The Design of Insurance Contracts When Liability Rules are Unstable

Neil A. Doherty*

ABSTRACT

Recent troubles in liability insurance markets have been associated with a destabilization of liability rules. A nautral policy response is seen in proposals for legislative reform of liability rules with the aim of increased stability. An alternative track explored here examines the design of insurance contracts. The effects of judicial and legislative process on the insurance pooling function are examined in a rational expectations setting. It is then shown that the conditions for closure of policies characterized as "random premium contracts" are weaker than for traditional contracts given instability of liability rules. Finally, it is shown that the insurance market has responded to recent crises by organizational and contractual innovations that embody the random premium design. These responses include the emergence of mutual like forms of organization and the "claims made" policy.

Baruch Berliner, 1982, outlined the ideal conditions for the functioning of an insurance market. Amongst other conditions is a requirement that there be a large number of independent loss exposures. This condition ensures the operation of the Central Limit Theorem and the insurer's risk will essentially be diversified away. Failure to satisfy this condition appears to be at the heart of many of the recent troubles in liability insurance markets. For example, the recent liability insurance "crisis" in the United States appears to be a response to a destabilization of the legal system.[1] Insurers argue that they are able to insure the liabilites of clients arising under an unchanging set of liability rules, but they cannot insure against changes in the rules themselves.

Liability insurance markets have responded with different designs of insurance or risk bearing contracts and with changes in the forms of organizations that bear risk. For example, liability insurers have introduced the so called "claims made" contract which, in effect, shortens the period over

---

*Neil A. Doherty is Professor of Insurance and Risk Management, Wharton School, University of Pennsylvania.

This research was supported by the National Science Foundation under grant #88-09299 for which I am most grateful.

The author wishes to thank Scott E. Harrington for helpful comments.

[1] See Wade, (1987) and Rabin, (1987). A particularly severe illustration is the pollution insurance market, see Kleindorfer (1986).

which changes in the level of expected loss may commonly affect all outstanding policies. A more dramatic change has been the shift from stock forms of insurance organization to mutuals, group captives, industry pools and reciprocals. These structures share the risk among a group of firms (or individuals) commonly exposed. Equity in the structure is held by the policyholders. Such structures are referred to in the current article as mutuals.

These changes in contract or organizational design have a similar effect. The premium for any given period of cover is random. It is subject to retroactive adjustment on the basis of new information concerning the aggregate loss in the pool. For example, the mutual may pay a dividend (positive or negative) to its policyholders which is related to the aggregate loss in the pool. The policy holder buying a claims made policy will find that losses which may have arisen, but which have not been presented as claims, within the policy year will be priced in the future in a future insurance contract. The policyholder does not know how much he or she will have to pay to insure this year's liabilities. The premium is, in effect, random.

The next section of the article describes how correlations arising both from the incidence of losses and the generation of new information can lead to undiversified insurance portfolios. This section also shows how judicial and legislative processes can lead to such problems. A subsequent section presents a formal analysis of optimal insurance purchase with random premiums showing the conditions under wich random premium contracts are preferred. The final section addresses the contractual and organizational innovations which have emerged in theses troubled markets.

The insurance markets described are mainly for corporate insurance buyers. This is certainly true of the markets for pollution liability and products liability. For this reason, the analysis is conducted for the corporate insurance buyer and is driven by the value maximization objective. The prevalence of mean variance asset pricing models points to the use of the two moment framework. Any generalization of the results to individuals can only be made by approximation since their utility functions may not be completely described in two moments.[2] The choice of a mean variance framework has another virtue. If, as suggested, the failure of diversification is primarily responsible for the troubles of the liability insurance market then portfolio properties need to be analyzed explicitly. This is most conveniently achieved in the mean variance framework.

---

[2] This caution is particularly relevant if we examine the medical malpractice market where many of the buyers are individual physicians. This market has suffered similar problems to the commercial liability insurance market. Some comfort in the mean variance approximation may be gleaned from the fact that this market has evolved similar solutions to the problems of legal uncertainty as the non commercial liability market. These solutions include the use of mutuals, reciprocals and claims made policies. See Danzon, (1986); Doherty and Dionne (1987).

Risk Pooling and the Law of Large Numbers

Diversification with Rational Expectations and New Information

From the law of large numbers it is known that an insurance market with a large number of independent exposures will substantially reduce portfolio risk. An obvious shortfall from the ideal conditions arises when the events that are insured are correlated. Other pooling problems may arise in more subtle ways. The process of concern here relates to the effects of new information revealed to the insurer during the life of the insurance contract on the loss reserves carried by the insurer. If this information reveals that the reserving errors are correlated, the effect will be similar to that of correlated losses. In this section, both types of correlation are examined and it is shown that reserve error correlation may be a major factor in the recent problems experienced in the markets for liability insurance.

Consider an insurance contract for policyholder "i" that is written at time 0 with the expectation that some random loss, L_i will be paid at some future time. A convenient assumption is that the timing of the loss is known but that its value is not. This is not crucial to the main results but it does help to highlight the role of information in insurance price setting. Time 1 is defined as immediately before the loss, and time 2 is immediately after the realization of loss. Based on the insurer's expectation of the loss at the time the contract is written, time 0, the insurer will form a reserve. Following Cummins and Outreville, (1987), rational expectations on the part of the insurer are assumed. Thus;

$$L_i = E(L_i \mid \phi) + e_i \tag{1}$$

where the term ϕ denotes information available to the insurer at time "t". The expectation formed by the insurer is conditional on this information. By assumption the error term, e, does not exhibit serial correlation and satisfies:

$$E(e_i \mid \phi_t) = 0 \text{ for all i}$$

Immediately before the loss, time 1, the insurer's expectation of loss is:

$$E(L_i \mid \phi_1) = E(L_i \mid \phi_0) + u_i \tag{2}$$

where u_i is the insurer's revision in expectations between time 0 and time 1. Thus,

