JOURNAL OF FINANCIAL INTERMEDIATION **6,** 3–38 (1997)
ARTICLE NO. JF960205

Price, Financial Quality, and Capital Flows in Insurance Markets

J. DAVID CUMMINS AND PATRICIA M. DANZON

The Wharton School, University of Pennsylvania, Philadelphia, Pennsylvania 19104

Received October 5, 1994

This paper develops a model of price determination in insurance markets. Insurance is provided by firms that are subject to default risk. Demand for insurance is inversely related to insurer default risk and is imperfectly price elastic because of information asymmetries and private information in insurance markets. The model predicts that the price of insurance, measured by the ratio of premiums to discounted losses, is inversely related to insurer default risk and that insurers have optimal capital structures. Price may increase or decrease following a loss shock that depletes the insurer's capital, depending on factors such as the effect of the shock on the price elasticity of demand. Empirical tests using firm-level data support the hypothesis that the price of insurance is inversely related to insurer default risk and provide evidence that prices declined in response to the loss shocks of the mid-1980s. *Journal of Economic Literature* Classification Numbers: G22, G32, G33. © 1997 Academic Press

In the mid-1980s, the market for general liability insurance experienced a "crisis," triggered by sharp increases in losses that caused a large negative shock to the equity capital of the insurance industry in 1984. Following the shock, aggregate premiums written for general liability (GL) insurance increased at over 70% per year in 1985 and 1986,[1] and limitations on the availability of coverage were widely reported (U.S. Department of Justice, 1986).[2] Insurers responded to the shock by raising unprecedented amounts

[1] This understates the increase in premium rates, assuming that demand is not totally inelastic, and understates the cost increase for a given level of real protection—expected out-of-pocket cost—because of increased use of deductibles, lower coverage limits, and policy exclusions. A survey of large corporations showed a median rate increase between 1985 and 1986 of 54% for GL primary coverage and 214% for GL excess coverage (U.S. General Accounting Office, 1988).

[2] To a lesser extent other insurance coverages also experienced unusual premium growth rates during this period. Premium growth in lines of insurance other than general liability was 19% per year during the 1985–1986 period, compared with about 5% in the early 1980s. General liability premium growth was near zero in the early 1980s. The data referenced here are from A.M. Best Company, "Best's Aggregates and Averages: 1990 Edition," Oldwick, NJ, 1990).

1042-9573/97 $25.00

of new equity capital in 1985 and 1986. Considerable debate has focused on the extent to which events during the crisis represented a rational response to increased loss expectations and capital depletions or resulted from insurance market and/or capital market imperfections. The purpose of this paper is to provide new information to help resolve the controversy by developing a theoretical model of the relationship between loss shocks, capitalization, and prices in insurance markets and testing the model using a pooled cross section, time series sample of insurers over a period including the 1980s liability insurance crisis.

The liability crisis has been interpreted as a particularly acute phase of the insurance underwriting cycle, in which periods of "hard" markets, with rising prices, reductions in coverage, and increases in deductibles, alternate with "soft" markets, with falling prices and ready availability of insurance. One important line of research has focused on capital shortages as a possible cause of underwriting cycles. In standard "arbitrage" models of insurance pricing (e.g., Kraus and Ross, 1982; Myers and Cohn, 1987), the price of insurance reflects the discounted value of expected losses and expenses, adjusted for taxes and market risk. Capital is assumed to be freely available as long as insurance is priced to yield the appropriate rate of return for securities of comparable risk. Capital shortages would not occur in the arbitrage model, and the model cannot explain why insurance would be unavailable or "unfairly" priced in the absence of regulatory intervention or some market imperfection.

The principal alternative to the arbitrage model that has been advanced to explain prices and capital structure in insurance markets is the "capacity constraint" theory (Gron, 1989, 1994; Winter, 1994).[3] The premise of this theory is that capital shortages and overages resulting from capital market imperfections are the primary cause of hard and soft markets for insurance, including the crisis of the 1980s. Insurance risk is assumed to be imperfectly diversifiable so that insurers must hold equity capital to ensure payment of claims that are larger than expected. The insurance industry is treated as a single price-taking firm and is assumed to be constrained, either by infinitely risk averse policyholders or by regulation, to write only that volume of business consistent with zero (or negligibly small) insolvency risk. Maintaining financial capital is assumed to be costly for insurers, due to factors such as corporate taxation and agency costs. Internal capital is assumed to be less costly than external capital, because "a cost is incurred in the 'round-trip' of paying out a substantial amount of retained earnings

[3] Among the other factors discussed as possible causes of the crisis are adverse selection, uncertainty about liability rules and interest rates, and underpricing by insurers (see, for example, Doherty and Kang, 1988; Harrington, 1988; Priest, 1987; Clarke *et al.*, 1988). Harrington and Danzon (1994) analyze price cutting in soft markets.

… and then immediately raising the same amount through the issuance of equity'' (Winter, 1994, pp. 382–383) or because of asymmetric information.

Because of the entry and exit costs for capital, the industry is hypothesized to go through periods of excess capacity, where the amount of capital in the industry is excessive relative to the demand for insurance, followed by periods of capital shortages triggered by loss shocks from imperfectly diversified risks that cause the zero insolvency-probability constraint to become binding. Prices are low relative to the present value of losses during periods of excess capacity, but increase when capacity falls as the price of existing capacity is bid up in the insurance market. The primary empirical prediction of the model is that premiums are not unbiased predictors of expected losses (as the arbitrage theory suggests) but rather that the difference between premiums and expected losses is inversely related to the stock of financial capital. Thus, premiums are predicted to be relatively low when capital is high and high when capital is low.

Although the capacity constraint theory provides useful insights into the operation of insurance markets, there are several reasons to reexamine the relationship between capital and price in insurance. First, recent insolvency experience casts doubt on the assumption that insurance is virtually free of insolvency risk.[4] This assumption plays a key role in the capacity constraint theory's prediction of an inverse relationship between price and capital. Likewise, the version of the arbitrage model that authors like Winter and Gron use to motivate their analyses does not incorporate financial quality.[5] More recent versions of the model view insurance pricing as analogous to the pricing of risky corporate debt (e.g., Doherty and Garven, 1986; Cummins, 1988). Such models predict a positive relationship between the price of insurance and capitalization, contrary to the capacity constraint theory, on the hypothesis that insurance prices will reflect the expected loss to policyholders due to insolvency (the insolvency put option). Second, the existing empirical evidence provides mixed support for the capacity constraint theory of the mid-1980s crisis. Winter's (1994) results, using aggregate, industry-wide data, support the theory for cycles prior to 1980 but

[4] Insolvencies of property-liability insurers increased from about 10 per year for the period 1969–1983 to 32 per year in 1984–1990, while average annual assessments by guaranty funds increased from $37 million to $492 million over the same period (Cummins *et al.* 1994). In recognition of the growing insolvency problem, the National Association of Insurance Commissioners (NAIC) adopted an aggressive ''Solvency Policing Agenda'' in 1989 that includes more extensive financial ratio and auditing tests of insurer solvency and an accreditation program for state insurance departments (see Klein, 1995). The NAIC adopted rigorous risk-based capital requirements for property–liability insurers in 1993. The private monitoring of insurer solvency has also increased since the early 1980s with several firms entering the insurance monitoring business since that time.

[5] We define a firm's financial quality as reflecting its default risk, with high quality firms having low default risk.

not for the period of the 1980s liability crisis. Gron's (1994) empirical results support the theory for short-tail lines of insurance but not for long-tail coverages such as general liability insurance, the line most affected by the crisis.[6] Third, the theory and the existing empirical analyses assume that the industry operates as a single, price-taking firm, and thus do not provide information on the cross-sectional relationship between price and capital. And, fourth, existing empirical analyses of liability insurance markets have not considered the role of capital flows, even though insurers raised substantial amounts of equity during the 1980s crisis. We estimate a simultaneous equations model for the joint determination of price and capital flows.

We extend the risky-debt version of the arbitrage theory of insurance pricing to provide a new explanation of the relationships between loss shocks, capitalization, and prices in insurance markets. Like Winter and Gron, we assume that an insurer's ability to pay claims depends on its capital stock. However, we assume that insurers are subject to insolvency risk. Thus, insurance policies in our model are analogous to risky corporate debt.

We assume that the demand for insurance is positively related to financial quality and imperfectly price elastic in the short run. This extends the standard risky debt model, which assumes that demand for corporate debt is perfectly elastic at the fair market price of the debt.[7] Thus, in our model the relationship between the price of insurance and capital is derived from policyholder demand for quality, and an optimal capital structure (ratio of assets to liabilities, or "safety") is implied. The model predicts that safer firms command higher prices, and the cross-sectional relationship between safety and price should be positive.

We also generalize the standard risky-debt model of the firm by incorporating two classes of liabilities—"old" liabilities, for which the insurer can collect no additional premiums, and "new" liabilities (policies) that are priced in the current period. This generalization is motivated by events during the liability crisis. Loss shocks due to unanticipated exposures (such as asbestos and toxic torts) and changes in liability rules affected unsettled claims on prior policies as well as expected claims for future policies. Loss shocks from old policies provide a potential impediment to the flow of new capital into insurance markets. Any improvement in firm safety resulting from new capital benefits old policyholders by reducing their expected loss

[6] Long-tail lines, such as general liability, are those where loss payment cash flows span a period of several years following the coverage year. In short-tail lines, such as fire insurance, virtually all loss cash flows occur in the coverage year and the immediately following year.

[7] The Winter–Gron model could be interpreted as assuming extreme quality-sensitivity, such that demand becomes zero if the insolvency probability is finite. Our assumption is that quality-sensitivity is not this strong so that the equilibrium level of the insolvency probability is finite.

due to insolvency (i.e., reducing the insurer's insolvency put), although they pay no additional premiums. Thus, new capital inflows following a loss shock may cause a loss in market value of the insurer's equity[8] to the extent that the new capital infusions reduce the default risk of the firm.

With regard to the effect of loss shocks on prices, an important contribution of our analysis is to identify conditions that would give rise to a price increase following a shock using a model that is more general and more realistic than the capacity constraint model. In particular, our model does not require the assumption that insurers are free from default risk or the presence of severe capital market imperfections. And, unlike the capacity constraint theory, our theory also can explain a price decrease following a shock.

The existing theoretical analysis most similar to ours is Cagle and Harrington (1995). They modeled an insurance market with endogenous insolvency risk where price-taking firms maximize value by choosing output and capital. With inelastic demand, they show that price increases following a loss shock but not sufficiently to fully restore capital to preshock levels. If demand is price and quality elastic, the amount of the price increase following a shock is reduced. Our analysis differs from theirs in its use of an explicit option pricing framework, the introduction of two classes of liabilities, and the analysis of conditions that could lead to a price decrease rather than an increase following a shock. We also extend their analysis by providing empirical tests of the predictions of the model.

Our empirical analysis, using firm-level data for the period 1980–1988, supports the hypothesis that the relationship between safety and price is positive, consistent with the view that insurance policies are like risky debt. The results suggest that prices decreased in response to the loss shocks of the mid-1980s. Empirical support is also provided for the hypotheses that insurers have optimal capital structures and that capital flows are positively related to changes in prices.

Section 1 of our paper provides a statistical overview of liability insurance markets in the 1980s, including capital flows. Section 2 develops the risky debt model of the insurance firm and analyzes the effects of a loss shock. Section 3 tests several implications of the model, using firm-level data for the period 1980–1988, and Section 4 concludes.

1. TRENDS IN INSURANCE PRICES, LOSSES, AND CAPITAL FLOWS

Trends in the price of insurance, defined as the loading or ratio or premium to expected losses for a given policy, cannot be directly measured

[8] That is, the market value of equity following the infusion of new capital may be less than the market value prior to the capital contribution plus the amount of the contribution.

from insurance accounting statements, which report aggregate premiums (price times quantity). The loading charge measure of price used in our firm-specific empirical analysis is the ratio of written premiums, net of underwriting expense and policyholder dividends, to the discounted present value of accident year (AY) incurred losses (see below), also referred to as the inverse loss ratio.[9] Accident-year (AY) losses include claims paid and reserves for claims outstanding and anticipated for policies written during that year.[10] The loading charge ratio thus represents the markup of premiums over discounted incurred losses to account for risk bearing and claims settlement services provided by the insurer.

