIES

Stewart C. Myers James A. Read, Jr.

ABSTRACT

This study shows how option pricing methods can be used to allocate required capital (surplus) across lines of insurance. The capital allocations depend on the uncertainty about each line's losses and also on correlations with other lines' losses and with asset returns. The allocations depend on the marginal contribution of each line to default value--that is, to the present value of the insurance company's option to default. The authors show that marginal default values add up to the total default value for the company, so that the capital allocations are unique and not arbitrary. They therefore disagree with prior literature arguing that capital should not be allocated to lines of business or should be allocated uniformly. The study presents several examples based on standard option pricing methods. However, the "adding up" result justifying unique capital allocations holds for any joint probability distribution of losses and asset returns. The study concludes with implications for insurance pricing and regulation.

INTRODUCTION AND SUMMARY

This paper clarifies how option pricing methods can be used to determine how much capital an insurance company should carry and how that capital requirement should be allocated across the company's lines of insurance business. It is the convention in the insurance business to refer to capital as "surplus."

Surplus is important because more surplus means more collateral for outstanding policies. Surplus is costly for at least two reasons. First, agency and information costs may attach to risk capital in any financial intermediary. These costs are described by Merton and Perold (1993). Second, the U.S. tax system subjects investment income to double taxation, first at the corporate level and again when it is realized by the corporation's shareholders. When an insurance company raises additional surplus

Stewart C. Myers is with the MIT Sloan School of Management and The Brattle Group, Cambridge, Mass. James A. Read, Jr., is with The Brattle Group. The authors thank the Automobile Insurers' Bureau of Massachusetts for research support and Richard Derrig, Robert Butsic, Robert Merton, and participants in seminars at The Wharton School, NBER, and MIT for many helpful comments. This paper's analysis and conclusions are the authors' responsibility and may not represent the views of the Bureau or its members.

545

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission w

and invests it in securities to collateralize policies, the company's equity investors are essentially holding these securities in a taxable mutual fund. The shareholders' investment is subject to a layer of taxes not encountered in direct investment in securities or in ordinary mutual funds. To survive, the insurance company has to recover these tax costs.

Because surplus is costly, competitive premiums-and fair regulated premiumsdepend on total surplus requirements and on their allocation to lines of insurance.¹ The line-by-line allocations are important. If surplus allocations are wrong, cost allocations will be wrong, too. In a competitive setting, allocation errors may lead firms to write unprofitable business and lose profitable business to competitors. In a regulated setting, some lines will cross-subsidize others, and companies will be tempted to "push" lines with too-low surplus allocations.

The use of option pricing concepts to analyze default risk and surplus requirements is well established. Early articles include Doherty and Garven (1986), Cummins (1988), and Derrig (1989).² These analyses are closely related to models of default risk, deposit insurance, and capital requirements in banking, a literature that started with Merton (1977).

But the application of these concepts is complicated. For example, most propertycasualty companies write several lines of insurance. Therefore, a company's total surplus requirement must depend on the composition of its business, and managers and regulators must consider how to allocate total surplus line by line. This allocation is necessary because lines must be priced individually. But the proper allocation is not obvious. There are gains from diversification, because each line coinsures the others. Each line's policyholders are paid as long as the firm as a whole is solvent.

There is a view, apparently widespread, that capital (surplus) cannot, or should not, be allocated to lines of insurance. For example, Phillips, Cummins, and Allen (1998, $p. 599$) conclude that:

Our model implies that it is not appropriate to allocate capital by line; rather, the price of insurance by line is determined by the overall risk of the firm and the line-specific liability growth rates. Thus, prices are predicted to vary across firms depending upon firm default risk, but prices for different lines of business within a given firm are not expected to vary after controlling for liability growth rates by line.

The authors of this article agree that any reduction in competitive prices (premiums) due to default risk will be the same for all lines for a given firm, other things equal. But the marginal contribution to default risk does vary across the lines. The authors show that these marginal contributions "add up" and argue that surplus should be allocated at the margin based on these contributions. Phillips', Cummins', and Allen's model does not contradict the authors' results, because their setup included no tax or other costs of holding surplus; there was no need to allocate surplus. The authors of this article comment further on their model below.

In principle, surplus could be allocated policy by policy, but the authors do not go to this level of disaggregation.

² See also Butsic (1994), Cox and Hogan (1995), Cummins and Sommer (1996), Garven (1992), and Shimko (1992).

Merton and Perold (1993, p. 30) stress the dangers of capital allocations for financial corporations generally:

For a given configuration, the risk capital of a multi-business firm is less than the aggregate risk capital of the businesses on a stand-alone basis. Full allocation of risk-capital across the individual businesses of the firm therefore is generally not feasible. Attempts at such a full allocation can significantly distort the true profitability of individual businesses.

It is true that the *reduction* in risk and required surplus due to less-than-perfect correlation between lines of business cannot be uniquely allocated back to the lines. However, if a company operates in two or more lines of business, then the lines' marginal surplus allocations are unique and do "add up."

Example

The distinction between stand-alone and marginal surplus requirements is critical to understanding this study. Table 1 provides an illustration. It assumes three lines of insurance, each generating future losses with present value of $$100 [PV_l(s) = $100]$. Panel A of the table assumes these lines are combined in one company with assets of \$450 and a surplus of \$150. The problem is to allocate this surplus across the three lines.

Because future losses are uncertain, there is a chance that total losses will exceed the future value of the company's assets. In this case, the company defaults. The authors measure the risk of default by the *default value*, defined as the premium the company would have to pay in a competitive market for a policy guaranteeing payment of its losses. In this case, the default value is \$0.93 or 0.31 percent of the current present value of losses. (The authors explain later how this and the other entries in Table 1 are calculated. The necessary formulas are given in Appendix 2.)

The middle column of Panel A shows the marginal surplus requirement for each line--that is, the marginal change in required total surplus in response to a marginal increase in PV(losses) for the line. This derivative is calculated with default value held constant as a percentage of the total of the lines' PV(losses). For example, an increase in Line 1's PV(losses) from 100 to 101 increases required surplus from 150 to 150.38. The right column of Panel A shows line-by-line surplus allocations based on these marginal requirements. Notice that the allocations add up exactly to the total surplus of \$150. The authors will show that this adding-up property is general and that the resulting allocations are correct from an economic point of view.

These allocations do depend, however, on the insurance company's composition of business-that is, on the relative sizes of the lines' losses and the losses' risk characteristics. Panel B of Table 1 shows the surplus requirements if each line is organized as a stand-alone company. The total surplus required for the three lines increases to \$171 because of loss of diversification. As Merton and Perold note, the reduction in surplus attributable to diversification *cannot* be allocated to the individual lines.³

There is also no sense or point for the diversified firm in Panel A to use the stand-alone surplus requirements in Panel B. But stand-alone requirements are sometimes used in practice in value-at-risk (VAR) systems in banking. See Saita (1999) for a survey of capital allocation in a VAR setting.

Examples of Surplus Allocations

Panel A shows marginal surplus requirements for three lines of insurance. Surplus allocations based on these marginal requirements add up to the total surplus carried by the firm. Panel B shows the stand-alone surplus requirements for each line. Panel C shows the total surplus required by each line, given the other two lines. In all cases default value is held constant at 0.31 percent of PV(losses).

Panel A: Marginal Surplus Requirements for Three Lines of Insurance

Panel B: Stand-Alone Surplus Requirements for Each Line

Panel C: Total Surplus Required by Each Line, Given the Other Two Lines

The allocations shown in Panel A do not apply to inframarginal changes. Panel C shows how surplus requirements change when Line 1, 2, or 3 is eliminated. For example, exit from Line 1 reduces total required surplus by \$35, from \$150 to \$115. In other words, a company that starts with Lines 2 and 3, and then adds Line 1, would have to contribute \$35 of additional surplus. The equivalent requirements for Lines 2 and 3 are \$46 and \$58, respectively, as shown in the right column of Panel C. These

requirements do not add up to the \$150 of surplus carried by the three-line firm in Panel A, and any allocation based on these requirements would be at best arbitrary and probably perverse.

Merton and Perold's statement that "Full allocation of risk capital is ... generally not feasible" is correct for inframarginal changes; for example, decisions to exit a line of business or to acquire a large block of new business. In fact, Merton and Perold explicitly focus on such decisions. But there is no reason for a firm to forswear allocation of its capital across its existing lines of business. This study shows how that allocation should be accomplished. We will also show that introduction of a new line of insurance will not necessarily change surplus allocations for existing lines to any material extent.

Summary of Results

The function of surplus is to reduce the risk of default to an acceptably low level. Option pricing theory provides a way to measure the cost of default insurance, namely by the value of a default put, which is the present value of possible future losses to policy owners. This "default value" depends on the level of surplus relative to the present value of future losses; on the degree of uncertainty about each line's losses; and also on the correlations between line-by-line losses and the returns on the insurance company's assets.

Each line of insurance contributes to the company's default value. The authors' key theoretical result is that the lines' marginal contributions to default value add up. A value-weighted sum of marginal default values exactly equals the default value for the company. This result holds for any joint probability distribution of line-by-line losses and asset returns.