$$L_i = E(L_i \mid \phi_0) + u_i + e_i \tag{3}$$

A zero correlation is assumed between u and e. Again, this is not crucial but it does simplify the exposition. To complete the assumptions, the pooling properties are examined for an insurance fund in which all policyholders are identical. This is interpreted to mean that the insurer's estimates of the loss distributions for all policyholders taken at any time prior to the realization, are identical. Again, this assumption simplifies without great loss of generality.[3] For the insurer to pool risk successfully, the variance of the revealed loss per policy must converge as follows:

$$\mathrm{VAR}[\Sigma\{L_i - E(L_i \mid \phi_0)\}/n] \rightarrow 0$$

as the number of policies, n, $\rightarrow \infty$

Substituting from (3) and taking the variance gives:

$$\mathrm{VAR}[\Sigma\{L_i - E(L_i\ \phi_0\}/n]$$

$$= (1/n)[\mathrm{VAR}u_i + \mathrm{VAR}\ e_i] + [(n-1)/n][\mathrm{COV}(u_i,u_j) + \mathrm{COV}(e_i,e_j] \quad (4)$$

I will refer to the portfolio risk described by equation 4 as the *insurer's risk*.[4] While the first expression on the right hand side of (4) converges conveniently to zero, the second expression does not unless the covariances are zero. This implies that there are two remaining sources of insurer's risk, correlation between the e's and between the u's. Theses are referred to as *event correlation* and *information correlation.*

(b) Liability Rules and Portfolio Correlations

How might the correlations described in equation 4 arise? Event correlations are straightforward. For example, the indicence of injury or harm from a defective product may be related to the incidence of use. Thus, if an automobile has a design defect that causes ineffective braking in poor weather conditions, the incidence of product liability suits would vary with weather conditions. Event correlation can and does arise to some degree in insurance markets. But the main concern here is with the liability rules themselves. Does the process by which these rules are made and enforced cause correlations, specifically information correlations, between the ultimate payoffs on different policies? Two different processes are described by which such correlations can arise.

(1) Changes in liability rules. When an insurance policy is written, the liability rules in place can be reflected in the policy terms, especially the premium. But the ultimate resolution of claims on a policy may be made months or even years after the policy was written and the cumulative effect of changes in liability rules in the interval can be substantial. Such rule changes are likely to have a common directional effect on all policies, or a subset, in the insurance portfolio. This will give rise to information correlations described above. Liability rules may change as a result of legislative or judicial processes. To provide some focus, illustrations will concentrate on environmental impairment liability (EIL). Though problems of insurability of EIL have been especially severe, even when compared with other liability insurance markets,[5] it does offer well documented illustrations of the process by which issues of stability and enforcement of liability rules affect insurability.

A dramatic example of legislative change in liability rules is the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA) of 1980,

---

[3] See Doherty (1985) for an analysis of diversification in heterogeneous insurance portfolios.

[4] This notation follows Houston, (1964) and Cummins, (1973).

[5] See Kleindorfer (1986), Kleindorfer and Kunreuther, (1987), Katzman (1988), and Doherty, Kleindorfer and Kunreuther (1988).

commonly known as "Superfund". This act imposed retroactive liability for the cost of cleanup of designated landfills on those who previously had dumped hazardous materials on the sites. The issue of whether general liability policy forms cover cleanup costs is a matter of dispute.[6] However, if the cleanup issue is resolved against insurers, the Act has, in effect, imposed retroactively a new area of cover on all policies issued piror to CERCLA.[7] Of course, new legislation can create new areas of liability and eliminate existing ones. The point is that the change simultaneously affects substantial number of policies *that already have been issued* thus causing common directional revision in the parameters of the loss distributions for these policies. Such effects are not diversifiable within the insurance pool.

Similar changes may arise through judicial changes in liability rules through the operation of precedent. An example might be a case which introduces strict liability in an area previously thought to be subject to the negligence standard. However, the precedent setting process and its implications for insurability under the next heading dealing with policy interpretation are described.

2) Policy interpretation. Since many policies are written on a general form (e.g. the comprehensive general liability policy) precedent setting re-interpretations can have similar effect on insurers to the establishment of a new liability on their policyholders. The problem arises through the operation of precendent through which the decisions of one court bind all future courts of similar or lower rank within the jurisdiction should the facts be materially similar. Of course, the material similarity of facts is a matter for judgement and precedents can be overturned in higher courts. Nevertheless, courts do make new laws in this manner, and major innovations in liability rules and policy interpretations can have a common directional effect on policies that already have been written and priced.

An often quoted example of the effect of precedent is the Jackson Township Case.[8] The decision in this case undermined the apparent intent of insurers to exclude gradual pollution from their policies. The policy excludes cover for events unless they are "sudden and accidental". In this case, the court interpreted these words as meaning unexpected and unintentional thereby allowing a claim for gradual pollution. The same court also provided an interpretaion of the word "occurrence" that applied the limits of cover, not to the event that resulted in injuries (i.e. the leakage of hazardous material) but separately to each person or piece of property affected by the event. In this case for example, 97 wells were affected by contaminated water. If this were counted as 97 occurrences the limits of cover under the policy would be increased by a factor of 97. Clearly, such interpretations, if followed, imply a significant expanison of cover for all existing policies. Whereas each policy has been priced on the assumption that the insurer would never have to pay more than a prescribed amount at any one time, the underwriter now finds that the maximum amount payable may have been multiplied by tens, hundreds or even thousands of times! The effect is simultaneous, uni-directional and undiversifiable within the insurance pool.[9]

Quotation of precedent setting cases courts the risk of accusation of bias; Jackson Township may not be in the mainstream of judicial decisions-it may be an outlier.

---

[6] See for example Cheek (1988).

[7] This statement is subject to the qualification that the policies issued prior to CERCLA were "occurrence" based. The occurrence policy was the dominant form at this time, "claims made" policies were introduced in the 1980's.