This section reports aggregate trends, to show the role of retroactive loss shocks and interest rates in explaining premium increases.[11] Figure 1 compares trends in GL premiums, undiscounted calendar year (CY) losses, and gross national product (GNP) from 1980 through 1988. CY losses rise roughly in proportion to GNP until 1984, when losses began to increase sharply, increasing by over 200% between 1983 and 1987. Premiums follow a more cyclical pattern, increasing more rapidly than losses between 1984 and 1986.

In order to distinguish retroactive vs prospective loss shocks, Fig. 2 decomposes CY incurred loss into two components: (1) claims paid and reserves set aside for policies written during that year (i.e., AY incurred losses), and (2) the loss reserve adjustment for prior accident years. The reserve adjustment doubled, from about 12% of total CY losses from 1980 to 1983 to 24% in 1985. AY losses accelerated most markedly in 1985 and 1986, which were also the years of the sharpest premium increases.[12] The

[9] This is similar to the economic loss ratio (the ratio of the present value of losses to premiums) used by Winter (1994). We use premiums written, which pertain to policies issued in a given year, rather than premiums earned, which reflect pricing during the current and preceding calendar years.

[10] Some other analyses of cycles use calendar-year (CY) incurred losses, which include the adjustment in loss reserves for prior years. Both CY and AY losses are gross of loss adjustment expenses, and are evaluated roughly 12–15 months from the start of the relevant year. For long-tail lines such as general liability and medical malpractice, the dollar-weighted mean pay-out period is around five years.

[11] The industry aggregate data used in this section are from A.M. Best Company, ‘‘Best's Aggregates and Averages,'' Oldwick, NJ, various years.

[12] To further analyze the effect of changing loss expectations, we decomposed accident year incurred losses into losses already paid and reserves for future claims, as of 15 months from the start of the accident year, for accident years 1980–1988. These data indicate that paid claims maintained a steady upward trend, while the growth in reserves first lags payments and then accelerates sharply in 1984–1986. Plots of paid and reserved claims evaluated at 27, 39, and 51 months show a similar pattern: the increase in AY losses incurred for 1984–1986 primarily reflects an increase in reserves for future claims rather than changes in claim payments. There are several theories of insurer incentives to manipulate loss reserves. Here we assume that reported loss reflects insurers' best estimates of expected loss. Under this assumption, these trends in reserves suggest a sharp upward revision of expectations regarding future payments on unsettled claims. Data for 1987–1988 show that these expectations were correct, as the paid claim index catches up with the reserve index.

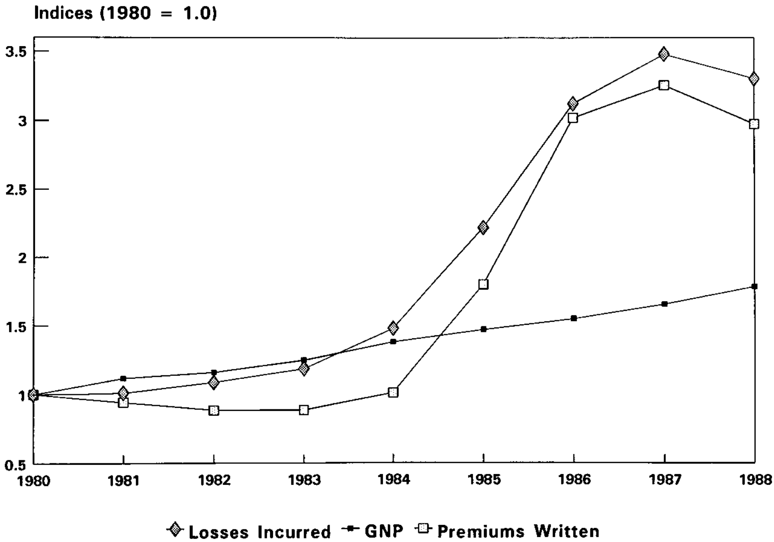

FIG. 1. General liability insurance net premiums written and losses incurred vs gross national product (GNP). Net premiums written = direct premiums written + reinsurance assumed − reinsurance ceded. Losses incurred = calendar year losses and loss adjustment expenses incurred.

experience for all long-tail lines combined follows a pattern that is similar but less extreme than that for GL. The loss reserve adjustment as a percentage of total CY losses increased from 4.8% in 1984 to 10.1% in 1985, and 8.7% in 1986; GL accounted for 40% of this total in 1984, 31% in 1985, and 32% in 1986. Thus, the shock was not confined to GL, although the GL experience was most extreme.

In order to show the trend in prices during the sample period, we plot the all lines and general liability implied loading ratios in Fig. 3.[13] Realized

[13] See Harrington (1988) for a previous application of this approach. The payout proportions are estimated using the industry-wide data from Schedule P, Part 3, from "Best's Aggregates and Averages," 1990 and 1991 editions. This gave 10 payout proportions for 1980 and 1981 and progressively fewer for later year. For the more recent years, where payments are not available for the longer lags, payout proportions were based on averages for earlier years. Claims were assumed to be fully paid after twelve years, so judgmental values were used for the payout proportions in years eleven and twelve. Testing based on alternative approaches revealed that the results were not very sensitive to the estimation approach for the payout proportions, in part because the payout tail was relatively stable over the sample period. Loss cash flows are estimated as AY losses multiplied by the estimated proportion of a given year's losses paid in each year of the payout tail. Losses are assumed to be paid at mid-year and are discounted using U.S. Treasury yield curves from Coleman *et al.* (1989). The discount factor is "risk-free" in that there is no adjustment for systematic risk of losses or for default risk.

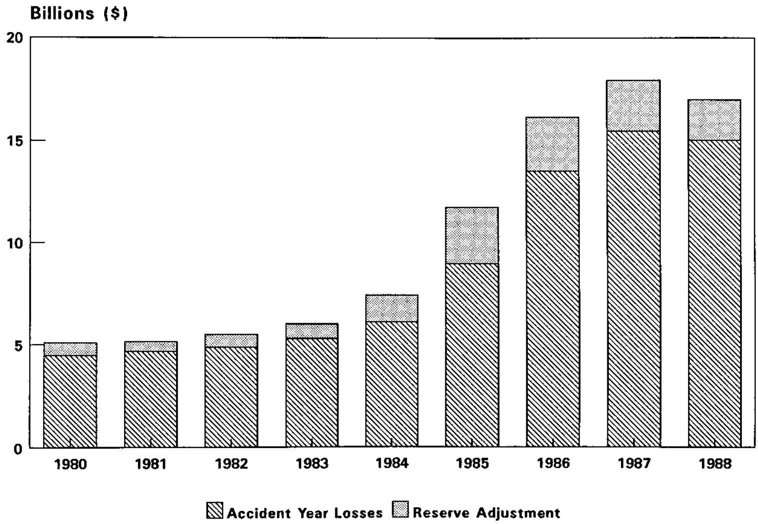

FIG. 2. Decomposition of calendar year losses incurred into accident year losses and reserve adjustment. Accident year losses = claim payments and reserves for policies written during a given year. Reserve adjustment = component of reported losses for a year reflecting adjustment in reserves for prior years' claims.

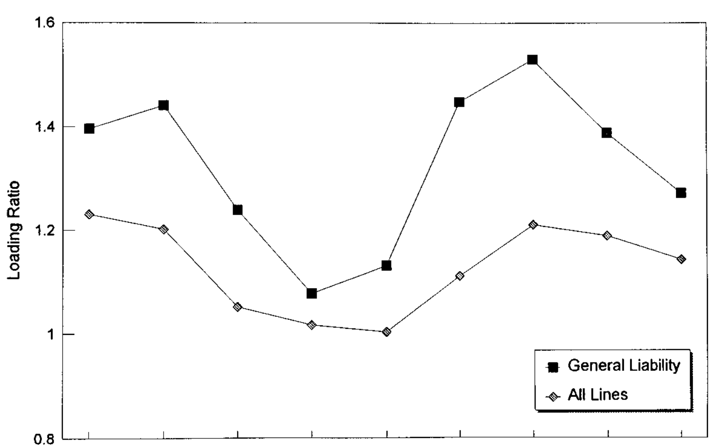

FIG. 3. Realized loading ratios: All lines and general liability. The loading ratio = (net premiums written − underwriting expense − policyholder dividends)/(present value of accident year losses incurred).

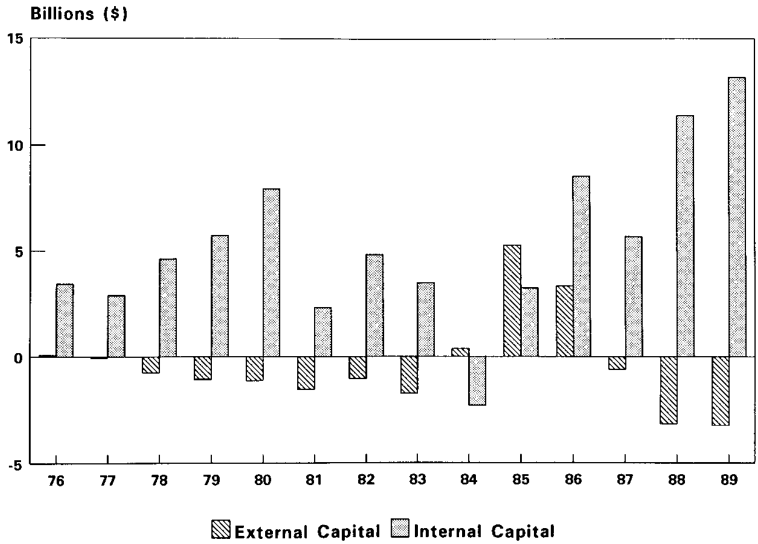

FIG. 4. External vs internal capital flows for stock property-liability insurers. External capital = new capital paid in − stockholder dividends. Internal capital = net income + unrealized capital gains.

loading ratios dropped significantly in 1982–1984 due to unanticipated losses and price cutting by insurers (Harrington and Danzon, 1994), rebounded in 1985–1986, and then dropped to levels similar to the early 1980s.[14] Because the loading ratio (markup of premiums over discounted incurred losses) was only slightly higher in 1985–1986 than during the more normal periods of the early and late 1980s, the high premiums of 1985–1986 do not appear to be due to abnormal markup ratios. Thus, consistent with Harrington (1988), we find that the atypical premium increases of 1984–1986 can be explained largely by the increase in expected losses during the period, the restoration of markups to more normal levels following the shock, and changes in interest rates.

The crisis years 1984–1986 were atypical not only in the premium increases but also in the inflows of new capital. Figure 4 shows the annual flows of internal capital (retained earnings plus unrealized capital gains)

[14] Some of the mid-1980s increase may reflect the corporate income tax increases of the Tax Reform Act of 1986, which affected long-tail lines more than short-tail lines. The Tax Reform Act of 1986 requires insurers to discount losses for purpose of computing the loss tax deduction. This has the effect of deferring part of the deduction. The deferral is greater for long-tail lines than for than short-tail lines because losses are paid out over a more extended period of time in the long-tail lines. See Cummins and Grace (1994).

and external capital (the difference between capital paid in and stockholder dividends) for stock insurers, which write more than 85% of GL premiums. Although stock insurers normally rely on retained earnings to generate equity capital, insurers responded to the 1984 loss shock by raising unprecedented amounts of external capital, with the largest flows in 1985–1986.[15]

In summary, the liability crisis of 1984–1986 was characterized by four phenomena: (1) sharp increases in expected losses on both old and new policies; (2) a large negative shock to capital in 1984; (3) large premium increases in 1985–1986 that are largely but perhaps not fully explained by shifts in the prospective loss distribution and interest rates; and (4) unusually large inflows of new external capital in 1985–1986. The following section develops a theoretical model of insurer choice of capital structure and pricing and analyzes response to a retroactive loss shock. The subsequent section provides empirical tests.