Each line's marginal default value depends on its marginal surplus allocation. For example, if expansion of Line 1 is supported by a large increase in surplus, then the expansion will add relatively little—or perhaps even subtract from—default value. If each line is supported by the same amount of surplus, then marginal default values will differ. But if we fix each line's marginal default value, then the marginal surplus requirements are fixed also. The resulting marginal surplus requirements also "add up" to the overall surplus held by the firm.

If the firm defaults on one line of policies, it defaults on all. As Phillips, Cummins, and Allen (1998) noted, default risk has the same proportional effect on each line's competitive premium. Therefore, each line's marginal surplus requirement should be set so that all lines' marginal contributions to default value are the same. This generates the premiums that would be observed in a competitive market when holding surplus is costly, and it prevents cross-subsidization of some lines of insurance by others.

The next section of the study reviews the option pricing framework for evaluating surplus requirements and default risk. The authors then present a formal model, derive formulas for marginal default values and surplus allocations, and work through numerical examples showing how surplus can be allocated line by line. The final sections of the study consider implications for pricing, regulation, and the optimal line-by-line composition of insurance business.

with permission of the copyright owner. Further reproduction prohibited without permission_W

The formulas and examples that follow assume for simplicity that the joint probability distribution of total losses and asset returns is either lognormal or normal. This does not limit or qualify the authors' theoretical results, because marginal default values and surplus requirements add up regardless of the form of the joint probability distribution of losses and asset returns. But practical application of the authors' results will probably require further research to identify the actual joint distributions of losses and returns, and to develop numerical methods for calculating marginal default values and surplus requirements.

THE OPTION PRICING FRAMEWORK

We start with a simple balance sheet for an insurance company.

For the moment, "losses" refers to future losses on all outstanding policies in all lines of business. Losses are net of loss expenses. For simplicity, the authors do not break out the present value of tax liabilities.

This balance sheet ignores the possibility of default. Losses are valued assuming claims will always be paid. Surplus is defined as the difference between the market value of assets and the default-free present value of losses. More surplus means more assets and greater assurance that the losses will actually be paid. But there is always some chance that losses and expenses will exceed the future value of assets. If this happens, the company defaults.

What happens in default? Several possibilities exist. In unregulated, wholesale insurance markets, the policyholders are left holding the bag; they can recover no more than the insurer's assets. This possibility reduces the present value of an insurance policy when issued and the premium a policyholder is willing to pay for it. The reduction in premium should equal the value of a put option written on the assets, where the put's exercise price is the amount of losses payable to the policyholder absent default. The competitive premium⁴ is:

Premium = PV(losses, assuming no default) + PV(surplus costs) - PV(default option),

where PV(surplus costs) captures the tax and other costs of surplus in an insurance company.

Plain vanilla options have fixed exercise prices. Here the exercise price is uncertain because future losses are not known when a policy is issued. The put is therefore an exchange option—the right to exchange the company's assets for its obligations to policyholders. The value of this exchange option is the *default value*:

Default value = PV (default option) = PV (option to exchange assets for liabilities).

In unregulated insurance markets, the default value is implicitly deducted from the premium paid. In regulated retail insurance markets, the risk of default is not absorbed by individual policyholders but by the industry as a whole (and ultimately by

In this study, "premiums" refers to net premiums—premiums less marketing expenses and commissions.

policyholders in the aggregate). For example, a state pool or fund typically guarantees payment of losses on policies of insolvent companies. Solvent companies contribute to the fund as required, usually after an insolvency.

In practice, the distinction between unregulated "wholesale" and regulated "retail" insurance markets is not as sharp as this suggests. Retail policy owners are not completely protected. The amount payable to any one policy owner is limited, and owners of policies issued by companies in default may face extra costs and delays in securing payment of claims. But the results of this study do not require a detailed analysis of pooling arrangements or guarantees, so the authors will simplify by treating retail policy owners as completely protected.

Industry pooling arrangements are tricky to model. If customers do not have to worry about default, premiums are set as if defaults never occur. The companies' premiums are not reduced by PV(default option). Thus, each company receives the default value but assumes a contingent liability for defaults by other companies. It's as if each company gets default insurance guaranteeing full payment of losses regardless of the future size of the company's losses or the future value of its assets.

The cost of the default insurance is the present value of the company's possible future payments to the industry pool. This cost cannot be calculated without modeling the interactions of many firms. But that cost is, for practical purposes, fixed from the point of view of any single firm. The key issue here is the default value—that is, the value of the default insurance the company obtains from the pool. This default value is the cost imposed on other companies in the industry. Sensible regulation will attempt to equalize default values (per dollar of liabilities) across companies and lines of insurance. Otherwise, policies will be mispriced and companies will be tempted to add risk and take advantage of other companies in the pool.

If payment of losses is guaranteed by an industry pool or fund, the initial (date $t = 0$)⁵ market value balance sheet, including default value, is:

In effect, the insurance company "buys" the default insurance, worth D , by accepting a liability to the guarantee fund.⁶

The authors have now distinguished equity (E) from surplus (S) . Surplus is defined as assets (V) less the present value of losses assuming no default (L) . (Think of L as the discounted face value of policyholders' claims.⁷) Equity is the market value of the residual claim.

⁵ The authors omit the time subscript for present values at date $t = 0$. For example, initial asset value is $V = V_0$. The uncertain future asset value will be written as V_1 .

⁶ If the firm defaults on payments to policyholders, then it also defaults on any obligation to the guarantee fund. Thus we could include G in L —that is, treat the guarantee fund as another line of business.

The discount rate would reflect the systematic risk of the losses but not the credit risk of the policy writer. See Myers and Cohn (1987).

REFERENCES

- Brealey, R. A., and S. C. Myers, 2000, Principles of Corporate Finance, 6th ed. (Homewood, Ill.: Irwin McGraw-Hill).
- Butsic, R. P., 1994, Solvency Measurement for Property-Liability Risk-Based Capital Applications, Journal of Risk and Insurance, 61: 656-690.
- Cox, S. H., and A. M. B. Hogan, 1995, Life Insurer Risk-Based Capital: An Option-Pricing Approach, Journal of Actuarial Practice, 3: 5-22.
- Cummins, J. D., 1988, Risk-Based Premiums for Insurance Guaranty Funds, Journal of Finance, 43: 823-839.
- Cummins, J. D., S. E. Harrington, and R. Klein, 1995, Insolvency Experience, Risk-Based Capital and Prompt Corrective Action in Property-Liability Insurance, Journal of Banking and Finance, 19: 511-527.
- Cummins, J. D., S. E. Harrington, and G. Niehaus, 1994, Risk-based capital requirements in property-liability insurance: A financial analysis of conceptual and measurement issues, in: The Future of the Insurance Industry, E. Altman and I. Vanderhoof, eds. (Homewood, Ill.: Irwin Professional Publishers).
- Cummins, J. D., and D. W. Sommer, 1996, Capital and Risk in the Property-Liability Insurance Markets, Journal of Banking and Finance, 20: 1069-1092.
- Derrig, R. A., 1989, Solvency Levels and Risk Loadings Appropriate for Fully Guaranteed Property-Liability Insurance Contracts: A Financial View, in: Financial Models of Insurance Solvency, J.D. Cummins and R. A. Derrig, eds. (Boston: Kluwer Academic Publishers), 303-354.
- Doherty, N. A., and J. R. Garven, 1986, Price Regulation in Property-Liability Insurance: A Contingent-Claims Approach, Journal of Finance, 41: 1031-1050.
- Garven, J. R., 1992, An Exposition of the Implications of Limited Liability and Asymmetric Taxes for Property-Liability Insurance, Journal of Risk and Insurance, 59:34-56.
- Ho, T. S. Y., 1999, Allocate Capital and Measure Performances in a Financial Institution, Financial Markets, Institutions and Instruments, 8: 1-23.
- Lamont, O., 1997, Cash Flow and Investment: Evidence from Internal Capital Markets, Journal of Finance, 52: 83-109.
- Margrabe, W., 1978, The Value of An Option to Exchange and One Asset for Another, Journal of Finance, 33: 177-186.
- Merton, R. C., 1998, Applications of Option-Pricing Theory: Twenty-Five Years Later, American Economic Review, 88: 323-349.
- Merton, R. C., 1993, Operation and Regulation and Financial Intermediation, in: Operation and Regulation of Financial Markets, P. Englund, ed. (Stockholm: Economic Council).
- Merton, R. C., 1977, An Analytic Derivation of the Cost of Deposit Insurance and Loan Guarantees, Journal of Banking and Finance, 1: 3-11.
- Merton, R. C., and A. F. Perold, 1993, Theory of Risk Capital in Financial Firms, Journal of Applied Corporate Finance, 6: 16-32.

Myers, S. C., 1999, Financial Architecture, European Financial Management, 5: 133-142.

- Myers, S. C., and R. Cohn, 1987, Insurance Rate of Return Regulation and the Capital Asset Pricing Model in: Fair Rate of Return in Property Liability Insurance, J. D. Cummins and S. Harrington, eds. (Norwell, Mass.: Kluwer-Nijhoff Publishing Company).
- Myers, S. C., and R. Rajan, 1998, The Paradox of Liquidity, Quarterly Journal of Economics, 3: 733-771.
- Phillips, R. D., J. D. Cummins, and F. Allen, 1998, Financial Pricing of Insurance in a Multiple-Line Insurance Company, Journal of Risk and Insurance, 65: 597-636.
- Saita, F., 1999, Allocation of Risk Capital in Financial Institutions, Financial Management, 28: 95-111.
- Shimko, D. C., 1992, The Valuation of Multiple Claim Insurance Contracts, Journal of Financial and Quantitative Analysis, 27: 229-246.