[8] Jackson Township Municipal Authority v. Hartford Accident and Indemnity Company, [45 A.2d 990, (1982)]. See also Waste Management of the Carolinas Inc. v. Peerless Insurance Company, [323 S.E. 2d 726 (1984)].

[9] For a discussion of similar controversial policy interpretations in EIL cases see General Accounting Office Study [1987].

Indeed some writers have suggested that the trend for courts to give "creative" interpretations against insurers may have reversed recently.[10] But this is beside the point. The process of making law by precedent enhances the weight of outliers; the outliers become the new norms. In this way, it is the outliers that simultaneously change the expectations of underwriters concerning the loss distributions of existing policies.

Comparison of Constant and Random Premium Contracts

Constant Premium and Random Premium Insurance Contracts

The simplest form of insurance contract addressed is one in which the premium is set at time 0 and the policyholder fully insures. The insurer bases the premium on information available at time 0 and, as a consequence, must absorb the risk shown by equation 4. This is defined to be a *constant premium contract*. Table 1 shows the competitive premium (ignoring transaction costs) for such a contract together with the insurer's risk from equation 4. A risk premium, R, is included in the insurance premium. An explanation for such a risk premium is addressed later.

A *random premium contract* is defined as one in which the price of insurance is unknown at the time the cover starts. The price or premium for such a contract includes any "up front" premium paid at the inception of cover plus any subsequent adjustment to that premium. A dividend paid to the policyholder would be counted as an adjustment as would any correction to the premium arising from information that becomes available after the contract is written. Any number of such random premium contracts could be designed. But two generic types are examined here. The first type, labelled Policy 2 in Table 1, bases the premium on information available to the insurer at time 1. The crucial feature being that the premium is fixed after the contract is written but before the loss is revealed. Table 1 shows the competitive premium set by the insurer on the basis of information at time 1, together with the insurer's risk. The advantage of such a contract to the insurer is that the effects of information arising between time 0 and time 1 can be purged from the insurer's risk. Thus, while a risk premium might still be charged, this will be based only on the event correlation.

The second type of random premium contract, policy 3, goes further. The insurer now bases the premium on information available after losses have been realized. The intention is not to nullify the effects of insurance against the policyholder's idiosyncratic risk, but to remove insurer risk. Thus the retroactive correction to the individual's premium is not based on his or her individual loss but on the collective loss in the pool. Thus, a policyholder dividend or what is sometimes called a "participating policy", results. For policy 3, the premium is shown as an initial contribution equal to the expectation of loss at time 0 plus an adjustment based on the difference between the realized aggregate loss and the expectation of aggregate loss at time 0.

---

[10] See Rabin [1987], p84. GAO [1987], p.65.

The respective contracts shown in table 1 differ in the allocation of the insurer's risk identified in equation 4. In the conventional constant premium contract, all insurer risk is borne by the insurer. At the other extreme, policy 3 allocates all the insurer risk to the policyholders, who only insure their idiosyncratic risk. Policy 2 is in between. That part of the insurer's risk arising from new information between times 0 and 1, is borne by the policyholder, the remainder is allocated to the insurer. In the next section, the optimal allocation of insurer risk for the corporate insurance buyer is addressed. Then, illustrations of random premium contracts will be described.

Table 1

Policy 1. Constant Premium, Premium Set at Time 0

Premium $= E(L_i \mid \phi_0) + R_0.$

Insurer's Risk $= (1/n)[\mathrm{VAR}\, u_i + \mathrm{VAR}\, e_i] + [(n-1)/n][\mathrm{COV}(u_i;u_j) + \mathrm{COV}(e_i;e_j)]$

Policy 2, Random Premium, Premium Set at Time 1

Premium $= E(L_i \mid \phi_1) + R_1.$

Insurer's Risk $= (1/n)\mathrm{VAR}\, e_i + [(n-1)/n]\mathrm{COV}(e_i;e_j)$

Policy 3, Random Premium, Premium Set at Time 2

Premium $= E(L_i \mid \phi_0) + [E(L_i \mid \phi_0)/\Sigma E(L_i \mid \phi_0)][\Sigma L_i - \Sigma E(L_i \mid \phi_0)]$

Insurer's Risk $= 0$

Corporate Objective Functions and Risk

The insurance contracts envisaged will be written between two firms. The particular form of organization is not addressed for the time being, except to say that the firms are ultimately owned by investors who are Markowitz efficient in their portfolio choices and share identical information. The firms will be assumed to wish to maximize their market values which are described by a single period, multi factor asset pricing model as follows;

$$V = E(Y) - \delta \mathrm{COV}(Y;M) \tag{5}$$

where Y is the terminal value of the firm
E(.) is the expectations operator
COV(Y,M) is the vector of covariances between Y and the returns on priced factors, M
δ is the vector of prices of these factors

The terminal value, Y, will be allocated amongst claimholders such as stockholders, bondholders, the taxman, etc. Some of these claims will be contingent claims. For example, given non lineraities in tax schedules, the tax claim may be considered as a call option on the pretax value. Such a simple schedule is described with "t" as the marginal tax rate:

$$\mathrm{TAX} = \begin{cases} 0; & \text{if taxable income} < \text{tax shield} \\ t(\text{taxable income}); & \text{if taxable income} > \text{tax shield} \end{cases} \tag{6}$$

The value of such an option is positively related to the variance of the underlying asset (i.e. taxable income).[11] This property applies to more complex progressive tax structures. As Scholes and Wolfson point out "... the progressive tax on the employer induces a degree of risk aversion over lotteries defined in terms of pre-tax cash flows" [1987, p7.][12]

Other claims on the firm will have similar "option like" qualities and may be valued in relation to the variance of the underlying asset. For example, the prospect of bankruptcy or liquidation implies that the receivers or liquidators have a call option on the firm to be exercised when the value of the equity becomes zero. Therefore, the present value of transaction costs will be positively related to variance. Short of bankruptcy or liquidation there may be less drastic reorganization costs that are induced by fluctuations in earnings for the firm or for its operating units. For example, a firm organized on a profit center basis may have to restructure internally, or sacrifice part of the incentive benefit of such a structure, should an individual unit reveal a catastrophic deviation from its expected performance.