2. A MODEL OF THE INSURANCE FIRM

The standard arbitrage model of insurance pricing implies that the competitive price of insurance is equal to the present value of expected losses, expenses, and federal income taxes (Myers and Cohn, 1987). This model does not explicitly reflect demand sensitivity to price or quality, and treats capital as exogenous. In this section, we present an alternative model that endogenizes capital and permits us to analyze the effects of retroactive loss shocks on prices.

Assumptions

The model is based on the following assumptions:

1. Insurers' losses are imperfectly diversifiable, due to such factors as dependence among risks and nonstationarity of the loss distribution. The firm's ability to pay claims therefore depends on its assets relative to liabilities.

[15] Mutuals rely almost exclusively on retained earnings to generate capital, except for initial capitalization. Their retained earnings were nearly zero in 1984 but rebounded in 1985 and 1986 as premiums rose and investment results improved. The A.M. Best data on capital flows reported here pertain to firms that write the great bulk of property casualty premiums, but exclude some foreign reinsurers and some captives. Anecdotal evidence suggests that capital flows in the omitted sectors followed similar patterns, with large losses in 1984 and inflows in 1985 and 1986 with increased rate of formation of captives. The data on external capital do not distinguish between equity issued in securities markets (whether raised by the individual insurer or an affiliated firm or parent on behalf of a subsidiary) and transfer of retained earnings from a noninsurance parent or holding company.

2. Holding capital is costly because of such factors as corporate income taxation and regulatory constraints on insurers' permissible asset mix that may prevent insurers from optimally structuring their asset and liability portfolios (e.g., managing the duration of equity). Insurers therefore do not hold sufficient capital to eliminate all insolvency risk.[16]

3. We consider two possible assumptions about information in capital markets. Our baseline assumption is informational efficiency, which we take to mean that managers seek to maximize the value of equity and cannot dilute old or new equity. Any information asymmetries that exist are unaffected by a loss shock, and there is no differential cost of external capital relative to internal capital. We also consider an alternative assumption, that a loss shock differentially raises the cost of external equity relative to the cost of retained earnings (e.g., Myers and Majluf, 1984), and test for the implied effects in the empirical analysis.

4. Since insurers are subject to default risk, insurer liabilities are analogous to risky debt and can be modeled using the standard risky debt model of the firm (e.g., Merton, 1974), with modifications discussed below.

5. Policyholder demand for insurance depends on expectations about the firm's financial quality, which are based on observable leverage ratios and possibly other characteristics. Whereas the standard risky debt model assumes a demand for debt that is perfectly elastic at the risk-adjusted price, our baseline assumption is that the insurer's demand function, conditional on quality, is imperfectly price elastic, due to information asymmetries that raise the costs of switching insurers for policyholders. We also consider the implications of assuming perfectly price elastic demand.

The Model

The insurance firm faces a demand for policies that depends on price and default risk (financial quality):[17] $Q(p, b(x); Z)$, where Q = quantity of insurance sold (dollars of promised liability payments); p = the price per unit of insurance; x = the firm's asset-to-liability ratio; $b(x)$ is the insurer's insolvency put option per dollar of liabilities, i.e., the current value of the owners' option to default if liabilities exceed assets at the claim payment date ($\partial b/\partial x = b_x < 0$); and Z is a vector of parameters of the policyholders' loss distribution.[18] Time subscripts and other arguments in

[16] These costs are not explicitly incorporated in the model to avoid unnecessarily complicating the notation. Likewise, marketing and administrative costs are not explicitly incorporated.

[17] An insurer's financial quality matters because guaranty funds limit the maximum amount payable per claim (e.g., \$300,000), are subject to delays in payment and may exclude commercial coverages such as general liability insurance.

[18] Thus, we assume that bankruptcy costs are zero, i.e., the loss to policyholders in the event of insolvency of the insurer = $\text{Max}(0, L - A)$, where L = promised loss payment and A = insurer assets.

the $b(x)$ function are temporarily suppressed for simplicity. Demand is assumed to be a continuous, concave, decreasing function of the insolvency put option and price: $Q_b < 0$, $Q_{bb} < 0$, $Q_p < 0$, $Q_{pp} < 0$, where subscripts are used to denote partial derivatives. Our basic assumption is that demand is imperfectly price elastic because of switching costs arising from asymmetric information in the product market. An insurer acquires information about the policyholder's risk class by initial screening and by experience over time. Similarly, buyers acquire information about the insurer's service quality by experience or costly investigation. Both stocks of information are destroyed if the policyholder switches insurers (D'Arcy and Doherty, 1990).

Following the literature on risky corporate debt (e.g., Merton, 1974), insurance liabilities are modeled as a risky discount bond maturing one period from the present. The firm begins the period at time 1 with assets A_1, preexisting liabilities L_1, and equity $E_1 = A_1 - L_1$. No additional premium can be collected on the preexisting liabilities, which mature at the end of the period (time 2). At time 1, the firm can issue new policies subject to the demand structure $Q_2(p_2, b(x); Z)$, where p_2 = price at time 1 of policies issued at time 1 and maturing at time 2. Let $L_2 = Q_2$ denote the face value of new liabilities, with revenues of $P_2 = p_2 Q_2$. The insurer may also issue new equity $E_2 > 0$ or pay a dividend, $E_2 < 0$. The issue of new policies and equity changes insurer assets by $A_2 = P_2 + E_2$. After the issue of new policies and equity at time 1, $A = A_1 + A_2$ is the market value of assets, $L = L_1 + L_2$ is the market value (unadjusted for default risk) of liabilities,[19] and $b(x)$ is the value of the insolvency put on the portfolio of old and new policies (see below).

The new liabilities are assumed to mature at time 2 along with the preexisting liabilities.[20] If at time 2 the realized value of assets (A^R) exceeds the realized value of liabilities (L^R), the policyholders receive L^R and the owners receive $A^R - L^R$. If $A^R < L^R$, the firm defaults and the policyholders receive A^R. Thus, at the beginning of the period, the value of the policyholders' claim is $L[e^{-r\tau} - b(x)]$, the riskless present value of the promised liability payments less the value of the insolvency put option.[21] We assume that if the firm defaults, each class of policyholders receives a share of

[19] That is, L is the market value of liabilities if they were paid at time 1. The option model is based on diffusion processes for assets and liabilities, and L is the value of the liability state variable at time 1.

[20] An alternative approach would be to use a compound option model (Geske, 1977), with L_2 maturing after L_1. This adds complexity without changing the basic result. In fact, it would exacerbate the problem discussed in the text, that capital added to support L_2 is at risk of benefitting L_1.

[21] This uses the homogeneity property of the option model, to factor L and express the put in terms of the asset-to-liability ratio x (see Cummins, 1988).

assets proportional to the face value of claims, $w_i = L_i/(L_1 + L_2)$, $i = 1, 2$.[22]

Given these assumptions, the value of equity at time 1, following the policy and equity issue decisions, is given by a call option on the firm's assets.[23]

$$C_S(A_1+A_2, L_1+L_2, \tau) = A_1 + A_2 - (L_1+L_2)e^{-r\tau} + (L_1+L_2)b(x; r, \tau, \sigma^2), \tag{1}$$

where A_i = market value of assets, $i = 1, 2$,
L_i = market value of liabilities, $i = 1, 2$,
$b(x; r, \tau, \sigma^2) = b(x)$ = the value of a put option on random variable x with exercise price 1,
$x = (A_1 + A_2)/(L_1 + L_2) = A/L$,
r = risk-free discount rate net of expected liability inflation,[24]
$\tau = 1$ = time until expiration of the liabilities,
σ^2 = the firm's risk parameter,
$= \sigma_A^2 + w_1^2\sigma_{L_1}^2 + w_2^2\sigma_{L_2}^2 - 2w_1\sigma_{A1} - 2w_2\sigma_{A2} + 2w_1w_2\sigma_{12}$,
σ_j = the diffusion parameter for process j ($j = A$ = assets, $j = L_1$ = liability class 1, and $j = L_2$ = liability class 2),
$\sigma_{A1}, \sigma_{A2}, \sigma_{12}$ = the covariance parameters between the asset process and

[22] This liquidation rule is consistent with the way insurance bankruptcies are handled in practice (National Association of Insurance Commissioners, 1993). With this rule, the put option is on the portfolio rather than on each block separately. Because an option on a portfolio is worth less than the appropriately weighted sum of the individual options, there is generally a gain from diversification if the firm writes the new policies. Note that the price at time 1 is based on the market value of liabilities at that time rather than the unknown realization, L^R.

[23] This is the analogue of the standard put-call parity relationship. The reader is referred to Cummins (1988) for the derivation of the option model with stochastic assets and a single class of stochastic liabilities and to Appendix A for a sketch of the derivation of the two-class model. Assets and both classes of liabilities are assumed to be stochastic and follow correlated geometric Brownian motion processes. The risk-free rate of interest is assumed to be nonstochastic. The derivation requires the additional assumption that the sums $A_1 + A_2$ and $L_1 + L_2$ can be treated as if they can be approximated by lognormal random variables. (Although products and quotients of lognormal random variables are lognormal, sums of lognormals are not lognormal.) For a previous analysis using this assumption see Doherty and Garven (1986). The assumption simplifies the theoretical discussion by permitting us to take advantage of the homogeneity property of the Black–Scholes option formula. The assumption would present a problem only if there were some reason to believe that it would change the conclusions derived from the analysis, for example, because the distributions of the sums $A_1 + A_2$ and $L_1 + L_2$ are markedly different from the lognormal. In extensive simulations of sums of two lognormal variables we were rarely able to reject the hypothesis that the sum could be described by a lognormal distribution (Shapiro–Wilk test).

[24] If insurance liability inflation were the same as economy-wide inflation, r would equal the real riskless rate of interest. However, the model allows insurance liability inflation to differ from economy-wide inflation. See Appendix A for a sketch of the derivation.

liability process 1, between the asset process and liability process 2, and between liability processes 1 and 2, respectively, and

$w_i = L_i/(L_1 + L_2)$, $i = 1, 2$.

In this model there are four classes of claimants to the firm's assets, old and new policyholders and old and new stockholders, which may overlap. The issue of new policies and/or new equity may affect the relative value of the claims of the four groups. The assumption of informational efficiency in capital markets precludes an insurance or equity issue decision that dilutes the value of old equity or imposes a capital loss on new equity. Thus, the market value of firm equity following the policy/equity issue must be at least as large as the value of equity prior to the policy/equity issue, $C_1(A_1, L_1, \tau)$, plus the value of new equity contributed, E_2

$$C_S(A_1 + A_2, L_1 + L_2, \tau) \geq C_1(A_1, L_1, \tau) + E_2. \tag{2}$$

Subject to this constraint the firm at time 1 chooses price (p_2) and new equity (E_2) to maximize the value added to equity[25]

$$\underset{p_2, E_2}{\text{MAXIMIZE:}}\ C_S - E_2 - C_1 = Q[p_2, b(x)][p_2 - e^{-r\tau} + b(x)] + L_1[b(x) - b_1(x_1)], \tag{3}$$

where $b_1(x_1; r, \tau, \sigma_1^2) = b_1(x_1)$ = the value of the insolvency put option at time 1 prior to issuing new policies and equity,

$x_1 = A_1/L_1$, and

$\sigma_1^2 = \sigma_A^2 + \sigma_{L_1}^2 - 2\sigma_{A1}$.