Surplus is an input, and equity is an output. The underlying assets (state variables) for the default option are assets (V) and losses (L) .⁸ Surplus is equal to the difference $(S = V - L)$, so allocating surplus amounts to allocating the excess of the insurance company's assets over its liability to policyholders. The amount of surplus therefore affects the market value of equity.

We now proceed to the option formulas. For simplicity, the authors will stick to a oneperiod (two-date) model. Thus they do not deal explicitly with risk created by lines with long tails of loss payments. However, the analysis is easily generalized. The definition of PV(losses) can be interpreted as including the present value of remaining losses from policies issued in previous periods. An analysis of default value and risk has to recognize the amount and uncertainty of these remaining obligations as if they were created by newly issued policies.

At the end of the period (date $t = 1$), the insurance company shareholders receive $V_1 - L_1$ if V_1 , the asset value at date 1, is sufficient to cover actual losses L_1 . (Note that the time subscript $t = 1$ denotes the uncertain future loss amount or asset value.) If assets are less than losses, the firm defaults and shareholders receive nothing. In other words, shareholders hold the option to pay off the insurance policies and thereby realize the residual value, if any, of the assets:

$$
E_1 = \max\{0, (V_1 - L_1)\}
$$

The end-of-period payoff to equity, E_1 , can also be expressed as the end-of-period value of the assets minus the policy losses plus the payoff to a default option:

$$
E_1 = V_1 - L_1 + D_1
$$

$$
D_1 = \text{Max} \{ 0, (L_1 - V_1) \}.
$$

The present value at date $t = 0$ of the default option—the default value—is:

$$
D = PV\left(\max\left\{0, \left(L_1 - V_1\right)\right\}\right)
$$

If a claim to the policy losses were actively traded, the present value of losses could simply be observed. It would also be possible to estimate the probability distribution for the present value of losses and compute the present value of the default option. Specifically, because the payoffs to the option depend solely on two underlying assets—a claim to the policy losses and a portfolio of investments—the option could be replicated dynamically by a combination of positions in those assets:

$$
D = \left(\frac{\partial D}{\partial L}\right) L + \left(\frac{\partial D}{\partial V}\right) V.
$$
\n(1)

Note that the option pricing methods do not value V and L . Asset value V is observable if assets are traded securities. The present value of losses usually has to be calculated by discounting. Contrary to Phillips, Cummins, and Allen (1998, p. 598 at n. 1), discount rates for losses must be estimated before options can be valued.

The quantities of the two assets in the replicating portfolio $\left(\frac{\partial D}{\partial L}\right)$ and $\left(\frac{\partial D}{\partial V}\right)$ depend on the present value of policy losses (L) , the market value of assets (V) , and their joint probability distribution. For example, if future losses and asset returns are jointly lognormal, we can invoke Margrabe's (1978) solution for the value of an exchange option. In that case, the weights in the replicating portfolio [Equation (1)] are $\partial D/\partial L = N\{z\}$ and $\partial D/\partial V = -N\{z - \sigma\}$ where $N\{\bullet\}$ is the cumulative probability function for the standard normal variable.

$$
z = \ln\left(\frac{L}{V}\right) / \sigma + \frac{1}{2}\sigma,
$$

and σ is the volatility of the asset-to-liability ratio.⁹ The volatility of the asset-toliability ratio in turn depends on the volatility of losses σ_L , the volatility of assets σ_v , and the covariance of the natural logarithms of losses and assets σ_{uv} :¹⁰

$$
\sigma = \sqrt{\sigma_L^2 + \sigma_V^2 - 2\sigma_{LV}}
$$

Thus, the value of the default option depends on only three variables—the present value of liabilities, the market value of assets, and the volatility of the asset-to-liability ratio:

$$
D=f(L,V,\sigma).
$$

In reality, claims on losses will not be explicitly traded, but standard option pricing methods can still be used, providing that financial markets are complete. "Complete" means that the menu of traded securities is sufficiently diverse that investors could match the option with positions in traded securities. This is a standard assumption in applied corporate finance. For example, consider a corporation seeking to maximize market value and evaluating a proposed capital investment project by discounted cash flow. The objective will be endorsed by all shareholders if investing in the project does not change the scope of risk characteristics attainable by investors. Strictly speaking, that in turn requires that investors could match the risk characteristics of the project's future cash flows with positions in traded securities. If this requirement is met, then the project has a well-defined market value even though it is not explicitly traded.

The same requirement applies in the insurance setting when PV(losses) is calculated by discounting expected future losses. The present (market) value is well-defined if the risk characteristics of the losses could be matched by positions in traded securities.¹¹ That seems a reasonable assumption, which is (implicitly) widely accepted. That assumption justifies both the calculation of PV(losses) and the use of option pricing methods to value the default option.¹²

The volatility of the asset-to-liability ratio is the same as the volatility of the liability-toasset ratio. Note that the time to expiration has been suppressed because it is fixed at one period.

 10 Because the time to expiration is fixed at one period, the volatility parameter is equal to the standard deviation of the natural logarithm of the end-of-period asset-to-liability ratio.

¹¹ The match does not have to be perfect, providing that the "tracking error" is diversifiable noise.

¹² For more on the application of option pricing methods to nontraded assets, see Merton (1998) and Brealey and Myers (2000), Ch. 21, especially pp. 636-637.

DEFAULT VALUES FOR INSURANCE PORTFOLIOS

This study is concerned with the allocation of default values and surplus requirements to lines of business in multiple-line insurance companies. The authors' goal, in other words, is to allocate the default values of *portfolios* of insurance business line by line. Because the default value depends on the present value of losses, the value of assets, and the associated risks, we need to be able to allocate these variables to lines of business.

It is straightforward to allocate liabilities and assets to lines of insurance. Aggregate liabilities equal the sum of the present value of losses on each line. If an insurance company is engaged in *M* lines of business, $L = \sum_{i=1}^{M} L_i$ where $L_i = PV(L_i)$. Aggregate surplus equals the sum of line-by-line surplus contributions, which are proportional to liabilities, so $S = \sum_{i=1}^{M} L_i s_i = Ls$ where $s_i \equiv \partial S / \partial L_i$ is the surplus required per dollar of liabilities in Line *i* and $s \equiv S/L$ is the aggregate surplus-to-liability ratio. Assets equal the sum of liabilities and surplus:

$$
V = \sum_{i=1}^{M} L_i (1 + s_i) = L(1 + s).
$$

Note that the aggregate surplus ratio is a weighted average of the line-by-line surplus requirements, $s = \sum_{i=1}^{M} x_i s_i$, where the weights are fractions of total liabilities, $x_i \equiv L_i/L$.

The authors define the line-of-business default allocations d_i as the *marginal* contributions to the default value of the company: $d_i \equiv \partial D / \partial L_i$. Appendix 1 shows that the default value for a multi-line insurance company can be allocated uniquely to lines of business. Specifically, the sum of the products of line-by-line liabilities and marginal default values is equal to the default value for the company:

$$
\sum_{i=1}^{M} L_i d_i = D \tag{2}
$$

The authors' conclusion that line-by-line marginal default values "add up" does not depend on restrictive assumptions about the probability distribution of losses or asset returns. The only requirement is that assets and line-by-line losses have welldefined present values. This does not require that claims to the losses be explicitly traded, only that financial markets are sufficiently complete that the losses would trade at well-defined market values.

To derive a formula for marginal default values, we need to specify the form of the probability distribution of future losses and asset values. The standard assumption

in the option pricing literature is the lognormal distribution. The lognormal is a natural choice because it implies that asset values are bounded from below at zero (negative values are ruled out) and are unbounded from above-that is, the probability distribution of future values is positively skewed. However, it presents a technical problem in a portfolio context, because a sum of lognormal variables is not itself lognormal. Specifically, if each line's future loss is lognormal, then the overall loss cannot be lognormal. Conversely, if the overall loss is lognormal, then the line-byline losses cannot be lognormal.¹³ As a result, an exact closed-form solution cannot be obtained if we assume that the line-by-line loss distributions are lognormal.

It is possible to derive an alternative closed-form solution for marginal default values if it is assumed instead that the distribution of future losses and asset values is joint normal. The probability distribution of a sum of normal random variables is normal, so if distributions of losses by line are normal, the aggregate loss is normal, too. If, in addition, the distribution of asset values is normal, then the distribution of surplus—the difference between future asset values and losses—is also normal. The cost of obtaining closed-form results is the assumption of symmetry in the distributions of losses, asset values, and surplus.

The following derivations of default values and surplus allocations assume that total losses (the sum of all lines' losses) and asset values are joint-lognormal. The authors believe that this is a reasonable approximation even when individual lines' losses are also lognormal. However, the authors' numerical illustrations include results for both the joint-normal and joint-lognormal cases.

Ultimately, the appropriate form of probability distribution is an empirical matter. It is an important matter because default is a deep out-of-the-money option, which implies that the "tails" of the distributions are influential. If the choice of distribution does not admit a closed-form solution for default values, analysts may have to resort to numerical methods, such as Monte Carlo simulation, to compute marginal default values and surplus requirements. But our basic result, that default values and surplus allocations are unique and "add up," holds for any joint probability distribution (see Appendix 1).