In addition to contingent claims on the firm, other costs of the firm may be variance related. For example, wages may include a risk premium that is related to the variance of operating revenues. This assumes that workers are risk averse, relatively undiversified and that the security of their employment is related to the riskiness of the firm's operating revenues.

It may be noted that each of these explanations of firm specific variance related costs rests upon some form of transaction costs, otherwise, all contracts could be reformed to eliminate any firm specific variance related cost. For example, with costless contracting, firms could arbitrage differences in their respective marginal tax rates thus removing tax non-linerities: employees could issue costless claims against their human capital thereby diversifying their individual portfolio risk, and so forth.

Taking note of these variance related claims, the terminal value may be decomposed into two components. The first, PRIM, represents a set of primary cash flows which are valued according to equation 6. The second represents a claim written on these primary flows which is related to their variance. Using value additivity and denoting the value of the variance related claims by the function f[.], the value of the firm is;

$$V = E(PRIM) - COV(PRIM;M) - f[VAR(PRIM);A_f] \tag{7}$$

where A_f is the vector of other arguments of f[.] The last term is simply the present value of the variance related flows. The valuation relationship, f[.],, might be the option pricing model if these claims are indeed contingent claims.

---

[11] The basic properties of options are described in Merton, [1973]. Majd and Myers [1986] have recently used option pricing to value tax liabilities. See also Scholes and Wolfson, [1987].

[12] Alberts and Hite [1983] show how variance related costs affect product market decisions. They consider the leverage decision and show how a financial innovation will be disseminated in a competitive product market. Variance related imperfections, such as costs of financial distress, will affect the product market equilibrium price levels as well as the optimal leverage decision.

The objective functions are now developed for the insurance firm and for a particular corporate insurance buyer. The buyer is identified by the subscript "j" and the transaction with the insurer involves the payment of premium P_j and the settlement of a possible claim L_j. A proportionate insurance contract is envisaged in which the buyer pays some proportion of the full insurance premium and receives a corresponding proportion of the loss from the insurer. The claim is certainly a random variable but so too may be the premium. The insurer already writes a portfolio of "n" policies. The premiums and losses are considered to be the primary cash flows and variances related claims such as taxes will be considered to be written on these flows. Using equation (7), the value of the insurer is:

$$\begin{aligned} V(INS) &= \sum^{n} E(P_i) - \sum^{n} E(L_i) + \alpha E(P_j) - \alpha E(L_j) \\ &\quad - (\delta)\, COV\{[\sum^{n} P_i - \sum L_i + \alpha P_j - \alpha L_j]; M\} \\ &\quad - f\{VAR[\sum P_i - \sum L_i + \alpha P_j - \alpha L_j]; A_f\} \end{aligned} \tag{8}$$

For the insurance buyer the primary cash flows comprise those from ordinary operations, W_j, plus those from the insurance policy and any uninsured loss. The function "g" is used to denote the value of variance related claims to the buyer. Using equation 7 again, the value of the buyer is:

$$\begin{aligned} V(BUY) &= E(W_j) - (1-\alpha)E(L_j) - \alpha E(P_j) \\ &\quad - (\delta)\, COV\{[W_j - (1-\alpha)L_j - \alpha P_j]; M\} \\ &\quad - g\{VAR[W_j - (1-\alpha)L_j - \alpha P_j]; A_g\} \end{aligned} \tag{9}$$

From equations 8 and 9, it is clear that risk affects the firm (insurer or buyer) in two ways. On the one hand, undiversifiable risk is priced by the market. Following the financial economics literature, is called SYSTEMATIC RISK. This is shown by the covariances with the priced factors. The second risk cost relates to the variance of the firms' operating revenues and the particular variance related claims which the firms may have on their operating revenues. Such costs are denoted VARIANCE RELATED COSTS. These costs are specific to the firm.

Optimal Insurance With Random Premiums

For brevity and simplicity, comparisons will be made of the extreme forms of policies shown in Table 1, i.e. the constant premium policy (Policy 1) and the random premium policy (Policy 3). Thus the notation will be abbreviated in this section, all expections are assumed to be made at time 0. The premium for the type 3 policy is;

$$P_i = E(L_i) + (E(L_i))/(\Sigma\, E(L_i))[\Sigma\, L_i - \Sigma\, E(L_i)] + K_i \tag{10}$$

The terms K is a loading levied by the insurer for risk taking and rent. If premiums of the form described by equation 10 are charged to all policyholders, then clearly total premiums must equal total losses plus the loading component *ex post*:

$$\Sigma P_i = \Sigma L_i + \Sigma K_i$$

With such contracts, the insurer's liability portfolio is risk free and the variance and covariance terms in equation 10 disappear. Thus, in a competitive insurance market, the loading charged to each policyholder, K_i, would be zero.

The following proposition may now be stated;

Proposition 1. In a competitive insurance market with random premiums, the Value Maximizing level of insurance for the buyer is that which minimizes the variance related costs of the buyer.

Proposition 1 is proved by maximizing the following:

$$\underset{\alpha}{\text{MAX}} \{V(BUY)\} \quad \text{s.t. equation (10)}$$

The first order condition is:

$$\begin{aligned} \partial\{V(BUY)\}/\partial\alpha &= [E(L_j) - E(P_j) + \delta\, COV(P_j;M) - \delta\, COV(L_j;M)] \\ &\quad - \{\partial g/\partial VAR(B)\}\{\partial VAR(B)\ /\partial\alpha\} \\ &= 0 \end{aligned} \tag{11}$$

where B is the buyer's operating revenue, or $B = W_j - (1-\alpha)L_j - \alpha P_j$.