The first order conditions for a maximum with respect to p_2 and E_2 are:

$$\{Q + Q_p[p_2 - e^{-r\tau} + b(x)]\} + \{Q_b[p_2 - e^{-r\tau} + b(x)] + Q + L_1\}b_x x_p = 0 \tag{4}$$

$$\{Q_b[p_2 - e^{-r\tau} + b(x)] + (Q + L_1)\}b_x x_e = 0, \tag{5}$$

where $Q_p, Q_b = \partial Q/\partial p_2$ and $\partial Q/\partial b$, respectively,

$x_p, x_e = \partial x/\partial p_2$ and $\partial x/\partial e$, respectively, and

$b_x = \partial b/\partial x$.

[25] Equation (3) is obtained by subtracting E_2 and $C_1(A_1, L_1, 1)$ from (1), using the put-call parity relation, $C_1 = A_1 - L_1[e^{-r\tau} - b_1(x_1)]$. This model assumes that buyers cannot "undo" the firm's capital structure choice through "homemade" increases or reductions in default risk.

The first expression in braces in Eq. (4) is the direct income effect of a price change. The second expression is the indirect "quality" effect of a price change on the insolvency put. From Eq. (5), this expression is zero if capital can be optimally adjusted using equity issue.

In Eq. (5) the first term $\{Q_2[p_2 - e^{-r\tau} + b(x)]\}b_x x_e$ is positive, reflecting the demand shift from increasing quality. The second expression, $(Q + L_1)b_x x_e$, is negative: adding equity reduces the value of the firm because the new equity improves safety not only for new policyholders but also for old policyholders who pay no additional premium. Thus, if demand is totally inelastic with respect to quality ($Q_b = 0$), equity is optimally zero. Given Q_b, a higher markup per new policyholder is required to raise new equity in firms with relatively large outstanding liabilities L_1. Rearranging yields

$$\frac{p_2 - e^{-r\tau} + b(x)}{p} = -\mathrm{E}_{Qp}^{-1} \tag{4'}$$

$$\frac{p_2 - e^{-r\tau} + (b(x)}{b(x)} = -\mathrm{E}_{Qb}^{-1}\left(1 + \frac{L_1}{Q}\right), \tag{5'}$$

where $E_{Qp} = (\partial Q/\partial p_2)(p_2/Q)$ = the elasticity of demand with respect to price, and
$E_{Qb} = (\partial Q/\partial b)(b/Q)$ = the elasticity of demand with respect to the put.

Equation (4′) is the standard formula for the profit maximizing markup of price over marginal cost. However, with limited liability, marginal cost is the expected value of liabilities net of the value of the put. Thus, assuming unbiased policyholder perceptions of quality, prices are lower with limited liability than with unlimited liability, and safer firms command higher prices.

Combining Eqs. (4′) and (5′) yields

$$\frac{Q_b}{Q_p} = \frac{Q + L_1}{Q} \quad \text{or} \quad \frac{\mathrm{E}_{Qb}}{\mathrm{E}_{Qp}} = \left(\frac{Q + L_1}{Q}\right)\frac{b(x)}{p_2}. \tag{6}$$

Equation (6) implies that the larger the ratio of old liabilities L_1 to new liabilities, the higher is the optimal insolvency put $b(x)$ (lower safety), relative to price (assuming $Q_{bb} < 0$, as above). In elasticity form, Eq. (6) is analogous to the Dorfman–Steiner condition for optimal expenditure on advertising because, like advertising expenditures, adding equity shifts demand.[26] Here, the optimal value of the limited liability put $(Q + L_1)b(x)$

[26] The Dorfman–Steiner condition states that the optimal ratio of advertising to sales is equal to the ratio of the demand elasticities with respect to advertising and price. See Scherer and Ross (1990), pp. 592–593.

relative to sales is equal to the ratio of the demand elasticities with respect to quality and price. Thus, in our model an optimal capital structure (assets/liabilities) is derived from the firm's demand structure, in contrast to other models where capitalization is exogenously determined by regulation or other factors (e.g., Myers and Cohn, 1987; Gron, 1994).

Firm Response to a Retroactive Loss Shock

Assume that prior to time 1, a retroactive shock to the loss distribution increases the face value of old liabilities L_1, thereby reducing the value of old equity and reducing safety below the optimal level. Before turning to comparative statics analysis of effects of this shock on the policy and equity issue decision at time 1, some straightforward implications of the model are worth stating.

If capital markets are informationally efficient and, contrary to our baseline assumption, the demand for insurance is perfectly price elastic conditional on quality, then the firm would not choose to improve its safety level (reduce the insolvency put) from its post-shock level $b_1(x_1)$. This result is immediate from Eq. (3). If insurance demand is perfectly price elastic, the price of insurance is simply the price of a risky discount bond $p_2 = [e^{-r\tau} - b(x)]$, implying that the first expression on the right-hand side of Eq. (3) is zero.[27] If safety were to improve following a shock so that $b(x) < b_1(x_1)$, capital would be penalized. An increase in safety would increase the value of old policyholders' claims (L_1). But since they pay no additional premiums, the market value of equity would decline by an equivalent amount. However, informational efficiency in capital markets implies that the firm would not choose to dilute either new or old equity (beyond the initial loss to old equity due to the loss shock), so that $b(x) \geq b_1(x_1)$ following a shock.[28]

However, if the demand for insurance is not perfectly price elastic, then it may be optimal to raise safety above its post-shock level. From (3), if the markup of price over marginal cost $[p_2 - e^{-r\tau} + b(x)]$ is positive, then $b(x)$ can be less than $b_1(x_1)$ without violating the constraint (2). The response of equity issue to a loss shock depends on the price and quality elasticities of demand and the cost of equity capital. Let p_2 and E_2 denote

[27] This assumes that insurance markets are competitive and that switching costs are sufficiently low, such that competition among insurers drives price down to the marginal cost, even if buyers are risk averse and hence willing to pay a markup over marginal cost.

[28] In theory the equity owners could increase the value of equity at the expense of old policyholders, by raising less equity or paying a dividend, thereby raising the put value ($b(x) > b_1(x_1)$). But such "go-for-broke" behavior is unlikely to be a value-maximizing strategy in a multiperiod context in which quality in period t affects demand in future periods, via reputation effects. Note that, even with informationally efficient capital markets and perfectly price elastic insurance demand, there is in principle no limit to the amount of capital or new insurance that can be issued as long as $b(x) \geq b_1(x_1)$.

the optimal price of new policies and quantity of new equity issued at time 2. Comparative statics analysis shows that the response of equity issue to an increase in L_1 ($\partial E_2/\partial L_1$) is ambiguous.[29] The increase in L_1 raises the ratio of old to new liabilities and, since old policies contribute no additional revenue, this raises the marginal cost of adding equity (see Eq. (5)). Nevertheless, it may be optimal to issue some new equity ($\partial E_2/\partial L_1 > 0$) if marginal revenue per new policy also increases, either because the quality elasticity increases (in absolute value) in response to the loss shock or because the markup per policy increases. Because the baseline assumptions of the model imply that demand becomes unambiguously more quality elastic following a shock, this effect alone or in combination with an increase in the markup must be sufficiently large in order for $\partial E_2/\partial L_1 > 0$.[30]

We now consider the response of price to a loss shock. Assuming informational efficiency in capital markets (i.e., that the cost of external equity relative to internal equity is unaffected by a loss shock), the optimal price change, $\partial p_2/\partial L_1$ is:

$$\frac{\partial p_2}{\partial L_1} = \frac{Q_p + Q_b + Q_{pb}[p_2 - e^{-r\tau} + b(x)]}{D}, \tag{7}$$

where $Q_{pb} = \partial^2 Q/\partial p_2 \partial b$, and D is the determinant of the second derivatives of the objective function.

Thus, given informational efficiency in capital markets, a sufficient condition for a negative relation between price and a loss shock ($\partial p_2/\partial L_1 < 0$) is that Q_{pb} is negative. Q_{pb} is a sufficient condition (under the baseline assumption that Q_b and Q_p are < 0) for demand to become more price elastic in response to an increase in b (i.e., a decline in the firm's safety level).[31] Among other things, a higher probability of insolvency reduces the expected value of the private information the insurer possesses regarding the buyer and of the buyer's private information on firm service quality, since this information would be lost in the event the firm becomes insolvent. As a result, an increase in b may make it more attractive for the buyer to reenter the market and find insurance elsewhere. Thus, an inverse relation-

[29] The comparative statics are presented in an appendix available from authors.

[30] The quality elasticity is $E_{Qb} = (\partial Q/\partial b)(b/Q) = Q_b b/Q$, where the subscript on b has been suppressed to simplify the notation. Differentiating, one obtains $\partial E_{Qb}/\partial b = b[QQ_{bb} - Q_b^2]/Q^2$. Under the baseline assumptions of the model, $Q_{bb} < 0$, and thus $\partial E_{Qb}/\partial b < 0$ unambiguously.

[31] The price elasticity is $E_{Qp} = (\partial Q/\partial p_2)(p_2/Q) = Q_p p/Q$, where the subscript on p has been suppressed to simplify the notation. Differentiating, one obtains $\partial E_{Qp}/\partial b = p[QQ_{pb} - Q_p Q_b]/Q^2$. Under the assumptions $Q_p < 0$ and $Q_b < 0$, $\partial E_{Qp}/\partial b < 0$ unambiguously if $Q_{pb} < 0$.

ship between price and a loss shock is a plausible (but not necessary) outcome in our model.

A positive relationship between price and a loss shock is also possible in our model, if either Q_{pb} is positive (and sufficiently large) or if a loss shock severely raises the cost of external capital, for example, due to increased risk that new equity will increase the value of old liabilities sufficiently to cause a capital loss. Under the extreme assumption that no adjustment of equity is possible, comparative statics analysis shows that the expression for $\partial p_2 / \partial L_1$ involves an additional term from the quality effect (see Eq. (4)), which is positive, reflecting the fact that higher prices boost safety and reduce the insolvency put. Thus, under the appropriate conditions, the model yields predictions that are consistent with those of the capacity constraint model. However, our model is more general in the sense that it can also explain a price decrease in response to a shock.

Factors not explicitly included in the model also could increase the likelihood of a price increase. For example, a shift in the underlying loss distribution in response to a shock could affect policyholders' demand for insurance. If new policyholders and old policyholders are the same individuals, by paying higher prices for new policies they contribute to the increased safety of their own old policies. While safety increases by less than the contribution (because strengthening equity also benefits equity holders and other policyholders), the fact that policyholders' existing claims are strengthened may lessen their resistance to higher prices. A positive change in price also might be more likely in response to loss shocks that are correlated across firms because an industry-wide shock is expected to make firm-specific demand less elastic.[32]

If we drop the assumption of informationally efficient capital markets and assume instead that equity markets fear that managers have underreported the true increase in old liabilities (L_1), the model implies that inflows of capital may be associated with price increases, which strengthen expectations of retained earnings and signal capital markets that equity will not be penalized because of increases the the value of old liabilities. One implication of the assumption that shocks to the loss distribution may affect both insurance demand and the cost of capital is that prices may

[32] With informational efficiency in insurance markets—including insurer information about policyholder risk (no adverse selection) and policyholder information about insurer quality—the threat of entry by new firms would constrain prices charged for new policies to their risk-adjusted value. But if new entrants face higher information costs than old firms, and policyholders face information costs in switching to new firms, then entry may not deter temporary price increases by old firms. In fact, much of the entry that occurred in the commercial liability insurance market during the 1980s was in the form of policyholder-initiated insurance firms, including captives, mutuals and risk retention groups. This is consistent with the hypothesis that asymmetric information acts as a barrier to entry.

respond differently to retroactive loss shocks than to capital shocks from other sources. To test this hypothesis we decompose the change in capital into the components due to retroactive loss shocks, new internal capital, and new external capital.

3. EMPIRICAL TESTS

The theoretical model yields some formal predictions and suggests other testable hypotheses. An important prediction of the model is that the price of insurance is positively related to capitalization, due to the risky debt nature of insurance policies. A related implication is that insurers have optimal capital structures because of buyer preference for financial quality. Thus, capital flows are predicted to adjust to attain optimal leverage ratios.