MARGINAL DEFAULT VALUES-LOGNORMAL CASE

Recall that if the distribution of aggregate losses and asset values is joint lognormal, the relevant measure of portfolio risk is the volatility of the asset-to-liability ratio:

$$
\sigma = \sqrt{\sigma_{\rm L}^2 + \sigma_{\rm V}^2 - 2\sigma_{\rm LV}}\,\,\cdot
$$

If the line-by-line loss volatilities are not large, the following expressions provide close approximations to the volatility of total losses and the covariance of the natural logarithms of losses and asset values ("log" losses and "log" assets):

¹³ The only case in which both line-by-line and total losses are lognormal occurs when the scale of each line is continuously rebalanced to keep its PV(losses) a constant proportion of the total. This could be an alternative justification for our formulas and examples based on lognormal losses.

$$
\sigma_L^2 = \sum_{i=1}^M \sum_{j=1}^M x_i x_j \rho_{ij} \sigma_i \sigma_j
$$

$$
\sigma_{LV} = \sum_{i=1}^M x_i \rho_{iv} \sigma_i \sigma_v.
$$

Correlations between log losses in two lines of insurance are denoted by ρ_{ii} . Correlations between log asset values and log losses in a single line are denoted by ρ_{iv} .

The next step is to calculate how the default value for a multiple-line insurance company is affected by a marginal change in PV(losses) for a line of business. A change in the amount of business written in Line i could affect all three of the option pricing parameters—the present value of liabilities, the market value of assets, and the volatility of the asset-to-liability ratio:

$$
\frac{\partial D}{\partial L_i} = \left(\frac{\partial D}{\partial L}\right)\left(\frac{\partial L}{\partial L_i}\right) + \left(\frac{\partial D}{\partial V}\right)\left(\frac{\partial V}{\partial L_i}\right) + \left(\frac{\partial D}{\partial \sigma}\right)\left(\frac{\partial \sigma}{\partial L_i}\right).
$$

The sum of the products of marginal default values and liabilities over all lines of business can be written as:

$$
\sum_{i=1}^{M} L_i \left(\frac{\partial D}{\partial L_i} \right) = \left(\frac{\partial D}{\partial L} \right) \sum_{i=1}^{M} L_i \left(\frac{\partial L}{\partial L_i} \right) + \left(\frac{\partial D}{\partial V} \right) \sum_{i=1}^{M} L_i \left(\frac{\partial V}{\partial L_i} \right) + \left(\frac{\partial D}{\partial \sigma} \right) \sum_{i=1}^{M} L_i \left(\frac{\partial \sigma}{\partial L_i} \right).
$$
 (3)

The authors will now show that the sum of the first two terms on the right side of Equation (3) equals the default value for the company, and that the third term on the right side equals zero.

The change in the present value of aggregate losses with respect to a change in the present value of losses in a given line of business is equal to one for all lines $(\partial L / \partial L_i = 1)$. The change in the value of assets with respect to a change in the value of liabilities in the *i*th line of business is equal to one plus the corresponding surplusto-liability ratio $(\partial V/\partial L_i = 1 + s_i)$. Therefore,

$$
\sum_{i=1}^{M} L_i \left(\frac{\partial L}{\partial L_i} \right) = L
$$
\n
$$
\sum_{i=1}^{M} L_i \left(\frac{\partial V}{\partial L_i} \right) = V.
$$

We have already seen that $D = (\partial D / \partial L) L + (\partial D / \partial V) V$, so the sum of the first two terms on the right side of the equation is equal to the default value for the company.

This leaves the third term on the right side. It can be shown that the partial derivative of the volatility of the asset-to-liability ratio with respect to the present value of losses in the *i*th line of business is given by:

$$
\frac{\partial \sigma}{\partial L_i} = \frac{1}{L \sigma} \Big[\big(\sigma_{iL} - \sigma_{L}^2 \big) - \big(\sigma_{iV} - \sigma_{LV} \big) \Big],
$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission ww

where σ_{il} is the covariance of log losses in the *i*th line of business with log losses on the insurance portfolio and σ_{ij} is the covariance of log losses on the *i*th line of business with log asset values. The sum of the products of these derivatives and the corresponding liabilities is equal to zero because the portfolio covariances add up. Specifically, (1) the weighted average covariance of policy losses with losses on the portfolio of insurance business is equal to the variance of losses on the portfolio of insurance business, and (2) the weighted average covariance of policy losses with the returns on assets is equal to the covariance of losses on the portfolio of insurance business with the returns on assets:

$$
\sum_{i=1}^{M} x_i \sigma_{iL} = \sigma_L^2
$$
\n
$$
\sum_{i=1}^{M} x_i \sigma_{iV} = \sigma_{LV}
$$
\n
$$
\sum_{i=1}^{M} L_i \frac{\partial \sigma}{\partial L_i} = \sum_{i=1}^{M} x_i \frac{1}{\sigma} \Big[\Big(\sigma_{iL} - \sigma_L^2 \Big) - \Big(\sigma_{iV} - \sigma_{LV} \Big) \Big] = 0.
$$

Substitute back into Equation (3) to confirm that the line-of-business default values add up to the default value for the company as a whole.

DEFAULT VALUES AND SURPLUS REQUIREMENTS FOR LINES OF INSURANCE

We turn now to the formulas for calculating marginal default values. The formulas are derived by asking how a change in the amount of business written in a single line affects the default value for the company as a whole. The authors use the formulas to answer two questions. First, how does changing the amount of business written in a given line of insurance affect the credit quality of the company? Second, how can surplus be allocated to lines of insurance for purposes of costing and pricing policies?

It will be convenient at this point to express the default value as the product of the present value of policy losses and the default-value-to-liability ratio $(d \equiv D/L)$. The default-value-to-liability ratio is a function of the asset-to-liability ratio ($v \equiv V/L$) and (if the distribution of losses and asset returns is joint lognormal) the volatility of the asset-to-liability ratio.¹⁴ Because the asset-to-liability ratio is identically equal to one plus the surplus-to-liability ratio ($v = 1 + s$), the default-value-to-liability ratio can also be written in terms of the surplus-to-liability ratio. Thus $d = f(s, \sigma)$ and $D = Ld$.

A general expression for marginal default values can be obtained by taking derivatives with respect to line-by-line liabilities. Each marginal default value is the sum of two terms, a scale term and a composition term:

$$
d_i \equiv \partial D / \partial L_i = d + \partial d / \partial x_i.
$$

¹⁴ If the distribution is joint normal, the default-value-to-liability ratio is a function of the normalized standard deviation of surplus.

with permission of the copyright owner. Further reproduction prohibited without permission ww

The scale term is the increase in default value due to an increase in the present value of liabilities, ignoring any change in portfolio composition. It is equal to d , the default-to-liability ratio for the insurance company as a whole. The composition term is the increase or decrease in the company default value attributable to changes in the mix of insurance business. Changes in the business mix could affect the aggregate surplus-to-liability ratio and/or the risk characteristics of the asset-to-liability ratio. Expressing marginal default values this way emphasizes the portfolio nature of the problem.

Appendix 2 derives expressions for the marginal default value by line of insurance in the lognormal case. Marginal default values depend on (1) the marginal surplus for the line of insurance, (2) the covariance of losses on the line with losses on other lines of insurance in the portfolio, and (3) the covariance of losses on the line with the returns on the insurance company's assets:

$$
d_i = d + \left(\frac{\partial d}{\partial s}\right)(s_i - s) + \left(\frac{\partial d}{\partial \sigma}\right)\left(\frac{1}{\sigma}\Big[\Big(\sigma_{iL} - \sigma_L^2\Big) - \Big(\sigma_{iV} - \sigma_{LV}\Big)\Big]\right).
$$
 (4)

The option "delta" $(\partial d / \partial s)$ is negative. Therefore, the higher the marginal surplus, the lower the marginal default value, other things being equal. The option "vega" $(\partial d / \partial \sigma)$ is positive. Therefore, the *higher* the covariance of losses with losses on the other lines of insurance in the portfolio, the higher the marginal default value; and the higher the covariance of losses with returns on the portfolio of assets, the lower the marginal default value.

How does expanding or contracting the amount of business written in a single line of insurance affect the credit quality of a multi-line insurance company? It depends on the risk characteristics of the lines, including the covariances of losses on each line with losses on the portfolio of policies and the returns on the company's assets. It also depends on how changes in the amount of business in the relevant line affect the surplus for the company as a whole. In other words, we need to know the marginal surplus line by line.

Suppose, for example, that the company has a policy of maintaining the same ratio of surplus to liabilities for every line of business. Under this policy, each line's marginal surplus equals the surplus ratio for the company: $s_i \equiv \frac{\partial S}{\partial L_i} = s$. The marginal default values by line of business are:

$$
d_i = d + \left(\frac{\partial d}{\partial \sigma}\right) \left(\frac{1}{\sigma} \left[\left(\sigma_{iL} - \sigma_L^2\right) - \left(\sigma_{iV} - \sigma_{LV}\right)\right]\right).
$$
\n(5)

With uniform surplus requirements, the marginal default values depend only on (1) the covariance of losses with losses on the portfolio of insurance and (2) the covariance of losses with returns on the portfolio of assets. Therefore, requiring the same marginal surplus for each line requires that the marginal default values vary by line.¹⁵

¹⁵ There is a special-case exception when $\sigma_{iL} - \sigma_{iL}^2 = \sigma_{iV} - \sigma_{LV}$ for all lines.