Proof of the proposition follows buy substituting from equation 10 and noting that, by the assumptions of random premiums and identical insureds, $E(L_j) = E(P_j)$ and $COV(L_j;M) = COV(P_j;M)$.[13]

This proposition has two immediate corollaries. First, in a competitive insurance market, the level of insurance that maximizes buyer value is also Pareto optimal. This follows from the fact that the premium was defined such that the insurer's value would be unchanged by the level of insurance sold to policyholder j at this price. Second, systematic risk plays no part in the Pareto Optimal and Value Maximizing insurance solutions. The result follows immediately from the observation that the systematic risk terms cancel out of the first order conditions. The intuition of this result rests on the capability of the capital market for spreading risk and on the additive properties of systematic risk. If the ownership claims of the buyer and seller are traded, then arbitrage between their respective equity would remove any comparative advantage in bearing systematic risk. This result is well known and has been extensively analyzed, (see Mayers and Smith, 1983 and Main, 1983).

[13] The equality of the expected loss and expected premium follows directly from the definition of the random premium with a zero loading. The equality of the covariances is shown; From the definition of the random premium;

$$COV[\Sigma L;M] = COV[\Sigma P;M]$$

thus

$$(1/n)COV[\Sigma L;M] = (1/n)COV[\Sigma P;M]$$

from the equality of insureds, it immediately follows that

$$COV[L_j;M] = COV[P_j;M]$$

Optimal Insurance with Constant Premiums

Insurance premiums that are constant and prepaid are more familiar than the random premiums envisaged in the previous section. In a competitive insurance market such premiums may be defined such that the effect of writing a policy at this premium leaves the value of the insurer unchanged. It is also assumed that the insurer is at the minimum point on its average cost curve. This implies that the insurer is an efficient risk bearer and has exploited opportunities available to such institutions for reducing the variability of its operating revenues. Such risk reducing activities might include the use of reinsurance, or "trading" in tax shields. The competitive premium is given by:

$$\partial V(INS)/\partial\alpha$$
$$= P_j - E(L_j) + COV(L_j;M) - \{\partial f/\partial VAR(I)\}\,\{\partial VAR(I)/\partial\alpha\} = 0 \tag{12}$$

where I is the operating revenue of the insurer or $I = \Sigma\, P_i - \Sigma\, L_i + \alpha P_j - \alpha L_j$

Proposition 2 can now be stated:

Proposition 2. In a competitive insurance market with constant premiums, the value maximizing level of insurance for the buyer is that which equates the marginal variance related costs of the buyer and seller.

Proposition 2 is proved by maximizing the following:

$$\underset{\alpha}{\text{MAX}}\ \{V(BUY)\}\ \text{s.t. equation (12)}$$

The first order condition is:

$$\partial\{V(BUY)\}/\partial\alpha =$$
$$-[P_j - E(L_j) + \delta\, COV(L_j;M) - \{\partial g/\partial VAR(B)\}\{\partial VAR(B)]/\partial\alpha\}]$$
$$= 0 \tag{13}$$

Substitution from equation (12) reveals that the condition is satisfied at the level of insurance at which:

$$\{\partial f/\partial VAR(I)\}\{\partial VAR(I)/\partial\alpha\} = -\{\partial g/\partial VAR(B)\}\{\partial VAR(B)/\partial\alpha\}$$

Satisfaction of 13 rests upon the comparative advantage in risk bearing between the insurer and the policyholder. The greater the comparative advantage of the insurer (at the margin), the lower the price of insurance and the more constant premium insurance will be purchased. The buyer will compare the price of the insurance (at the margin) with his or her own risk bearing cost. Thus one should expect that constant premium insurance contracts would flourish where the insurer has a comparative advantage in risk bearing but that other devices for apportioning risk would arise where the comparative advantage was reversed. This is addressed further in the next section.

This result rests on the respective variance related costs of the buyer and seller of insurance. As with the random premium contract, systematic risk has no role to play since the opportunity cost of bearing this risk is the same for the investors who hold the equity of the insurer and of the policyholder. It may also be noted that the value maximizing level of insurance for the buyer is also Pareto optimal since the premium is defined to leave the insurer indifferent as to how much insurance is sold.

Comparison of Random and Constant Premium Contracts

The preceding results depend on the variance related costs of the buyer and of the seller. These costs depend, in turn, on the contracts and obligations of the firms and upon the riskiness of their operating earnings. In the absence of a strong positive correlation between losses (a negative value) and other revenues, an increase in insurance will likely reduce the variance of operating revenue. Thus it is plausible and likely that an increase in the level of insurance will reduce that variance related costs to the buyer. Under the random premium contract the buyer will simply continue to purchase insurance as long as this condition holds. If the level of purchase is constrained at full insurance ($\alpha = 1$) then a corner solution may result.

The condition for insurance with constant premiums is somewhat stronger. Insurance will be purchased as long as the insurer has a comparative advantage in bearing risk related costs. This comparative advantage could arise in two ways. First, the insurer may be able to diversify the insurable risk more efficiently than the policyholder. Second, the claims on the insurer's revenues are less sensitive to variance than those of the policyholder. This second condition might be fulfilled if, for example, the insurer was able to use its tax shields fully at all levels of income (see Smith and Stultz 1985) or if its organizational form were such that the expected value of transaction costs in bankruptcy were low relative to those for the buyer (see Mayers and Smith 1983).

A comparison now may be made between the optimal insurance purchases with random and with constant premiums. The comparison is made under the assumptions that the variance related cost functions are concave in α for the insurer and convex for the policyholder. These assumptions are consistent with the second order conditions for a maxima. Figure 1 shows the marginal variance related costs (MVRC) of the insurer and the buyer. The sign of the buyer's MVRC is reversed such that it is shown in positive space alongside that of the insurer. If the insurer's MVRC curve is A, the first order conditions (equations 11 and 13) show a higher level of coverage under the random premium contract, α_R, than under the constant premium contract, α_c. This depicts a market in which the insurer does not have a comparative advantage in bearing risk, such as the markets with highly correlated exposures addressed in this article.