The predictions of the model under alternative assumptions about the effects of shocks on demand elasticities and other important conditions are summarized in Table I. Under the baseline assumptions of the model, we predict an unambiguous inverse relationship between price and a loss shock provided that $Q_{pb} < 0$, a sufficient condition for an increase in price elasticity of demand following a shock. However, if Q_{pb} is positive and sufficiently large, prices may respond positively to a loss shock. Under these conditions the price elasticity of demand is likely to decrease following a shock. If we relax the baseline assumptions and assume instead that no new capital can be raised following a shock and/or that the demand curve shifts to the right following a shock, then prices are more likely to increase in response to a shock. Table I also shows that capital will be raised following a shock under the baseline assumptions of the model if the quality elasticity of demand and/or price markups increase sufficiently following a shock.[33]

Data and Variable Construction

Our sample consisted initially of all firms that wrote at least 0.5% of the general liability (GL) market in any year between 1976 and 1987, a total of 50 firms (affiliated firms are consolidated into groups).[34] Three firms were eliminated from the sample because they either became insolvent or stopped issuing new policies during the sample period. These firms were excluded because the theory applies to ongoing firms and because our generalized least-squares (GLS) estimation technique requires complete

[33] Recall that an increase in the quality elasticity of demand is an unambiguous prediction under the baseline assumptions of the model. The effect of a shock on the markup is ambiguous in our model.

[34] Captives were excluded in selecting the sample because their principal customers are their parent firms, which have different information sets than the customers of commercial insurers.

TABLE I

PREDICTED RESPONSE OF PRICE AND NEW EQUITY TO A RETROACTIVE LOSS SHOCK

Baseline assumptions:

1. Informational efficiency in capital markets (no dilution of equity, internal and external costs of capital are identical).
2. Demand inversely related to price ($\partial Q/\partial p = Q_p < 0$) but not infinitely price elastic.[a]
3. Demand inversely related to expected insolvency costs ($\partial Q/\partial b = Q_b < 0$) but not infinitely elastic.

Response of price to loss shock ($\partial p_2/\partial L_1$):[b]

$\partial p_2/\partial L_1 < 0$ requires:

1. Baseline assumptions, and either
2. $\partial^2 Q/\partial p \partial b = Q_{pb} \leq 0$ (a sufficient but not necessary condition for demand to become more price elastic following a shock), or
3. $Q_{pb} > 0$ but $Q_p + Q_b < -Q_{pb}(p_2 - e^{-r\tau} + b(x))$. ($Q_{pb} > 0$ is a necessary but not sufficient condition for demand to become less price elastic following a shock).

$\partial p_2/\partial L_1 > 0$ requires:

Case A: 1. Baseline assumptions, and
2. $Q_{pb} > 0$ and sufficiently large so that $Q_p + Q_b > -Q_{pb}(p_2 - e^{-r\tau} + b(x))$. The larger is Q_{pb}, the more likely it is that demand is less price elastic following a shock.

Case B: Some combination of the following factors is present and of sufficient magnitude to lead to a price increase:
1. The cost of capital rises following a shock and the quality effect is strong enough to permit a price increase.
2. The demand for insurance increases following a shock.
3. Price elasticity of demand is low because of significant overlap between the old and new policyholders.
4. Price elasticity of demand is low because shocks are correlated across the industry rather than being firm-specific.

Response of new capital to loss shock ($\partial E_2/\partial L_1$):[c]

$\partial E_2/\partial L_1 > 0$:

Case A: 1. Baseline assumptions with some combination of the following factors present in sufficient strength such that $\partial E_2/\partial L_1 > 0$:
2. The markup per policy is sufficiently high, and/or
3. Demand is sufficiently quality elastic, and/or
4. Old liabilities are sufficiently small relative to demand for new policies.

Case B: 1. Demand is perfectly price elastic.
2. Equity is introduced to meet increased demand without lowering the insolvency put $b(x)$.

$\partial E_2/\partial L_1 \leq 0$:

Baseline assumptions but factors (2), (3), and (4) in Case A above are not sufficient for a positive partial derivative.

---

[a] $Q(p, b(x))$ = insurance demand function (demand is expressed in terms of dollars of promised liability payments), where p = price per unit of insurance, x = the firm's asset-to-liability ratio, $b(x)$ = the insurer's insolvency put option per dollar of liabilities, i.e., the value of the owners' option to default if liabilities exceed assets at the claim payment date with $\partial b/\partial x = b_x < 0$.

[b] In this expression, p_2 = price of insurance issued at time 1, and L_1 = preexisting liabilities of the insurer at time 1.

[c] $E_2 > 0$ new equity capital raised by the insurer at time 1; $E_2 < 0$ dividend to shareholders paid at the beginning of period 2.

panel data. One firm was eliminated because it was acquired by another sample firm in 1986, and one was eliminated based on statistical analysis of outliers. The 45 firms in the final sample accounted for 75% of total industry premium volume and 85% of general liability premium volume during the sample period. The sample firms are shown in Appendix B.

The data are from the annual statements filed with state insurance departments, as reported by the A.M. Best Company. These data are more detailed and comprehensive than the data used in most existing studies of underwriting cycles, which rely on annual aggregate data for the entire industry rather than firm-specific data.

Even though the "crisis" was most extreme in the line of general liability (GL) which includes product liability, we base our analysis of price on the company-wide, weighted average price across all lines. Although it would be useful to evaluate the GL-specific prices during the sample period, the data available to us were not adequate to support this type of analysis. Use of the all-lines price could be viewed as theoretically appropriate, because the insolvency put value is company-wide, rather than line-specific. The company's equity backs all policies, and no specific category of policies gets priority in bankruptcy. Because price may vary by line due to differences in risk and anticipated claims inflation and because GL accounted for a substantial share of the 1984–1986 price increase, we include the proportion of premiums in GL as a control variable.

Price. Our measure of price is the ratio of premiums paid to the discounted present value of expected losses for the total business written by the company. The numerator is premiums written, which reflects pricing on policies issued during a given year, measured net of underwriting expense and dividends paid to policyholders. Incurred losses are net of changes in prior years' reserves, i.e., on an accident year (AY) basis, and are discounted using the U.S. Treasury yield curves published in Coleman *et al.* (1989).[35] Losses include loss adjustment expenses since legal expense coverage is part of the protection of a liability insurance policy.

Financial quality, $b(x)$. Since policyholders cannot observe $b_t(x)$, we assume that they form an expectation of quality based on indicators of leverage in period $t-1$. To test whether prices are affected by the source of capital, either because prices respond differently to a shock to prior liabilities than to other sources of capital change or because external capital is more costly than internal capital, we decompose capital at the end of year $t-1$ ($K_{t-1} = A_{t-1} - L_{t-1}$) into its components as

$$K_{t-1} = K_{t-2} + \text{AYINC}_{t-1} + \text{NEWCAP}_{t-1} - \text{LRAL}_{t-1} \qquad (8)$$

[35] We use the all-industry payout tails for Schedule P lines and the Schedule O payout tail for all other lines. The years used to estimate the payout tails are discussed in footnote 13.

where K_{t-1} = $\text{ASSETS}_{t-1} - \text{LIAB}_{t-1}$ = assets minus liabilities at end of year $t-1$; AYINC_{t-1} = premiums + net investment income + unrealized capital gains − accident year incurred losses; NEWCAP_{t-1} = capital paid in from external sources net of stockholder dividends paid out in year $t-1$; and LRAL_{t-1} = losses paid + change in reserves in AY_{t-1} for accident years $< t-1$. All $t-1$ variables are normalized to LIAB_{t-2}.[36] The loss shock LRAL_{t-1} measures the effect on capital in period $t-1$ of changes in incurred losses for prior years. As another indicator of financial quality, we also define a dummy variable (BESTSA) that takes a value of 1 if the A.M. Best Company rating for a firm is A or A+.

Capacity constraint. The Winter/Gron capacity constraint hypothesis predicts that price is inversely related to capacity, defined as the level of surplus relative to the demand for insurance. To test whether the relative level of surplus adds additional explanatory power, after controlling for the firm-specific variables suggested by our model, we include Winter's (1994) measure of capacity—real capital in $t-1$ relative to the average value over the preceding five years, K_{t-1}/K_5, where K_5 = average real capital over the period $t-2$ to $t-6$. In order to distinguish the industry-wide effect from the firm-specific effect, we decompose the firm-specific capacity variable K_{t-1}/K_5 into two orthogonal components: INDUSK_{t-1} is the predicted value from a regression of K_{t-1}/K_5 for the firm on the analogous value for the industry as a whole; FIRMK_{t-1} is the residual from this regression. The orthogonalization removes a source of multicollinearity and provides a cleaner test of the capacity constraint hypothesis, which emphasizes industry-wide capital shortages (INDUSK_{t-1} varies over time but not cross-sectionally).

New capital. Our measure of new equity NEWCAP is capital and surplus paid in from external sources. This includes new equity issues and transfers from noninsurance parent corporations.[37]

Other variables. We include as a control variable the proportion of the firm's premium volume in general liability (GLSHARE), since the price (inverse loss ratio) is expected to differ across lines, depending on line-specific loss adjustment and underwriting expenses, risk and taxes. The proportion of premiums in all other long-tailed lines was also initially included, but was not significant and was dropped from the equations reported here.

[36] We normalized to LIAB_{t-2} rather than LIAB_{t-1} since LIAB_{t-1}, the measure of retroactive liabilities, is a key explanatory variable.

[37] Transfers from a parent are appropriately treated as new equity in our model, assuming that the market would penalize the stock of the parent for transfers that reduce equity value. Our measure of new equity nets out equity flows among insurers within the same insurance group.

A dummy variable distinguishes stock insurers from mutuals and reciprocals. Stocks may differ from mutuals for reasons such as risk diversification, access to capital, agency costs, and control of moral hazard and adverse selection.[38] We do not attempt to distinguish among these factors. Another dummy variable distinguishes publicly traded stock insurers from the other firms in the sample to control for any differences arising from a firm's trading status. Table II reports means and standard deviations of all variables.

Specification

The theory implies a two equation model in price level p_t and additions to equity from external sources E_t:

$$p_t = a_0 + a_1 K_{t-1} + \mathbf{a}_2 \mathbf{X}_t + \mathbf{a}_3 \mathbf{X}_{t-1} + u_t \tag{9}$$

$$E_t = b_0 + b_1(p_t - p_{t-1}) + \mathbf{b}_2 \mathbf{Y}_t + \mathbf{b}_3 \mathbf{Y}_{t-1} + \nu_t, \tag{10}$$

where $\mathbf{X}$ and $\mathbf{Y}$ are vectors of predetermined variables, a_i and b_i ($i = 0, 1$) are scalar coefficients, $\mathbf{a}_i$ and $\mathbf{b}_i$ ($i = 2$ and 3) are coefficient vectors, and u_t and ν_t are error terms. Some of the coefficients in $\mathbf{a}_i$ and $\mathbf{b}_i$, $i = 2, 3$, are constrained to zero based on the theory. We hypothesize that policyholder expectations of quality in period t depend primarily on actual values in period $t - 1$. Thus, most of the elements of $\mathbf{a}_2$ are constrained to zero. On the other hand, inflows of new equity are more likely to be affected by contemporaneous values of the explanatory variables, so that most elements in $\mathbf{b}_3$ are constrained in zero. The rationale for this approach is that detailed data on the financial condition of insurers is made available to the public and regulators annually, at year-end, whereas the insurers themselves can base equity issuance decisions on more recent, internal data. The use of different lag configurations in the two equations leads to identification of the model.

The equations reported here are based on a pooled, cross-section time-series analysis of 45 firms for the period 1980–1988. Testing revealed that the regression error terms in Eqs. (9) and (10) were characterized by serial correlation and heteroskedasticity. With lagged endogenous variables and autocorrelated errors, OLS estimates will be inconsistent. To estimate (9) and (10), we therefore use a generalized least-squares (GLS) model for panel data developed in Kmenta (1986, pp. 616–625). The model is adapted to control for the problem of lagged endogenous variables using the autoregressive two-stage least squares approach proposed by Kmenta (1986, pp. 708–710). Both the autocorrelation and heteroskedasticity adjustments are

[38] See, for example, Mayers and Smith (1988) and Danzon (1984).