Does it make sense to allocate different amounts of default risk to different lines of insurance? The answer is no, because if the company defaults on one policy, it defaults on all policies. Policyholders bear the default risk of the company, not the marginal default risk in a single line of business.¹⁶ The cost of insurance should be calculated accordingly. Specifically, surplus should be allocated to lines of insurance to equalize marginal default values: $d_i \equiv \partial D / \partial L_i = d$. If the default-value-to-liability ratio is the same for all lines of insurance, then surplus must be allocated line by line:

$$
s_i = s - \left(\frac{\partial d}{\partial s}\right)^{-1} \left(\frac{\partial d}{\partial \sigma}\right) \left(\frac{1}{\sigma} \left[\left(\sigma_{iL} - \sigma_L^2\right) - \left(\sigma_{iV} - \sigma_{LV}\right)\right]\right).
$$
 (6)

Because the option delta and vega are negative and positive, respectively, Equation (6) says that the higher the covariance of losses with losses on the other lines of insurance, the *higher* the surplus requirement, and the higher the covariance of losses with returns on assets, the lower the surplus requirement. Finally, note that a weighted average of the marginal surplus requirements s_i equals the overall surplus ratio, because the weighted averages of the covariance terms $\sigma_{iL} - \sigma_L^2$ and $\sigma_{iv} - \sigma_{LV}$ both equal zero.

EXAMPLES

Other things equal, increasing surplus reduces the default value. Given surplus, the default value depends on the uncertainty about losses, the correlation between losses across the company's lines of insurance, the correlation between losses and asset returns, and the uncertainty about asset returns.

The determinants of default values are illustrated in Tables 2 to 5. These examples assume three lines of insurance with the same liability of \$100 per period. Losses are paid in a lump sum at the end of the period. For simplicity, taxes are ignored.

Each table contains two panels. For the moment, the authors concentrate on the top panel, starting in Table 2. Here the standard deviations of losses are 10, 15, and 20 percent, respectively, for lines 1, 2, and 3.¹⁷ The correlation coefficients for the lines' losses are identical at 0.5. The correlation between each line's losses and the return on the company's assets is negative, -0.2 . (This is consistent with a negative beta for L .) The standard deviation of the return on assets is 15 percent.

From these inputs, the authors calculate the covariances of each line's losses with aggregate losses and with the returns on assets. For example, the covariance of Line 1's losses with the total losses of all three lines is 0.0092. Line 1's covariance with the asset returns is -0.0030. The authors also calculate the standard deviation of total losses (σ _L = 0.1236) and the covariance of total losses with the asset returns $(\sigma_{LV} = 0.0045)$.

¹⁶ Here the authors agree completely with Phillips, Cummins, and Allen (1998).

¹⁷ The standard deviations apply to portfolios of policies for each line. Uncorrelated policyby-policy risk is assumed diversified away. The standard deviation of each line's losses reflects factors common to all policies in that line; for example, inflation of claims paid.

Default Value and Surplus Allocations
Risky Assets, Base-Case Correlations

The remaining input is the surplus-to-liability ratio. The authors assume 50 percentthat is, \$150 surplus for \$300 liability. Thus, liabilities are collateralized by \$450 of assets.

Results are provided for two cases: normal (in which the input parameters are used as standard deviations and covariances of losses and assets) and lognormal (in which the inputs are used as standard deviations and covariances of log losses and log assets).¹⁸ For now, the authors focus on the lognormal results. With the inputs in Table 2, the default value per dollar of liability is 0.31 percent for the lognormal case:

$$
d = f(s, \sigma) = 0.0031
$$

\n
$$
s = 0.50
$$

\n
$$
\sigma = \sqrt{\sigma_V^2 + \sigma_L^2 - 2\sigma_{LV}} = 0.2163.
$$

If the industry were obligated to pay the losses in default—for example, through a state pool—a contingent liability costing 0.31 percent of the company's liabilities would be imposed on other companies. If the company had to buy default insurance guaranteeing payment of its losses, it would have to pay 0.31 percent of liabilities.¹⁹

This default value, of course, varies as inputs are changed. Compare Table 2 to Tables 3 through 5:

- In Table 3, a change from risky to safe assets essentially eliminates the default value.
- Table 4 assumes the three lines have the same 15 percent standard deviation but low correlations. (Think of risky property and casualty lines, but written in completely different geographic regions.) The default value is 0.20 percent.
- Table 5 assumes high correlation across lines. (Think of three cohorts of longtailed policies in the same line of business, one issued at date 0, the others one and two periods previously. The remaining losses on the three cohorts are likely to be highly correlated.) The default value rises to 0.43 percent.

Other variations are, of course, possible. In general, the default value increases if surplus is reduced, if the volatility of asset returns or of total losses is increased, and if the correlation between losses and asset returns is reduced. (A positive correlation would provide a partial hedge, reducing the odds of default.)

The authors do not intend Tables 2 through 5 to imply that default values and required surplus-to-liability ratios can be calculated to two or three decimal points. But the examples do give qualitative insights. For example, Table 3 shows how much a

¹⁸ Formulas for marginal default values and sensitivity parameters in the lognormal case are given in Appendix 2. Formulas for average default values, marginal default values, and sensitivity parameters in the normal case are given in Appendix 3.

¹⁹ This assumes that the cost of default insurance is not paid from the assets of the company, but by an additional equity contribution. If so, the assets collateralizing losses remain at \$450. If the cost of default insurance is paid out of assets, collateral decreases and the default value increases.

Default Value and Surplus Allocations Safe Assets Case

Panel A: Portfolio Assets and Lightlities

Panel B: Line-by-Line Allocations

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission w

Default Value and Surplus Allocations
Geographic Diversification Case

Panel B: Line-by-Line Allocations

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission wv

Default Value and Surplus Allocations Long-Tail Case

Panel A: Portfolio Assets and Liabilities

Panel B: Line-by-Line Allocations

move from risky to safe assets reduces the risk of default. With plausible inputs for insurance losses, an insurance company that held safe assets could operate safely at a much lower surplus-to-liability ratio. If the object is to reduce the cost of writing insurance, this model suggests safe assets as the first choice, because in efficient and competitive financial markets, no present value is lost by not investing in risky assets. Of course there would be no point in following such an investment strategy if regulators and rating agencies did not significantly reward the safe-asset company with a lower required surplus ratio.²⁰

The default values discussed up to this point apply to a company, not to individual lines of insurance. For example, the 0.31 percent default value in Table 2 is the cost of guaranteeing payment of total losses on all three lines of insurance. The marginal default values by line of business depend on marginal risk characteristics and marginal surplus allocations.

Consider two policies for allocating surplus. The first allocates the same surplus to each line as a percentage of its liabilities. The second allocates surplus so that the marginal default value for each line is equal to the default value for the company.

The bottom panel of Table 2 shows the two cases. If surplus is fixed at 50 percent of each line's liabilities, then the marginal contribution of each line to the total default value varies. For example, the 0.62 percent default value for Line 3 means that a \$1 increase in Line 3 liability would increase the total default value by 0.62 cents.²¹ A \$1 increase in Line 1 liability would increase the total default value by 0.02 cents. Line 1's losses have a low standard deviation and low covariances with total losses and assets. So adding Line 1 business supported by a 50 percent surplus-to-liability ratio adds almost nothing to the value of the default option.

If surplus is allocated so that each line's marginal contribution to the total default value is equal, then surplus allocations vary. Line 1's surplus-to-liability ratio is 38 percent, Line 2's is 50 percent, and Line 3's is 63 percent. If an additional \$1 of Line 1 liability is supported by 38 cents of additional surplus, all default values, total and allocated, remain at 0.31 percent.²² The required surplus-to-liability ratio would therefore decline slightly: the marginal ratio is 38 percent, not 50 percent. A similar \$1 increase in the Line 3 liability would not affect default values if the surplus allocated to Line 3 is simultaneously increased by 63 cents. The overall surplus-to-liability ratio would increase slightly.

Tables 2 through 5 also report results based on the assumption that the underlying losses and asset returns have joint normal distributions. Notice that the portfolio de-

²⁰ If assets are liquid (i.e., easily bought and sold on short notice), then the risk of assets can change on short notice. That means a regulator could not reward a safe-asset company without checking asset composition frequently. The company's investment strategy would have to be transparent to the regulators. If these conditions are too costly or awkward to meet, regulators may have to assume that assets are not safe. See Myers and Rajan (1998).

²¹ The increase in Line 3 liability changes the portfolio weights slightly, and this in turn affects default values. This is, however, a second- or third-order effect, so the authors ignore it in discussing these examples.

²² Adding \$1 to one line's liability changes the portfolio weights, feeding back into surplus requirements. But for reasonable numerical values, the feedback is exceedingly small.

fault values (Panel A) are greater in the normal case than in the lognormal case. In Table 2, for example, the default value in the normal case is 0.43 percent versus 0.31 percent in the lognormal case. Notice, too, that the deviations of marginal default values and surplus requirements around the portfolio values (Panel B) are somewhat less in the normal case relative to the lognormal case. In Table 2, marginal surplus requirements range from 41 to 59 percent in the normal case versus 38 to 63 percent in the lognormal case, for example. These results reflect the symmetry of the normal probability distribution in relation to the positive skewness of the lognormal.