Shifting the comparative advantage changes these results. For example, if the insurer's MVRC curve is C, no cover would be purchased under the constant premium contract but full insurance (a corner solution) would be purchased with a random premium. Conversely, curve B depicts a comparative advantage in risk bearing for the insurer. At levels of coinsurance in the feasible range $0<\alpha<1$, the insurer's MVRC curve is lower than that of the buyer. This situation favors the constant premium contract. These simple contracts are almost universally traded in markets such as fire, automobile and homeowners insurance where correlations are low.

Contractual and Organizational Responses

Mutual Insurance Companies

The random premium risk sharing arrangements envisioned in the previous section may be achieved or approximated in various ways. Two recent market responses are described. One is a change in the organizational form; the other a choice of contract design. Respectively, these are mutual insurance and claims made policies.

A mutual insurance firm is one which is owned by its policyholders; thus the policyholder holds a portfolio comprising an insurance policy and an equity claim on the insurance pool. The total contribution of the policyholder to maintain the pool is the sum of the premium, assumed to be prepaid, and the dividend. Suppose that the mutual distributes to its equityholders the difference between its aggregate premiums and aggregate losses. If this value is negative it is paid by policyholders as an assessment. Also suppose that the premium for the individual is the expected value of loss.[14] Thus the total contribution (premium plus dividend) is exactly that shown in equation 10. The retroactive nature of the dividend/assessment plays the same role as a retroactive adjustment of premium based on collective loss experience. Even if the collective surplus is not distributed, the ownership value of that surplus will presumably have the same value as the dividend.

Mutuals come in different shapes and sizes and other structures that are functionally similar to mutuals also exist. Thus some of the purity of the above arrangement may be lost if the mutual is unable to charge assessments or if the State claims part of the undistributed surplus in the event of liquidation. Other forms that have similar functional effect to mutuals are reciprocal insurance firms, which are unincorporated mutuals, and group captive insurance firms. Group captives are insurance firms owned by a group of parents and formed for the express purpose of insuring the risk of the parents. Essentially the group captive is a corporate mutual, though such firms may be subject to different regulation and taxation depending on location. Some alternatives are described by Mayers and Smith (1986). In effect, each of these forms superimposes the equity claims on to the policyholder claims and achieves the random premium.

Mutuals, and mutual like forms of organization have appeared in liability markets that have been beset by the problems described above. Danzon (1985)

[14] If a higher up front premium is charged, the insurer will reduce or avoid the problems of enforcing negative dividends against policyholders.

has described the problems in the medical malpractice market during the 1970s and 1980s. Before the "crisis" in that market in the 1970s the predominant form of organization was the large stock firm. Now about half of the market is supplied by the so called "bedpan mutuals" which are physician owned mutuals or reciprocals. An even more dramatic transformation to this form has been observed very recently in the pollution insurance field. Following the withdrawal of all major U.S. insureres from this market (with one or two exceptions such as the A.I.G.) new innovations have appeared. Usually through the medium of a major broker, groups of chemical companies have pooled their risks under such firms as the Hypercept, North American Casualty Co-operative, Resource, EPIC, etc. (see Berkowitz 1987). Outside the pollution field groups ranging from cab drivers, to local governments have responded to the liability insurance "crisis" by forming pools that mutualize their risk. Finally, it should be noted that such developments should not come as a surprise. Over a decade ago, Marshall (1974) noted the Pareto superiority of such forms in the face of undiversifiable risk.

Changes in Contract Design; Claims Made v. Occurrence

Another approach lies in contract design. The particular change is the switch being made from occurrence to claims made policies. The occurrence contracts resemble Policy 1 described in Table 1 and the claims made contracts resemble Policy 2. It is shown that claims made policies can reduce the insurer risk in the insurance pool. To do this, compare the riskiness of the earnings generated from liabilities annually contracted under each arrangement. First, the respective contracts must be characterized.

Occurance Policies. Under an occurrence policy the insurer accepts liability for any occurrences arising within the policy period that give rise to liability claims against the policyholder. Claims do not have to be settled in the policy period and may indeed be settled many years after. The occurrence and claims made policies will be compared on the basis of a two year development cycle. All tort awards are made either in the year of occurrence, year r, or in the subsequent year, s. At the end of each year, information is made available. ϕ_t refers to information made available at the end of $t-1$ on which period t decisions are based. At the beginning of each year policies are written. Losses are incurred immediately after the policies are written but not all losses are known to, or filed with, the insurer. Those claims filed in the policy year are paid immediately. the remainder are filed at the beginning of the subsequent year and are paid immediately they are filed. For simplicity, the discount rate is assumed to be zero, though this is not necessary to generate the results.

Some refinement of notation is necessary. Defining L_{irs} as the loss for policyholder i incurred in year r but settled in year s. P_{ir} is the premium charged to i for a policy written in year r. Now the premiums will equal the expected value of loss payments under such contracts (based on information at inception) plus a risk premium R_i. Thus:

$$P_{i1} = E(L_{i11} \mid \phi_1) + E(L_{i12} \mid \phi_1) + (R_{i1})$$

$$P_{i2} = E(L_{i22} \mid \phi_2) + E(L_{i23} \mid \phi_2) + (R_{i2}) \quad (14)$$

etc.