TABLE II

Summary Statistics: Variables Included in the Regressions: Means and Standard Deviations, 1980–1988

Variable	Mean	Standard deviation
Price	1.1284	0.1537
New equity capital	\$38,251	\$102,355
Change in price	−0.0006	0.1399
Assets	\$4,637,460	\$4,930,660
Liabilities	\$3,587,984	\$3,569,421
Capital	\$1,049,476	\$1,608,452
Accident year income	\$170,464	\$232,731
Loss Reserve Adjustment	\$54,380	\$123,433
Premiums written	\$2,102,531	\$2,506,955
General liability premiums written/total premiums written	0.115	0.094
Stock: Dummy variable = 1 if the insurer is a stock company, 0 otherwise	0.800	0.400
Traded: Dummy variable = 1 if the insurer is a publicly traded stock company, 0 otherwise	0.467	0.500
Best's A or A+ rating = BESTSA Dummy variable = 1 if A or A+, 0 otherwise	0.881	0.324
FIRMK($t-1$)	0	0.160
INDUSK($t-1$)	1.096	0.098
Sample size	405	
Percent of industry premiums written by sample firms (1980–1988)	74.5%	
Percent of general liability premiums written by sample firms (1980–1988)	84.7%	

Note. Dollar valued variables are in thousands. All insurer financial statement data are from regulatory annual statements as reported on the A.M. Best Company data tapes. Lags are indicated by the notation $t - i$, t = current period. PRICE = (net premiums written—underwriting expenses—dividends to policyholders)/(present value of accident year losses incurred). Net premiums written = direct premiums written + reinsurance premiums assumed − reinsurance premiums ceded. Losses are discounted using U.S. Treasury yield curves from Coleman *et al.* (1989) and industry-wide loss payout proportions from A.M. Best Company, "Best's Aggregates and Averages," 1990, 1991. Accident year income is defined as calendar year net income plus the loss reserve adjustment, where the loss reserve adjustment is the component of a given year's calendar year losses incurred attributable to adjustments in reserves for all prior years. BESTA = dummy variable = 1 if the insurer is rated A or A+ by the A.M. Best Company, 0 otherwise. Real equity capital is obtained by deflating equity capital by the CPI, base year 1982. To obtain FIRMK($t-1$) and INDUSK($t-1$), we define $K(t-1)/K(5)$, where $K(t-1)$ = real capital in year $t-1$ and $K(5)$ = average real capital over the period $t-2$ to $t-6$. INDUSK($t-1$) = the predicted value from a regression of $K(t-1)/K(5)$ for the firms in the sample on the analogous value for the industry as a whole: FIRMK($t-1$) = the residual from this regression.

allowed to vary by firm.[39] OLS results are also reported for comparison; and standard two-stage least-squares (TSLS) estimates of the capital equation are reported as well, because of the presence of p_t in Eq. (10).[40]

Empirical Results

The price level equation. Table III reports estimates of three specifications of the price level equation, with different measures of financial quality. The first equation includes only a single composite measure of lagged financial quality (equity capital/liabilities). This variable is significantly positively related to price, consistent with the risky debt hypothesis that safer firms command higher prices.

The second equation decomposes lagged capital into the two-period lagged capital stock (K_{t-2}) and a vector of three measures of capital change: new internal capital (AYINC_{t-1}), new external capital (NEWCAP_{t-1}), and the loss shock from old liabilities (LRAL_{t-1}). Under the null hypothesis that policyholders care only about overall leverage but are indifferent to the source of capital, the coefficients on these four components should be equal (in absolute value). Equation (3) adds the variables, INDUSK_{t-1} and FIRMK_{t-1}. The capacity constraint theory predicts a negative coefficient on INDUSK_{t-1}.

The results in Eqs. (2) and (3) support the risky debt hypothesis: the internal and external capital variables have statistically significant positive coefficients and the loss reserve adjustment (LRAL_{t-1}) has a negative coefficient that is statistically significant in the GLS equations, providing evidence of an inverse relationship between prices and loss shocks during

[39] We use a first-order autoregressive specification for the error term in the GLS regressions: $u_{it} = \rho_i u_{i,t-1} + \varepsilon_{it}$, where u_{it} = the error term for company i in year t; $E(u_{it}^2) = \sigma_i^2$ (heteroskedasticity); $\varepsilon_{it} \sim N((0, \sigma_{\varepsilon i}^2)$, $E(\varepsilon_{it}\varepsilon_{i,t-i}) = 0$, $i > 0$; and $E(u_{it}u_{jt}) = 0$, $i \neq j$. The equations are iterated until the coefficient estimates and other parameters converge. Such iterated Aitken estimates have been shown to converge to maximum likelihood estimates if estimation is begun with initial consistent estimates.

[40] We also tested an error component (random effects) model, with the error term specified as $u_{it} = \xi_i + \nu_t + \varepsilon_{it}$, for company i in year t, with the usual assumptions about the error components (see Kmenta, 1986, p. 625 for the details). Unlike the first-order autoregressive model, which implies that the covariance of the errors, $E(u_{it}u_{is})$, $t \neq s$, declines geometrically as the time distance between disturbances increases, the error components model implies that $E(u_{it}u_{is}) = \sigma_\xi^2$, $t \neq s$; i.e., the covariance of the disturbances over time remains unchanged regardless of the time distance between observations. As mentioned above, the autoregressive specification is indicated as more appropriate for our sample. We also tested a fixed effects model, which is estimated using OLS with time and company dummy variables added to the specification. The fixed effects specification does not adjust for first-order serial correlation but rather allows the constant term in the regression to vary by year. The regressions results based on the random and fixed effects models are similar to the OLS regressions reported in Tables III and IV.

TABLE III
PRICE LEVEL EQUATION: 1980–1988

Variable	Equation 1: GLS	Equation 1: OLS	Equation 2: GLS	Equation 2: OLS	Equation 3: GLS	Equation 3: OLS
Constant	0.94 41.514	0.860 27.498	0.977 43.616	0.917 29.844	0.895 20.798	0.834 12.835
Stock: Dummy variable = 1 if the insurer is a stock company, 0 otherwise	0.047 2.912	0.069 3.572	0.038 2.477	0.047 2.559	0.034 2.286	0.048 2.620
BESTSA($t-1$) = Best's A or A+ rating($t-1$) Dummy Variable = 1 if A or A+, 0 otherwise	0.006 0.331	−0.016 −0.717	0.009 0.501	−0.008 −0.376	0.013 0.744	−0.007 −0.314
Traded: Dummy variable = 1 if the insurer is a publicly traded stock company, 0 otherwise	0.044 2.708	0.023 1.642	0.040 2.823	0.018 1.337	0.035 2.490	0.017 1.310
GLSHARE($t-1$) = other liability premiums written($t-1$)/total prem written($t-1$)	0.541 6.749	0.547 7.456	0.424 5.150	0.347 4.866	0.413 4.646	0.331 4.434
$K(t-1)$ = total equity capital($t-1$)/liabilities($t-2$)	0.293 5.171	0.476 8.471				
$K(t-2)$ = total equity capital($t-2$)/liabilities($t-3$)			0.121 2.358	0.198 3.413	0.104 1.988	0.209 3.497
AYINC($t-1$) = accident year income($t-1$)/liabilities($t-2$)			0.654 7.489	1.233 8.504	0.551 4.877	1.103 5.981
NEWCAP($t-1$) = [capital paid in($t-1$) − div($t-1$)]/liabilities($t-2$)			0.595 5.260	0.785 4.677	0.640 5.088	0.752 3.854
LRAL($t-1$) = loss reserve adjustment($t-1$)/liabilities($t-2$)			−0.406 −2.783	−0.297 −1.479	−0.462 −2.757	−0.214 −0.929
FIRMK($t-1$)					−0.038 −1.638	0.004 0.155
INDUSK($t-1$)					0.072 2.059	0.071 1.421
Adjusted R^2	0.249	0.247	0.309	0.349	0.302	0.349
Sample size	405	405	405	405	405	405

Note. Dependent variable = PRICE = (net premiums written − underwriting expenses − dividends to policyholders)/(present value of accident year losses incurred). For each variable, the top number = the coefficient and the lower number = t ratio. R^2 is redefined in the GLS models to use weighted sums of squares (Kmenta, 1986, equation (2.12)). GLS = generalized least squares; OLS = ordinary least squares; lags are indicated by the notation $(t-i)$, t = current period. The dependent variable is measured at period t. Insurer financial statement data are from regulatory annual statements as reported on the A.M. Best Company data tapes. In the PRICE variable, net premiums written = direct premiums written + reinsurance the sample period. This is consistent with the baseline assumptions of the model with Q_{pb} either ≤ 0 (an increase in price elasticity) or > 0 but not large enough to offset the negative terms in the numerator of Eq. (7) (see also Table I).

Statistical tests reject the hypothesis that the coefficients on the three components of the change in capital are equal.[41] In particular, the coefficient on the loss reserve adjustment LRAL_{t-1} is smaller in absolute value than the coefficients of AYINC_{t-1} and NEWCAP_{t-1}, providing evidence that a retroactive loss shock has a less negative effect on price than capital shocks from other sources.[42] Thus, there is also some evidence that a shift in

[41] In the OLS versions of Eqs. (2) and (3), F tests lead to rejection at the 1% confidence level of the null hypothesis that the coefficients on the three components of the change in capital are equal. In the GLS version of Eq. (2), the null hypothesis of coefficient equality is rejected at the 10% level based on a χ^2 test. The hypothesis is rejected at the 1% level in the GLS version of Eq. (3). The likelihood ratio test (a χ^2 test) is appropriate for testing coefficient restrictions in the GLS regressions because we use the iterated version of the GLS estimator, which converges to maximum likelihood estimates (Kmenta, 1986, p. 620). The test results are as follows: Equation (2), OLS: computed F statistic = 15.5 vs critical F(0.01, 2, 396) = 4.61, where $F(a, b, c)$ = critical F statistic at confidence level a with b and c degrees of freedom. Equation (2), GLS: computed χ^2 statistic = 5.5 vs critical $\chi^2(0.05, 2) = 5.99$ or $\chi^2(0.10, 2) = 4.61$, where $\chi^2(a, b)$ = chi-square critical value at confidence level a with b degrees of freedom. Equation (3), OLS: computed F statistic = 13.1 vs critical F(0.01, 2, 394) = 4.61. Equation (3), GLS: computed χ^2 statistic = 9.3 vs critical $\chi^2(0.01, 2) = 9.21$.

[42] Thus, the effect on price of capital depletions resulting from loss shocks is mitigated somewhat relative to capital changes from other sources. For example, assume that LRAL_{t-1} is 20% of liabilities at the end of the prior year (a 20% loss shock) and that half of the dollar value of the resulting capital loss is offset by retained earnings (AYINC_{t-1}) and half by new capital (NEWCAP_{t-1}). Because the coefficient of LRAL_{t-1} is smaller in absolute value than the coefficients of AYINC_{t-1} and NEWCAP_{t-1}, price would increase even though the ratio of capital to liabilities would be lower due to the shock. The implied increase in price at the sample mean price ratio based on Eq. (2) of Table II is 3.9% based on the GLS version of the equation and 12.6% based on the OLS version (the results based on equation (3) are similar). This is a hypothetical example; we do not intend to imply that any restoration of capital would necessarily be in the amounts of proportions used in the example.

premiums assumed − reinsurance premiums ceded; and losses are discounted using U.S. Treasury yield curves from Coleman, Fisher, and Ibbotson (1989) and industry-wide loss payout proportions from A.M. Best Company, "Best's Aggregates and Averages," 1990, 1991. BESTSA = 1 if the insurer is rated A or A+ by the A.M. Best Co., 0 otherwise. Accident year income = calendar year net income plus the loss reserve adjustment, where the loss reserve adjustment is the component of calendar year losses incurred attributable to adjustments in reserves for all prior years. In the variable NEWCAP($t-1$), div($t-1$) = dividends to stockholders in year $t-1$. To obtain FIRMK($t-1$) and INDUSK($t-1$), we define $K(t-1)/K(5)$, where $K(t-1)$ = real equity capital in year $t-1$ and $K(5)$ = average real equity capital over the period $t-2$ to $t-6$. Real capital is obtained by deflating equity capital by the CPI, with 1982 as the base year. INDUSK($t-1$) = the predicted value from a regression of $K(t-1)/K(5)$ for the firms in the sample on the analogous value for the industry as a whole; FIRMK($t-1$) = the residual from this regression.

demand, capital constraints, or an increase in the cost of capital may have mitigated the negative impact of loss shocks on price during the sample period. The coefficient of NEWCAP_{t-1} is smaller than the coefficient of AYINC_{t-1}, except in the GLS version of Eq. (3), contrary to the hypothesis that external capital is more costly than internal capital.