These results, of course, depend on the input assumptions, but given the assumptions, the marginal default values and surplus allocations are unique and not arbitrary. The weighted average of the marginal default values is equal to the default value for the company, and the weighted average of the marginal surplus-to-liability ratios is equal to the surplus-to-liability ratio for the company. The authors believe that this result and their general procedure for calculating marginal default values and surplus allocations are not in the insurance literature.

IMPLICATIONS FOR PRICING

Marginal default values in a multiple-line insurance company depend not only on line-by-line risk characteristics but also on the mix of insurance the company writes. Therefore, when the mix of insurance changes, the marginal default values change, too. Are marginal default values—and surplus allocations derived from marginal default values—sufficiently robust to changes in the mix of business to be useful for costing and pricing insurance?

To shed light on this question, consider the line-by-line surplus requirements for a series of hypothetical insurance companies with varying degrees of diversification. Suppose each of these companies starts with an equal fraction of its liabilities in each of N existing lines of insurance and then adds liabilities in an $N + 1$ st "new" line. The fraction of liabilities in the $N+1^{st}$ line increases from zero to $1/(N+1)$, and the fractions in each of the N existing lines simultaneously decrease from $1/N$ to $1/(N+1)$. We hold default value constant at 2.24 percent of liabilities. Consider what happens to the required surplus ratio (1) for the company, (2) for the existing line(s) of insurance, and (3) for the new line of insurance. To keep the experiment as simple as possible, the authors assume that the standard deviation of losses is 30 percent in every line of insurance, that losses in every line are uncorrelated with losses in every other line, that losses are uncorrelated with the return on assets, and that the standard deviation of the return on assets is 15 percent.

Figure 1 displays results for three cases: a company that operates in two lines of insurance, a company that operates in four lines, and a company that operates in ten lines.²³ Look first at the results for the two-line company. When the new (2nd) line of insurance is absent from the portfolio, the surplus requirement for the company and the marginal surplus requirement for the existing (1st) line are 50 percent, whereas the marginal surplus requirement for the new line is negative at -24 percent. As the new line becomes a larger fraction of liabilities, the surplus requirements for the com-

 23 These results were computed using the lognormal formula for surplus allocation [Equation (6)].

FIGURE 1

Surplus Requirements and Diversification

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.ww

pany and the existing line both fall and the marginal requirement for the new line rises. When the new line is a quarter of the liabilities, for example, the requirement for the company is 36 percent, the marginal requirement for the existing line is 45 percent, and the marginal requirement for the new line is 8 percent. The requirements for the company and the existing line continue to decline, and the marginal requirement for the new line continues to rise until the new line is an equal fraction (one-half) of liabilities, at which point all three requirements are equal at 31 percent.²⁴ Notice that the marginal requirement for the existing line declines much more slowly than the marginal requirement for the new line rises.

The marginal requirement for the new line is much more sensitive to changes in the composition of business than is the existing line. Therefore, setting a surplus allocation for a new line is more difficult. The surplus required for a new line's initial business will be quickly out of date as the line grows.

The four- and ten-line companies exhibit the same patterns as does the two-line company, with the marginal surplus requirements for existing lines declining slowly and the marginal requirement for the new line rising rapidly as the fraction of liabilities in the new line increases from zero up to an equal share. For example, as the new line increases from zero to a tenth of liabilities in the ten-line case, the marginal surplus requirement for existing lines falls from 50 to 49 percent, while the requirement for the new $(10th)$ line rises from 28 to 49 percent. These results suggest that for companies with even a modest degree of diversification, marginal surplus requirements for existing lines of business are reasonably robust to the introduction of a new line of business, and that marginal surplus requirements for existing lines can be approximately correct even as new lines are added or existing lines phased out.

IMPLICATIONS FOR REGULATION

To calculate the cost of insurance on a line-by-line basis, surplus must be allocated in such a way as to equalize marginal default values across lines. In the context of regulated, retail insurance allocating surplus to equalize marginal default values is necessary to avoid cross-subsidies and to eliminate incentives for insurance companies to push riskier lines.

Consider regulation Massachusetts-style, where "fair" premiums are calculated for different lines of insurance. Suppose that regulators assume the same surplus-toliability ratios for all lines (and companies) and calculate retail premiums ignoring default risk. (This is in fact current practice in Massachusetts.) Then companies concentrating on riskier lines—that is, lines with higher marginal default values—will be subsidized, through the industry pool, by companies concentrating on lines with low marginal default values.

Take Table 2 as the example. If each line's premiums are calculated assuming a 50 percent surplus-to-liability ratio, as in Panel A, then Line 3 is at the margin 0.60 percent more profitable, on a present value basis, than Line 1. Companies concentrating on Line 3 would in effect be subsidized by the industry as a whole. All companies would be tempted to push Line 3 and to cut back on Line 1 policies.

²⁴ If the liabilities in the new line increase beyond that point, the surplus requirement for the company begins to rise again.

The problem with default insurance provided by an industry pool is that the insurance is not explicitly paid for. Each company sees the benefit and cost as:

If a company increases the benefit, say, by investing in riskier assets, the cost does not increase; it is fixed.²⁵ There is no mechanism forcing the company to pay more for its default insurance.²⁶

The moral hazard has nothing to do with regulation of insurance premiums. The same temptations arise in jurisdictions where regulators are concerned with solvency only. If they express that concern by mandating a uniform surplus-to-liability ratio, companies are again tempted to shift to riskier assets or riskier lines of business.

Regulators could in principle eliminate the moral hazard by varying required surplus depending on asset risk and on the proportions of liabilities in different lines of insurance. The objective should be to equalize total default values company by company²⁷ and marginal default values line by line. (Note again that default values are expressed as a percentage of liabilities.)

If this objective is reached, the moral hazard disappears. Imagine an insurance industry in which each company matches the example in Table 2. Each company sells three lines of insurance in equal proportions and holds equally risky assets. All default values equal 0.31 percent of liabilities.

Suppose the insurance regulator sets the following rules:

- 1. The base-case surplus-to-liability ratio is 50 percent.
- 2. If a company changes its composition of business—i.e., moves a greater proportion of its business to a particular line—then its required surplus will be adjusted according to the marginal surplus-to-liability ratios given in Panel B.
- 3. If a company changes the risk of its assets,²⁸ its required surplus will be adjusted to keep the company's total default value at 0.31 percent.
- ²⁵ The cost is not totally fixed. If a company defaults on loss payments to policyholders, it goes out of business and defaults on its obligation to the pool. Thus, an action that increases the value of default insurance from a pool also decreases the present value of the company's obligation to the pool. Note that this effect reinforces the moral hazard created by participation in the pool.
- ²⁶ The regulators cannot penalize a risk-taking company by reducing the premiums it is allowed to charge. That would give it a competitive advantage.
- ²⁷ In a competitive insurance market with full information, there would be no need to equalize default values across companies. Riskier companies would simply receive lower premiums. This is a zero-NPV result as long as wholesale buyers have enough information and sophistication to evaluate credit risk accurately. The main problem would be the insurance company's incentive to bait and switch: it would be tempted to sell an *initial* portfolio of policies and then to try to surprise policyholders with a shift into more risky assets or lines of insurance. If successful, this trick would increase the default value, reduce the present value of the initial policies, and increase the value of the company's equity. On the other hand, the possibility that the company would forfeit intangible assets (e.g., its reputation) in the event of default could mitigate this temptation.
- ²⁸ A change in risk might mean a change in the correlation of asset returns with losses.

With these rules, there is no moral hazard—no artificial incentive to push risky lines or increase asset risk. The benefit and cost of participating in the industry pool are equal for every company. Each company receives default insurance worth 0.31 percent of liabilities and assumes a pro rata share of the cost of providing default insurance to other companies. But because all companies' default values are equal, the present value of each company's pro rata share of the default costs of the industry pool is also 0.31 percent of that company's liabilities.

EFFICIENT SURPLUS REQUIREMENTS

The examples in this study imagine a degree of fine tuning that is impossible in real life. But the example explains the desired result—elimination of moral hazard—and shows how varying required surplus-to-liability ratios can achieve that result. The example is incomplete because it assumes a base-case composition of business. What compositions are likely to be efficient? What are the implications for regulated surplus requirements and premiums?

Double taxation means that surplus is costly. Yet insurance companies have to carry enough surplus to keep default values acceptably low, low enough to satisfy regulators (and wholesale customers in unregulated markets).²⁹ The maximum acceptable default value is a constraint on surplus.

Given this constraint, surplus can be reduced by diversification across lines of business or by reducing asset risk. In this context, effective diversification means adding lines with low correlations of losses and if possible high correlations of losses with asset returns. (Ideally, losses would hedge asset returns.) Geographic diversification also helps.

The marginal financial gains from such diversification are high at first for a company starting with one or a few highly correlated lines. These gains decline as more lines and new geographic areas are added. Then, at some point, costs increase.³⁰ Diversification proceeds until the marginal gain from reducing surplus equals marginal cost (See Figure 2).

Efficient diversification does not minimize required surplus. It minimizes the total cost of issuing, administering, and collateralizing policies. This establishes the efficient composition of business.