Now, by assumption:

$$L_{irs} = E(L_{irs} \mid \phi_s) + e_{irs}.$$

$$= E(L_{irs} \mid \phi_r + u_{irs} + e_{irs} \quad (15)$$

where $E(e_{irs}) = 0$

The effect of claims made policies considered later will be to reduce or remove the effects of information correlation. To focus attention on this source of correlation, further assume that the e's, are uncorrelated, thus:

$$COV(e_{irs};e_{jrs}) = 0$$

The ultimate earnings from liabilities written in any year can only be determined at the end of the subsequent year. This is not too bothersome in a two year example but, in reality, the final determination could only be made after all statutes of limitations have expired and all pending cases have been settled. As a paractical compromise, annual earnings may be estimated using the reserve necessary to meet liabilities for outstanding claims and those incurred but not yet known to the insurer. Such reserves may be estimated on the basis of information available at the end of the policy year. Since losses paid in year 2 may differ from the amount reserved at the end of the occurrence year, attribute this final adjustment to the recorded profit for the subsequent year. Thus incurred earnings may be tallied at year end. The incurred earnings are thus:

$EARN_r$ = Premiums earned in year r
- Losses incurred and paid in year r
- Expected (reserved) value of losses incurred in year r but to be paid in subsequent year s
- Amount by which losses incurred in year q (the year prior to r), but paid in year r were under reserved

Thus:

$$\begin{aligned} EARN_r &= \sum \{ [E(L_{irr} \mid \phi_r) + E(L_{irs} \mid \phi_r) + R_{ir}] \\ &\quad - [E(L_{irr} \mid \phi_r) + e_{irr}] \\ &\quad - [E(L_{irs} \mid \phi_r) + u_{irs}] \\ &\quad - [e_{iqr}]\} \\ &= \sum R_{ir} - e_{irr} - e_{iqr} - U_{irs} \end{aligned} \quad (16)$$

Now consider the characteristics of the insurer's portfolio of liabilities. Continue to assume that the portfolio comprises identical policies. The expected value of earnings averaged per policy at inception is $E[EARN_r/n]$

which is independent of the number of policies, n. The variance of earnings per policy is:

$$\mathrm{VAR}[\mathrm{EARN}_r/n] \tag{17}$$
$$= \{(1/n)[\mathrm{VAR}(e_{irr}) + \mathrm{VAR}(e_{iqr}) + \mathrm{VAR}(u_{irs})] + [(n-1)/n]\mathrm{COV}(u_{irs};u_{jrs})\}$$
$$\text{which} \rightarrow \mathrm{COV}(u_{irs};u_{jrs}) \text{ as "n"} \rightarrow \infty$$

It is this feature that explains the existence of the risk premium, R. Moreover, the presence of the risk premium will deter the demand for insurance.

Claims Made Policies. Under the claims made policy, cover is offered for claims that are filed with the insurer during the policy period. The claim does not have to be settled in the policy period but, in the example given here all claims are settled once they are made or filed with the insurer. The premiums paid for each year of cover relate only to the liabilities accepted under that cover. These premiums do not include a risk premium since, as will be shown, all risk is diversifiable in the insurance pool.

$$P_{i1} = E(L_{i11} \mid \phi_1) + E(L_{i01} \mid \phi_1) \tag{18}$$
$$= E(L_{i11} \mid \phi_1) + E(L_{i01} \mid \phi_0) + u_{i01}$$
$$P_{i2} = E(L_{i22} \mid \phi_2) + E(L_{i12} \mid \phi_1) + u_{i12} \qquad \text{etc.}$$

A consistent definition of earnings must relate to the earnings from liabilities incurred under the policies written in the year:

EARN_r = Premiums written in year r
− Losses incurred in year r and paid in r
− Losses incurred in year q and paid in r

$$\mathrm{EARN}_r = \sum \{[E(L_{irr} \mid \phi_r) + E(L_{iqr} \mid \phi_q) + u_{iqr}]$$
$$- [E(L_{irr} \mid \phi_r) + e_{irr}]$$
$$- [E(L_{iqr} \mid \phi_q) + u_{iqr} + e_{iqr}]\}$$

Thus;

$$\mathrm{EARN}_r = -\sum\{e_{irr} + e_{iqr}\} \tag{19}$$

for which the expected value $E(\mathrm{EARN}_r/n) = 0$ and $\mathrm{VAR}(\mathrm{EARN}_r/n) \rightarrow 0$ as "n" $\rightarrow \infty$

Notice that it has only been shown that the claims made policy can reduce (in this case remove) portfolio risk that arises from information correlation. By assumption, the e's, were uncorrelated. Had there been event correlation, the portfolio of claims made policies would still be left with this risk. Thus is should be stressed that the switch to claims made cannot be expected to remove all portfolio risk in practice.

If insurance is purchased in successive periods, the essential difference between the occurrence and claims made policies lies in the premium. In occurrence policies, the premium for occurrences in a given year is decided on the basis of information available at the beginning of that year. Though changes in liability rules may cause the insurer to revise its reserves for outstanding liabilities, this will not affect the premium which already has been paid. The insured, assuming full insurance is purchased, is jointly insured against the loss, L_{irs}, and any changes in the liability rules which are picked up in the term u_{irs}. This contrasts with the claims made policy. Here the insurer can defer its estimate of the premium for occurrences giving rise to claims after the occurrence year. This can be seen by comparing the premiums ultimately paid for occurrences within a given year, year 1. From equation 14, the premium for occurrences in year 1 under as occurrence policy is:

$$E(L_{i11} \mid \phi_1) + E(L_{i12} \mid \phi_1) + R_{i1} \tag{20}$$

But for equation 18, the insured will pay for coverage of year 1 occurrences partly in the year 1 premium and partly in the year 2 premium as follows;

$$E(L_{i11} \mid \phi_1) + E(L_{i12} \mid \phi_1) + u_{i12} \tag{21}$$

Comparison of equations 20 and 21 shows that the claims made policy transfers the risk of changing liability rules between time 1 and 2 to the policyholder. The policyholder effectively self insures for such changes in liability rules as shown by u_{irs}. The risk premium R may be thought of as the price of transferring the risk of changing liability rules to the insurer under the occurrence policy.