Equation (3) includes the two measures of capacity, INDUSK_{t-1} and FIRMK_{t-1}. These variables are not implied by our model but are added to test the Winter–Gron hypothesis. Comparing Eqs. (2) and (3), adding the two measures of capacity does not raise the adjusted R^2. Both variables are insignificant in the OLS version of the regression. In the GLS version, the systematic component INDUSK_{t-1} is significantly positively related to price, contrary to the capacity constraint theory.[43] The coefficient of FIRMK_{t-1} is negative and marginally significant in the GLS version of Eq. (3), perhaps providing some evidence of a positive price response to firm-specific shocks that decrease the firm's capacity.[44]

Stock firms have significantly higher prices than mutuals; and publicly traded stock firms tend to have higher prices than nontraded firms, perhaps reflecting the fact that more information is available on publicly traded insurers, reducing information asymmetries and boosting prices. The proportion of business in GL also has a significant positive relationship with price. The Best's rating is not significant at conventional levels, after controlling for capitalization.

Contributions of new equity. The theoretical model implies that an insurer has an optimal capital structure, defined as an optimal ratio of assets to liabilities (which may change over time depending upon market conditions). Denoting the optimal asset-to-liability ratio as k, we have

$$A_t = \text{kLIAB}_t = A_{t-1} + \text{AYINC}_t + \text{NEWCAP}_t - \text{LRAL}_t + (\text{LIAB}_t - \text{LIAB}_{t-1}). \tag{11}$$

Rearranging terms,

$$\text{NEWCAP}_t = (k-1)\text{LIAB}_t - K_{t-1} - \text{AYINC}_t + \text{LRAL}_t. \tag{12}$$

[43] Winter's (1994) results also imply a positive relationship between price and transitory deviations in surplus in the 1980s, in contrast to his finding of a negative relationship in previous decades.

[44] However, this result is most likely due to the fact that FIRMK_{t-1} to some extent double counts information already included in the three capital change variables (LRAL_{t-1}, AYINC_{t-1}, and NEWCAP_{t-1}). When the capital change variables are omitted from the equation, the coefficients of both INDUSK_{t-1} and FIRMK_{t-1} are positive and statistically significant. The bivariate correlations between FIRMK_{t-1} and AYINC_{t-1}, NEWCAP_{t-1}, and LRAL_{t-1}, are 0.22, 0.30, and -0.17, respectively.

If the asset-to-liability ratio is below k as the result of a loss shock, the extent to which the optimal capital structure is attained through retained earnings vs new equity depends on the effect of the shock on parameter values (see Table I). A positive relationship between new capital and loss shocks in our empirical tests would provide evidence that the conditions for new capital issuance were satisfied during our sample period.

Table IV reports OLS, TSLS, and GLS estimates of Eq. (12) with other control variables added. The GLS version of the model uses an instrumental variables approach to control for the endogeneity of the price change variable. The results are consistent with the hypothesis that firms have optimal capital structures. New capital flows are positively related to the growth in liabilities (LIAB) and inversely related to new flows of internal capital (AYINC). New capital is also positively related to the retroactive loss shock (LRAL), suggesting that firms tended to raise capital following shocks during the sample period. This in turn suggests that shocks increase the quality elasticity of demand and/or the markup of price over marginal cost sufficiently so that it is optimal to issue new equity (see Table I).

Capital flows are inversely related to the two capacity variables (INDUSK and FIRMK).[45] These results could be interpreted as contrary to the capacity constraint theory, which implies that capital depletions due to shocks are almost always replenished through retained earnings. On the other hand, the results with the capacity variables could imply that shocks during this period were large enough to cause insurers to reach boundary points at which raising new capital is optimal under the assumptions of the capacity constraint theory (see Winter, 1994).

The issue of new capital is positively related to the change in price (between years t and $t-1$). The relationship is significant at the 5% level in the OLS model and at the 10% level in the GLS model (one-tail tests), but is not significant in the TSLS model. The results thus provide some support for the hypothesis that capital markets look to price changes for assurance that infusions of new equity will not raise the value of old liabilities sufficiently to penalize shareholders.

Stock insurers raised more capital from external sources than did mutuals, as expected since the sources of new capital to mutual insurers are very limited. There is a positive relationship between the traded-firm dummy variable and new equity flows, although this variable is significant only in the GLS equation. New equity issue is not significantly related to the GL premium share. The Best's rating variable has a negative coefficient and is significant in the OLS and TSLS regressions but not significant in the GLS regression.

In equations not reported, we also estimated the new capital equation using lagged values of the asset and liability variables, which would be appropriate

[45] The findings with respect to the other variables were qualitatively unchanged when the two capacity variables were omitted from the equation.

TABLE IV
FLOWS OF EXTERNAL CAPITAL: 1980–1988

Variable	OLS	TSLS	GLSIV
Constant	−0.039	−0.037	−0.010
	−1.827	−1.680	−0.746
Stock: Dummy variable = 1 if the insurer is a stock company, 0 otherwise	0.011	0.011	0.009
	2.448	2.323	3.614
BESTSA($t-1$) = Best's A or A+ rating($t-1$) Dummy variable = 1 if A or A+, 0 otherwise	−0.011	−0.011	−0.002
	−1.940	−2.001	−0.478
Traded: Dummy variable = 1 if the insurer is a publicly traded stock company, 0 otherwise	0.004	0.004	0.008
	1.258	1.251	−3.146
GLSHARE(t) = other liability premiums written(t)/total premiums written(t)	−0.004	−0.004	0.015
	−0.216	−0.204	1.516
LIAB(t) = liabilities(t)/liabilities($t-1$)	0.090	0.091	0.056
	6.237	6.254	5.468
$K(t-1)$ = total equity capital($t-1$)/liabilities($t-2$)	−0.023	−0.024	−0.070
	−1.402	−1.489	−5.034
AYINC(t) = accident year income(t)/liabilities($t-1$)	−0.083	−0.074	−0.092
	−2.165	−1.842	−4.702
LRAL(t) = loss reserve adjustment(t)/liabilities($t-1$)	0.202	0.216	0.185
	4.302	4.343	6.015
price change(t) = PRICE(t) − PRICE($t-1$)	0.028	0.013	0.028
	2.266	0.553	1.535
FIRMK($t-1$)	−0.011	−0.011	−0.013
	−1.647	−1.648	−2.893
INDUSK($t-1$)	−0.031	−0.034	−0.021
	−2.919	−3.003	−3.513
Adjusted R^2	0.273	0.264	0.390
Sample size	405	405	405

Note. Dependent variable = (inflows of new equity capital(t))/(total liabilities($t-1$)). For each variable, the top number = the coefficient and the lower number = asymptotic t ratio. R^2 is redefined in the GLS models to use weighted sums of squares (Kmenta, 1986, Eq. (2.12)). OLS = ordinary least squares; TSLS = two-stage least squares; GLSIV = generalized least squares with instrumental variables uded to adjust for endogeneity of PRICE CHANGE(t). The lag structure is indicated by the notation $(t-i)$, t = current period. BESTSA = dummy variable = 1 if the insurer is rated A or A+ by the A.M. Best Co., 0 otherwise. Accident year income is defined as calendar year net income plus the loss reserve adjustment, where the loss reserve adjustment is the component of a given year's calendar year losses incurred attributable to adjustments in reserves for all prior years. In the PRICE variable, net premiums written = direct premiums written + reinsurance premiums assumed − reinsurance premiums ceded; and losses are discounted using U.S. Treasury yield curves from Coleman, Fisher, and Ibbotson (1989) and industry-wide loss payout proportions from A.M. Best Company, "Best's Aggregates and Averages," 1990, 1991. To obtain FIRMK($t-1$) and INDUSK($t-1$), we define $K(t-1)/K(5)$, where $K(t-1)$ = real equity capital in year $t-1$ and $K(5)$ = average real equity capital over the period $t-2$ to $t-6$. Real capital is obtained by deflating equity capital by the CPI, with 1982 as the base year. INDUSK($t-1$) = the predicted value from a regression of $K(t-1)/K(5)$ for the firms in the sample on the analogous value for the industry as a whole; FIRMK($t-1$) = the residual from this regression.

if expected current values are best estimated by prior values. The results are generally similar, although significance levels are lower.[46]

4. CONCLUSION

This paper provides new theoretical analysis and empirical evidence on the relationship between loss shocks, capitalization, and price in insurance markets. In our model, insurance is supplied by firms that are subject to default risk. Demand is inversely related to insurer default risk and imperfectly price elastic due to factors such as information asymmetries and private information in insurance markets. We generalize the standard risky-debt model of the firm to include two classes of liabilities, "old" liabilities generated by prior policies on which no further premiums can be collected and "new" liabilities for policies priced and issued in the present. This enables us to model the effect on price of a retroactive loss shock that affects prior liabilities such as the shock associated with the liability insurance crisis of the mid-1980s.

Our model predicts that price should be positively related to financial quality. The model also implies that firms have optimal capital structures because of the relationship between insurance demand and financial quality. Thus, firms are predicted to raise capital in response to loss shocks and other changes that increase leverage, unless such shocks severely increase information asymmetries in capital markets. An inverse relationship between prices on new policies and loss shocks to prior liabilities is a plausible (but not necessary) prediction of our model, depending on demand and capital market conditions. Thus, our model is more general than the capacity constraint theory, which imposes a zero insolvency-probability constraint and predicts a positive relationship between loss shocks and price.

The empirical results are consistent with the predictions of the model. Using firm-level data, we find that price is positively related to financial quality, measured by the ratio of equity to liabilities. Price is inversely related to loss shocks to prior liabilities during our sample period. However, price is less responsive to depletions of capital from loss shocks to prior liabilities than to changes in capital from other sources, suggesting that a shift in demand, capital constraints, or an increase in the cost of capital

[46] We also conducted Hausman tests of the endogeneity of $\mathrm{LIAB}_{(t)}$, $\mathrm{AYINC}_{(t)}$, and $\mathrm{LRAL}_{(t)}$. These variables are possibly endogenously related to the dependent variable because insurers can affect the reported values by manipulating loss reserves. If such manipulations are more likely when new capital is being raised, an endogeneity problem could be present. Our tests failed to reject the null hypothesis that these variables are exogenous. This may imply that capital is more likely to be raised after announcements of legitimate reserve strengthening and price increases that reassure capital markets that significant hidden liabilities are not present.

may have mitigated the negative impact of loss shocks on price during the sample period. Price is inversely related to firm-specific deviations of capital from normal levels (GLS equation) but positively related to industry-wide capital deviations.

Inflows of external capital are positively related to loss shocks, consistent with the hypothesis that firms have optimal capital structures. We also find some evidence that capital flows are positively related to price increases, as predicted if capital markets require assurance that equity will not be penalized to benefit prior liability-holders.