In practice, this efficient composition will not be unique or sharply defined. For example, a company may specialize in one or a few closely related lines of insurance and still compete effectively. It can do so if its costs are low enough to offset the tax cost of the extra surplus it has to carry to ensure creditworthiness. Note that the extra surplus does not justify a higher premium. The specialist company must match the diversified companies' premiums to compete and survive.

²⁹ As a practical matter, even sophisticated wholesale customers would find it inefficient to buy policies from a company with a high risk of default. The customer would have to investigate the seller, price out the policy for default risk, and continue to monitor the seller's financial health. All of these activities are costly.

³⁰ Reinsurance reduces risk and required surplus, but it, too, is costly.

FIGURE 2

The trade-off of reduced surplus requirements against operating and administrative costs determines the efficient composition of business. As more lines are added, diversification decreases required surplus, but operating, administrative, and perhaps agency costs increase.

Efficient diversification does not minimize required surplus. It minimizes the total cost of issuing, administering, and collateralizing policies. This establishes the efficient composition of business.

In practice, this efficient composition will not be unique or sharply defined. For example, a company may specialize in one or a few closely related lines of insurance and still compete effectively. It can do so if its costs are low enough to offset the tax cost of the extra surplus it has to carry to ensure creditworthiness. Note that the extra surplus does not justify a higher premium. The specialist company must match the diversified companies' premiums to compete and survive.

If premiums are regulated—that is, calculated by the regulators—the implications of this analysis are as follows:

- Premiums should assume a base case of efficient diversification (that is, an effi- \bullet cient composition of business) and normal asset risk. Premiums should cover all costs, including the tax costs of required surplus.
- The required surplus for a line of insurance should depend on the line's risk and the correlations of its losses with other lines' losses and with asset returns. The formal allocation rules are described above and illustrated in Tables 2 to 5.
- No company should be required to adopt a particular composition of business. But if a company deviates from the base case in a manner that would increase default risk, its required surplus should be adjusted up to keep default values at or below the maximum acceptable level. If a deviation reduces default risk, the company should be allowed to carry less surplus.

Implementation of these principles in real life faces various practical and political problems. Two examples are given here. The first problem of implementation is to gain a clearer picture of the costs of diversification in insurance. This may go beyond

the increase in the costs of administering dissimilar operations. Consider the conglomerate discount that is, the tendency of diversified companies to sell in the stock market for much less than the sum of the values of their parts. It may be that excessive diversification brings managerial slackness or complacency. There is also recent research showing that diversified companies do not allocate capital efficiently.³¹

Second, regulation can in practice encourage companies to operate inefficiently. For example, too-tight state regulation can drive out national companies and encourage local ones, thereby increasing default risk.

The base-case composition of property and casualty business would almost surely be geographically diversified. A company writing policies in only one state would have to carry extra surplus to achieve default values as low as a national company. When a state regulator squeezes a national company's premiums, the benefits go to in-state policyholders and the costs are absorbed by shareholders in all states. Default is a second- or third-order concern, because the national company's surplus secures all its business. The national surplus cross-subsidizes the in-state policies.

If the regulatory squeeze is too long or too hard, the national companies gradually 32 leave the state, and local companies take over their business. The local companies' surplus has to be higher, other things equal, or default risk increases because of less effective diversification. The higher surplus is inefficient because of the double tax on corporate income. But now the regulator has to face up to default risk: it is no longer spread out nationally, but internalized.

CONCLUSIONS

This study clarifies the implications of option pricing methods for surplus requirements by line of insurance. The importance of surplus levels for insurance regulation is obvious. But here the authors are presenting theory and principles, not detailed recommendations for practical pricing and regulation. Before doing so, several issues would have to be solved or in some way accommodated. For example:

- How much diversification is efficient? Is there in practice a clear "efficient" composition of business from which base-case surplus requirements can be set? Should regulators demand more surplus from companies that do not diversify across lines? The authors have mentioned that price regulation by states encourages companies to specialize by state. Such companies' higher default risks, compared to geographically diversified companies, are an indirect cost of the regulation.
- How should the inputs necessary to calculate a line's contribution to default risk be estimated? As the examples in this study show, default value depends not just on uncertainty about the line's losses but also on the correlation with other lines' losses and with asset returns. Also, lines with long tails of losses contribute to default value—because ultimate losses are still uncertain—for years after the premiums have been received.

³² They do not leave immediately because of the cost of re-entry if in-state profitability improves, and also because state regulators may set up exit barriers or charges.

³¹ See, for example, Lamont (1997). Possible reasons for the conglomerate discount are reviewed in Myers (1999).

- How much default risk is acceptable? Conservatively high surplus requirements can make default values so low that moral hazards effectively disappear. But in this case customers pay more for insurance because of the costs of the extra surplus.
- The easiest way to reduce surplus requirements is to reduce asset risk. What are the costs of doing so? In efficient securities markets, there is no loss of present value when risky assets are replaced with safe ones.

This study has concentrated on regulatory issues, especially in the latter parts. But the authors' analysis has managerial implications for costing and pricing insurance. To set the premiums for a policy, an insurance company must estimate the surplus required to support that policy. This study shows how required surplus depends on the composition of the company's business, the risk and correlations of losses line by line, the risk of its assets, etc. The authors believe that the numerical procedures illustrated by Tables 2 through 5 will have some practical application. Precise input numbers will not be available, but rough and reasonable estimates should be possible. These would suffice for qualitative analyses and what-if calculations.

Finally, although the authors have limited this study to insurance, they believe that the proofs that default values and surplus (i.e., capital) requirements "add up" apply generally to all financial institutions and should have implications for risk management and regulation in banking, securities, and derivatives trading, etc.

APPENDIX 1

This appendix proves generally that marginal default values "add up" and that surplus allocations are therefore unique and not arbitrary. The proof requires no assumptions about the joint probability distributions of line-by-line losses and returns on the firm's portfolio of assets. The only requirement is frictionless financial markets and fixed state-contingent prices for all relevant outcomes.

For simplicity, we will consider two lines, a and b, and two dates, $t = 0$ and $t = 1$. Present values (PVs) and outcomes are:

The outcomes are expressed as the product of the initial present values and gross returns \overline{R}_V , \overline{R}_a , and \overline{R}_b . The gross returns are the state variables. Each can be zero but not negative.

Each line has a surplus requirement, denoted s_a or s_b . From the balance sheet identity,

$$
V = L_a (1 + s_a) + L_b (1 + s_b) = L (1 + s)
$$

where $L = L_a + L_b$ and s is a weighted average of s_a and s_b . The default value D is:

$$
D = \text{PV}\bigg[\text{Max}\bigg(L_a\tilde{R}_a + L_b\tilde{R}_b - V\tilde{R}_V, 0\bigg)\bigg]
$$

The authors assume that the assets are safe and the returns R_v known. They explain below how their proof can be extended to accommodate an uncertain return \overline{R}_v .

Figure 3 plots the region in which default occurs and the payoff to the default value is positive. The two axes are \tilde{R}_a and \tilde{R}_b . If $\tilde{R}_b = 0$, default occurs only when the loss on Line *a* exceeds VR_V , that is, when \tilde{R}_a exceeds $\overline{R}_a = R_V V / L_a$. When $\tilde{R}_a = 0$, default requires that \tilde{R}_h exceeds $\overline{R}_h = R_V V / L_h$. Thus the default region lies in Z , above a line connecting \overline{R}_a and \overline{R}_b .

FIGURE 3

Conditions for Default

The payoff to the default option is positive if the sum of losses $L_a \overline{R}_a + L_b \overline{R}_b$ exceeds the end-of-period asset value $V\tilde{R}_V$. Given \tilde{R}_V , positive payoff occurs in Region Z, above the line defined by the intercepts \overline{R}_a and \overline{R}_b .

We assume well-defined market prices for the return distributions \tilde{R}_{v} , \tilde{R}_{a} , and \tilde{R}_{b} . If that assumption holds, we can write the present value of D as:

$$
D = \int_{Z} \pi(z) \big(L_a R_a(z) + L_b R_b(z) - VR_v \big) dz.
$$
 (A1-1)

Here $\pi(z)$ is a state-price density for a state z, defined by the paired outcome $(\tilde{R}_a, \tilde{R}_b)$.¹ The integral over Z means a summation over all possible states that fall within Re-

¹ In general, z and $\pi(z)$ would be defined by the triple outcome $\left(\tilde{R}_{a}, \tilde{R}_{b}, \tilde{R}_{v}\right)$.

gion Z in Figure 3. Note that a proportional change in L_a, L_b , and V does not change the location of the line in Figure 3. Therefore, it does not change D as a fraction of V or L .