Our analysis of industry aggregate data indicates that prices, as measured by average loading ratios rather than premium levels, increased only modestly over pre-crisis levels following the sharp decline of the mid-1980s. The more dramatic increases in premium levels seem to have been driven primarily by increased loss expectations and declines in interest rates, and capacity constraints do not appear to have played a significant role. Although our model and empirical evidence provide an explanation for movements in prices and capital flows during the crisis period, they do not attempt to explain the shortages of coverage, changes in contract terms, and "nonlinear" pricing which reportedly occurred during the crisis. These features of the crisis were more likely caused by information asymmetries and uncertainty about liability rules that made it unusually difficult for insurers to price coverage for some classes of risks in the rapidly changing environment of the mid-1980s. These factors are explored more fully elsewhere in the literature (e.g., Priest, 1987, 1991; Berger and Cummins, 1992).

APPENDIX A: DERIVATION OF THE TWO-CLASS OPTION MODEL

Consider an insurer with stochastic assets and two classes of stochastic liabilities. Assume that assets and liabilities follow diffusion processes

$$\begin{aligned} dA &= \mu_A A\,dt + \sigma_A A\,dz_A \\ dL_1 &= \mu_{L_1} L_1\,dt + \sigma_{L_1} L_1\,dz_{L_1} \\ dL_2 &= \mu_{L_2} L_2\,dt + \sigma_{L_2} L_2\,dz_{L_2}, \end{aligned} \tag{A1}$$

where A, L_1, L_2 = market values of assets and liabilities (classes 1 and 2),
μ_A, σ_A = drift and diffusion parameters for assets,
μ_{L_i}, σ_{L_i} = drift and diffusion parameters for liability class i, $i = 1, 2$,
dz_A, dz_{L_1}, dz_{L_2} = possibly dependent standard Brownian motion processes.

The Brownian motion processes are related as follows: $dz_A\, dz_{L_1} = \rho_{A1}\, dt$, $dz_A\, dz_{l_2} = \rho_{A2}\, dt$, $dz_{L_1}\, dz_{L_2} = \rho_{12}\, dt$, where ρ_{Ai}, $i = 1, 2$, = instantaneous correlation coefficients between the Brownian motion processes for assets and liability classes 1 and 2, respectively, and ρ_{12} = instantaneous correlation coefficient for liability classes 1 and 2.

Both assets and liabilities are assumed to be priced according to an intertemporal asset pricing model, such as the intertemporal capital asset pricing model (ICAPM). The ICAPM implies the following return relationships:

$$\mu_A = r_f + \pi_A\text{, for assets, and}$$

$$\mu_{L_i} = r_{L_i} + \pi_{L_i}\text{, for liability classes } i = 1, 2,$$

where r_{L_i} = the inflation rate in liability class i, and

π_j = the market risk premium for asset $j = A, L_1, L_2$.

The Fisher hypothesis is assumed to hold so that $r_f = r_r + r_I$, where r_r = the real rate of interest and r_I = economy-wide rate of inflation. The economy-wide rate of inflation will not in general equal the inflation rates on the two classes of insurance liabilities. If assets (and liabilities) are priced according to the ICAPM, the risk premium would be:

$$\pi_j = \rho_{jm}(\sigma_j/\sigma_m)[\mu_m - r_f],$$

where μ_m, σ_m = the drift and diffusion parameters of the Brownian motion process for the market portfolio, and ρ_{jm} = the correlation coefficient between the Brownian motion process for asset j and that for the market portfolio.

The value of an option on the two-liability insurance company can be written as $P(A, L_1, L_2, \tau)$, where τ = time to expiration of the option. Differentiating P using Ito's lemma and invoking the ICAPM pricing relationships for assets and liabilities yields the differential equation

$$\begin{aligned} Pr_f = {} & r_f P_A A + r_{L_1} P_{L_1} L_1 + r_{L_2} P_{L_2} L_2 - P_\tau \\ & + \frac{1}{2}\sigma_A^2 P_{AA} A^2 + \frac{1}{2}\sigma_{L_1}^2 L_1^2 P_{L_1L_1} + \frac{1}{2}\sigma_{L_2}^2 P_{L_2L_2} L_2^2 \\ & + P_{AL_1} A L_1 \sigma_{A1} + P_{AL_2} A L_2 \sigma_{A2} + P_{L_1L_2}\sigma_{12} L_1 L_2. \end{aligned} \tag{A2}$$

Risk and the drift parameters (μ_j) have been eliminated by using the ICAPM and taking expectations. It is also possible to do this by using a hedging argument, provided that appropriate hedging securities are available.

The next step is to use the homogeneity property of the options model to change variables so that the model is expressed in terms of the asset-to-liability ratio x, the option value-to-liability ratio $p = P/L$, and the liability proportions $w_1 = L_1/L$ and $w_2 = L_2/L$, where $x = A/L$ and $L = L_1 + L_2$. This requires the assumption that a lognormal diffusion approximation can be used for $L_1 + L_2$. The assumption about additivity of lognormals is also used in the discrete time option pricing literature (e.g., Stapleton and Subramanyam, 1984). The result is the differential equation

$$pr = xp_x r - p_\tau + \frac{1}{2}x^2 p_{xx}\sigma_n, \tag{A3}$$

where $r = r_f - w_1 r_{L_1} - w_2 r_{L_2}$,
$\sigma_n = \sigma_A + w_1^2\sigma_{L_1}^2 + w_2^2\sigma_{L_1}^2 - 2w_1\sigma_{A1} - 2w_2\sigma_{A2} + 2w_1w_2\sigma_{12}$,
σ_j = the diffusion parameter for process j($j = A$ = assets, $j = 1$ = liability class 1, and $j = 2$ = liability class 2), and
σ_{jk} = the covariance parameter for processes j and k.

Equation (A3) is the standard Black–Scholes differential equation, where the optioned asset is the asset-to-liability ratio (x).

APPENDIX B: COMPANIES IN THE SAMPLE

Aetna Life & Casualty Group
Allstate Insurance Group
American Financial Group
American General Group
American International Group
American Universal Group
Chubb Group of Insurance Companies
Cigna Group
Cincinnati Financial Group
Commercial Union Insurance Companies
Continental Insurance Companies
CNA Insurance Companies
Crum & Forster Companies
Electric Mutual Group
Employers Reinsurance Group
Employers of Texas Group
Liberty Mutual Group
Lincoln National Group
Motors Insurance Group
Munich Group
Nationwide Group
Northwestern National Group
Ohio Casualty Group
Orion Group
PMA Group
Prudential of America Group
Reliance Insurance Companies
Royal Insurance Group
St. Paul Group
Sentry Insurance Group
Skandia America Group
State Farm Group

Farmers Insurance Group
Federated Mutual Group
Fireman's Fund Companies
General Reinsurance Group
Hartford Insurance Group
Home Group Insurance Companies
Kemper Group
Swiss Reinsurance Group
Transamerica Insurance Group
Travelers Insurance Group
United States F&G Group
W.R. Berkley CP Group
Zurich Insurance Group—U.S.

REFERENCES

Berger, L. A., and Cummins, J. D. (1992). Adverse selection and equilibrium in liability insurance markets, *J. Risk Uncertainty* **5,** 273–288.

Cagle, J. A. B., and Scott E. Harrington, S. E. (1995). Insurance supply with capacity constraints and endogeneous solvency risk, *J. Risk Uncertainty* **11,** 219–232.

Clarke, R. N., Warren-Boulton, F., Smith, D., and Simon, M. J. (1988). Sources of the crisis in liability insurance: An economic analysis, *Yale J. Reg.* **5,** 367–395.

Coleman, T., Fisher, L., and Ibbotson, R. (1989). ''U.S. Treasury Yield Curves: 1926–1988.'' Moody's Investors Service, New York.

Cummins, J. D. (1988). Risk-based premiums for insurance guaranty funds, *J. Finance* **43,** 823–839.

Cummins, J. D., and Grace, E. (1994). Tax management and investment strategies of property-liability insurers, *J. Bank. Finance* **18,** 43–72.

Cummins, J. D., Harrington, S. E., and Niehaus, G. (1994). A financial analysis of risk-based capital requirements for property liability insurers, *in* ''The Financial Dynamics of the Insurance Industry'' (E. Altman and I. Vanderhoof, Eds.), Irwin Professional, Homewood, IL.

D'Arcy, S., and Doherty, N. A. (1990). Adverse selection, private information, and lowballing in insurance markets, *J. Business* **63,** 145–164.

Danzon, P. (1984). Tort reform and the role of government in private insurance markets, *J. Legal Stud.* **13,** 517–550.

Doherty, N. A., and Garven, J. (1986). Price regulation in property-liability insurance: A contingent-claims approach, *J. Finance* **41,** 1031–1050.

Doherty, N. A., and Kang, H. B. (1988). Interest rates and insurance price cycles, *J. Bank. Finance* **12,** 199–214.

Fairley, W. B. (1979). Investment income and profit margins in property-liability insurance: Theory and empirical results, *Bell J.* **10,** 192–210.

Geske, R. (1977). The valuation of compound options, *J. Finan. Econ.* **7,** 63–81.

Gron, A. (1989). ''Property-Casualty Insurance Cycles, Capacity Constraints, and Empirical Results,'' Ph.D. dissertation, Department of Economics, Massachusetts Institute of Technology, Cambridge, MA.

Gron, A. (1994). Capacity constraints and cycles in property-casualty insurance markets, *Rand J. Econ.* **25,** 110–127.

Harrington, S. E. (1988). Prices and profits in the liability insurance market, *in* ''Liability: Perspectives and Policy'' (R. E. Litan and C. Winston, Eds.), The Brookings Institution, Washington, D.C.

HARRINGTON, S. E., AND DANZON, P. M. (1994) Price cutting in property liability insurance, *J. Business* **67,** 511–538.

KLEIN, R. (1995). Insurance regulation in transition, *J. Risk Insur.* **62,** 363–404.

KMENTA, J. (1986). "Elements of Econometrics," 2d ed. Macmillan, New York.

KRAUS, A., AND ROSS, S. (1982). The determination of fair profits for the property-liability insurance firm, *J. Finance* **33,** 1015–1028.

MAYERS, D., AND SMITH, C. JR. (1988). Ownership structure across lines of property-casualty insurance, *J. Law Econ.* **31,** 351–378.

MERTON, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates, *J. Finance* **29,** 449–470.

MYERS, S., AND COHN, R. (1987). Insurance rate regulation and the capital asset pricing model, *in* "Fair Rate of Return In Property-Liability Insurance" (J. D. Cummins and S. E. Harrington, Eds.), Kluwer Academic, Norwell, MA.

MYERS, S., AND MAJLUF, N. S. (1984). Corporate financing and investment decisions when firms have information that investors do not have, *J. Finan. Econ.* **13,** 187–221.

NATIONAL ASSOCIATION OF INSURANCE COMMISSIONERS (1993). Insurers rehabilitation and liquidation model act, *in* "Model Laws Regulations and Guidelines," Kansas City, MO.

PRIEST, G. (1991). The modern expansion of tort liability: Its sources, its effects, and its reform, *J. Econ. Perspect.* **5,** 31–50.

PRIEST, G. (1987). The current insurance crisis and modern tort law, *Yale Law J.* **96,** 1521–1590.

SCHERER, F. M., AND ROSS, D. (1990). "Industrial Market Structure and Economic Performance," 3rd ed. Houghton Mifflin, Boston.

STAPLETON, R. C., AND SUBRAHMANYAM, M. G. (1984). The valuation of multivariate contingent claims in discrete time models, *J. Finance* **39,** 207–238.

WINTER, R. (1994). The dynamics of competitive insurance markets, *J. Finan. Intermed.* **3,** 379–415.

UNITED STATES DEPARTMENT OF JUSTICE (1986). "Report of the Tort Policy Working Group on the Causes, Extent, and Policy Implications of the Current Crisis in Insurance Availability and Affordability." Washington, D.C.

UNITED STATES GENERAL ACCOUNTING OFFICE (1988). "Liability Insurance: Effects of Recent 'Crisis' on Businesses and Other Organizations." GAO/HRD-88-64. Washington, D.C.