The derivative of D with respect to L_a and L_b is somewhat complicated. A change in L_a or L_b shifts the boundary for the states in which default occurs. Both intercepts of the boundary line in Figure 3 $(\overline{R}_a$ and \overline{R}_b) are affected. The derivatives of the intercepts are:

$$
\frac{\partial \overline{R}_a}{\partial L_a} = -\frac{R_V L_b (1 + s_b)}{L_a^2}
$$
 (A1-2a)

$$
\frac{\partial \overline{R}_a}{\partial L_b} = \frac{R_V (1 + s_b)}{L_a} \tag{A1-2b}
$$

and

$$
\frac{\partial \overline{R}_b}{\partial L_a} = \frac{R_V (1 + s_a)}{L_b} \tag{A1-2c}
$$

$$
\frac{\partial \overline{R}_b}{\partial L_b} = -\frac{R_V L_a (1 + s_a)}{L_b^2}.
$$
\n(A1-2d)

We can now show that line-by-line marginal default values add up; that is:

$$
L_a \frac{\partial D}{\partial L_a} + L_b \frac{\partial D}{\partial L_b} = D \tag{A1-3}
$$

The marginal default values can be expressed as:

$$
\frac{\partial D}{\partial L_a} = \int\limits_Z \pi(z) \Big(R_a(z) - R_V \left(1 + s_a \right) \Big) dz + G_a \frac{\partial \overline{R}_a}{\partial L_a} + G_b \frac{\partial \overline{R}_b}{\partial L_a} \tag{A1-4a}
$$

$$
\frac{\partial D}{\partial L_b} = \int\limits_Z \pi(z) \Big(R_b(z) - R_V \big(1 + s_b \big) \Big) dz + G_a \frac{\partial \overline{R}_a}{\partial L_b} + G_b \frac{\partial \overline{R}_b}{\partial L_b} \,. \tag{A1-4b}
$$

Here G_a represents the *change* in value of D as a function of a shift in the vertical intercept of the line-boundary in Figure 3. That is, $G_a \equiv \partial D / \partial R_a$. G_b is the change in value due to a shift in the horizontal intercept.

Calculate the weighted as here defined by Equation $(A1-3)$, using the derivatives in Equations (A1-4a) and (A1-4b). This gives

$$
\int_{Z} \pi(z) \Big(L_a R_a(z) + L_b R_b(z) - R_V \Big(L_a \big(1 + s_a \big) + L_b \big(1 + s_b \big) \Big) \Big) dz
$$
\n
$$
+ L_a G_a \frac{\partial \overline{R}_a}{\partial L_a} + L_a G_b \frac{\partial \overline{R}_b}{\partial L_a} + L_b G_a \frac{\partial \overline{R}_a}{\partial L_b} + L_b G_b \frac{\partial \overline{R}_b}{\partial L_b}.
$$
\n(A1-5)

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission w

But a check of the derivatives in Equations $(A1-2a)$ and $(A1-2c)$ confirms that the terms in G_a cancel. Equations (A1-2b) and (A1-2d) imply that the terms in G_b also cancel. Then the first term in Equation $(A1-5)$ equals D by definition, thus proving Equation (A1-3) and establishing that marginal default values add up.

We can repeat this analysis conditional on any future asset return \tilde{R}_{v} . Because the proof follows for all possible values of \overline{R}_V , it immediately generalizes to the case of uncertain asset returns. The proof also generalizes when there are three or more lines of insurance.

The marginal surplus requirements are set at $t = 0$ and are constants in Equations (A1-4a) and (A1-4b). Marginal default values add up regardless of how s_a and s_b are determined. However, s_a and s_b can be set to equate $\partial D/\partial L_a$ to $\partial D/\partial L_b$. As explained in the main text of the study, this is the most logical procedure for purposes of calculating the cost of insurance.

APPENDIX 2

This appendix provides the formulas used to compute the numerical results reported in Table 1 and in Tables 2 through 5 for the cases in which the probability distribution of losses and asset values is joint lognormal. It also derives the formula for marginal default values by line of business.

Default Values and Sensitivity Parameters

If the probability distribution of losses and asset values is joint lognormal, the default-value-to-liability ratio for a company is a function of its surplus-to-liability ratio and the volatility of its asset-to-liability ratio. To calculate the results reported in Tables 2 through 5, the authors applied the following formula:

$$
d = f(s, \sigma) = N\{z\} - (1+s)N\{z - \sigma\},
$$

where $N\{\bullet\}$ denotes the cumulative probability function for the standard normal variable and

$$
z = \frac{-\ln(1+s)}{\sigma} + \frac{1}{2}\sigma
$$

$$
\sigma = \sqrt{\sigma_L^2 + \sigma_V^2 - 2\sigma_{LV}}
$$

The partial derivatives of the default-value-to-liability ratio with respect to the surplus-to-liability ratio and the volatility of the asset-to-liability ratio were computed as follows:

$$
\frac{\partial d}{\partial s} = -N\{z - \sigma\}
$$

$$
\frac{\partial d}{\partial \sigma} = N'\{z\},
$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission_W

where $N^1\{\bullet\} - N'$ denotes the probability density function for the standard normal variable. These parameters are referred to in the tables as "delta" and "vega," respectively.

Marginal Default Values

Consider a small change in the fraction of the insurance portfolio written in the ith line of insurance. A change in the mix of business could affect the default-value-toliability ratio by changing (1) the surplus-to-liability ratio or (2) the volatility of the asset-to-liability ratio:

$$
\frac{\partial d}{\partial x_i} = \left(\frac{\partial d}{\partial s}\right)\left(\frac{\partial s}{\partial x_i}\right) + \left(\frac{\partial d}{\partial \sigma}\right)\left(\frac{\partial \sigma}{\partial x_i}\right).
$$

The change in the surplus-to-liability ratio with respect to a change in the fraction of liabilities in the ith line equals the difference between the surplus-to-liability ratio in the *i*th line of business and the surplus-to-liability ratio for the company:

$$
\frac{\partial s}{\partial x_i} = s_i - s
$$

The partial derivative of the volatility of the asset-to-liability ratio with respect to the fraction of liabilities in the *i*th line is:

$$
\frac{\partial \sigma}{\partial x_i} = \frac{1}{\sigma} \Big[\Big(\sigma_{iL} - \sigma_L^2 \Big) - \Big(\sigma_{iV} - \sigma_{LV} \Big) \Big].
$$

Putting these results together gives the general formula for calculating marginal default values on a line-by-line basis:

$$
d_i = d + \left(\frac{\partial d}{\partial s}\right)(s_i - s) + \left(\frac{\partial d}{\partial \sigma}\right)\left(\frac{1}{\sigma}\Big[\Big(\sigma_{iL} - \sigma_L^2\Big) - \Big(\sigma_{iV} - \sigma_{LV}\Big)\Big]\right).
$$

APPENDIX 3

This appendix provides the formulas used to compute the numerical results reported in Tables 2 through 5 for the cases in which the probability distribution of losses and asset values is joint normal. It also derives the corresponding formula for marginal default values by line of business.

Default Values and Sensitivity Parameters

If the probability distribution of losses and asset values is joint normal, the defaultvalue-to-liability ratio for a company is a function of its surplus-to-liability ratio and the normalized standard deviation of surplus θ (the standard deviation of surplus per dollar of liabilities). To calculate the results reported in Tables 2 through 5, the authors applied the following formula:¹

¹ The authors derived this formula by assuming that asset and liability values are expected to appreciate at the risk-free rate of interest and then computing and discounting payoffs to the default option accordingly.

$$
d = f(s, \theta) = -sN\{-z\} + \theta N'\{z\},
$$

where $N\{\bullet\}$ and N' denote the cumulative probability function and the probability density function, respectively, for the standard normal variable, and:

$$
z = \frac{s}{\theta}
$$

$$
\theta = \sqrt{\theta_L^2 + (1+s)^2 \theta_V^2 - 2(1+s)\theta_{LV}}
$$

The partial derivatives of the default-value-to-liability ratio with respect to the surplus-to-liability ratio and the standard deviation of surplus were computed as follows:

$$
\frac{\partial d}{\partial s} = -N \{-z\}
$$

$$
\frac{\partial d}{\partial \theta} = N' \{z\}.
$$

Marginal Default Values

In this case, a change in the mix of business could affect the default-value-to-liability ratio by changing (1) the surplus-to-liability ratio or (2) the standard deviation of surplus:

$$
\frac{\partial d}{\partial x_i} = \left(\frac{\partial d}{\partial s}\right)\left(\frac{\partial s}{\partial x_i}\right) + \left(\frac{\partial d}{\partial \theta}\right)\left(\frac{\partial \theta}{\partial x_i}\right).
$$

The change in the surplus-to-liability ratio with respect to a change in the fraction of liabilities in the ith line equals the difference between the surplus-to-liability ratio in the ith line of business and the surplus-to-liability ratio for the company:

$$
\frac{\partial s}{\partial x_i} = s_i - s
$$

The partial derivative of the standard deviation of surplus with respect to the fraction of liabilities in the ith line is:

$$
\frac{\partial \theta}{\partial x_i} = \frac{1}{\theta} \Big[\Big(\theta_{iL} - \theta_{L}^2 \Big) - (1+s) \Big(\theta_{iV} - \theta_{LV} \Big) + \Big(s_i - s \Big) \Big((1+s) \theta_{V}^2 - \theta_{LV} \Big) \Big].
$$

Putting these results together gives the general formula for calculating marginal default values on a line-by-line basis:

$$
d_i = d + \left(\frac{\partial d}{\partial s}\right)(s_i - s) + \left(\frac{\partial d}{\partial \theta}\right)\left(\frac{1}{\theta}\Big[\left(\theta_{iL} - \theta_L^2\right) - \left(1 + s\right)\left(\theta_{iV} - \theta_{LV}\right) + \left(s_i - s\right)\left(\left(1 + s\right)\theta_V^2 - \theta_{LV}\right)\Big]\right)
$$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission_W