The current analysis of claims made and occurrence policies has interesting implications for recent anti trust suits brought against the insurance industry by the attorneys general in several states.[15] These sutis allege that the insurers conspired to replace the occurrence form with the claims made form and that this substitution was harmful to policyholders. More particularly, the plaintiffs actions had adverse effects on the price, availability and level of coverage. While the current analysis is silent on the correctness of the process by which the new policy forms were introduced, it does challenge the implication of the attorney generals' position that the claims made form is harmful to the interest of consumers. Indeed, our analysis suggests that this form has facilitated the closure of contracts thereby it has helped to revive a flagging market. Indeed, the claims made form represents a preferred form of contracting under conditions of non independence between insurable risks.

Conclusion

The conclusions of this article are:

1. Many of the recent problems in insurance markets arise because the insurance pool is not able to substantially reduce risk. In liability insurance, these problems arise because the long delays between the writing of the contract and the realization of loss permit a substantial cumulative change in the information base on which loss values are commonly estimated. These

[15] See for example Insurance Review, May 1988, pp 26–32.

changes arise both from legislative and judicial changes in liability rules and from judicial precedents which re-interpret insurance contract wordings.

2. Using a corporate decision framework in which systematic risk is priced in the capital market and in which there are organization specific variance related costs, it is shown that the optimal amount of insurance under a random premium contract is determined only by its effects on the variance related costs of the buyer. The optimal contract under constant premiums depends on the comparative advantages of the buyer and seller in bearing such risk.

3. In markets where the insurer does have a comparative advantage in bearing risk (e.g. where insured events have a low correlation or when the insurer can use its tax shields more effectively than the insured), there will be an active market in conventional, constant premium, insurance contracts. If risk is not easily diversifiable in the insurance pool, the market will favor random premium contracts.

4. Random premium contracts may be achieved in several ways, either through organization design or through contract design. Mutual, and mutual like, insurers effectively randomize premiums in the manner envisaged here. As predicted by the model, these forms of organization have started to emerge in markets plagued by problems of correlation. Medical malpractice and pollution liability insurance markets are examples.

5. An illustration of a change in contract design that has similar effect to the random premium contract, is the claims made liability insurance policy. The insurance industry has introduced such contracts to replace occurrence policies, which, in effect, are constant premium contracts. While not universal, a significant proportion of new liability insurance is now sold on a claims made basis. Again, this trend is consistent with the predictions of the random premium model. Unlike the mutual form which mitigates the effect of event correlation and information correlation, the claims made policy only mitigates the effects of the latter. This conclusion challenges the basis of recent anti-trust suits brought against the insurance industry which suggest that introduction of the claims made policy is harmful to consumers.

REFERENCES

Alberts, W. W. and G. L. Hite 1983, The Modigliani-Miller Leverage Equation Considered in a Product Market Context, *Journal of Financial and Quantitative Analysis,* 18: 425–37.

Berkowitz, J. B. 1987, EIL: Insurance Coverage Isn't Enough, *National Underwriter,* March, 30: 57–68.

Berliner, B. 1982, *Limits of Insurability of Risks,* Prentice Hall, Englewood Cliffs, N.J.

Cheek, L. III 1988, Insurance Issues Associated with Cleaning Up Inactive Hazardous Waste Sites, forthcoming, *Geneva Papers on Risk and Insurance.*

Cummins, J. D. 1973, Insurer's Risk: A Restatement, *Journal of Risk and Insurance,* 41: 147–157.

Danzon, P. M. 1985, *Medical Malpractice: Theory, Evidence and Public Policy,* Harvard University Press, Cambridge, Mass.

Doherty, N. A. 1985, *Corporate Risk Management: A Financial Exposition,* MacGraw Hill, New York.

Doherty, N. A. and G. Dionne, 1987, Insurance with Undiversifiable Risk, Working Paper, Center for Research in Risk and Insurance, University of Pennsylvania.

Doherty, N. A., P. Kleindorfer and H. Kunreuther, 1988 Insurance Perspectives on Hazardous Waste Management Strategy forthcoming in *Geneva Papers on Risk and Insurance.*

General Accounting Office, 1988, *Hazardous Waste Insurance Availability,* RCED-88-2.

Houston, D. B. 1964, Risk, Insurance and Sampling, *Journal of Risk and Insurance,* 31: 511-38.

Keeton, R. E. 1981, The Impact on Insurance of Trends in Tort Law in J. D. Long (ed), *Issues in Insurance,* American Institute for Property and Liability Underwriters, Malvern, PA.

Kleindorfer, P. R. 1986, Enviornmental Liability Insurance: Perspectives on the U.S. Insurance Crisis, Seminar fuer Versicherungswissenshaft der Universitat Hamburg.

Kleindorfer, P. R. and H. Kunreuther (eds), 1987, *Insuring and Managing Hazardous Risks; From Seveso to Bhopal and Beyond,* Springer Verlag, Berlin and New York.

Main, B. G. M. 1983, Corporate Insurance and Taxes *Journal of Risk and Insurance,* 50: 197-223.

Majd, S and S. C. Myers, 1984, Valuing the Government's Tax Claim on Risky Corporate Assets, *N.B.E.R. Working Paper.*

Marshall, J. M. 1974, Insurance Theory: Reserves versus Mutuality, *Economic Inquiry,* 476-92.

Mayers, D. and C. W. Smith Jnr., 1983, On the Corporate Demand for Insurance, *Journal of Business,* 55: 281-96.

______. 1987, Ownership Structure Across Lines of Property Casualty Insurance, *Journal of Law and Economics,* 16: 73-98.

Rabin, R. L. 1987, Environmental Liability and the Tort System, *Houston Law Review,* 24: 27-52.

Scholes, M. S. and M. A. Wolfson, 1987, Taxes and Organization Theory, *Working Paper,* Stanford University.

Smith, C. W. and R. Stultz, 1985, The Determinants of Firm's Hedging Policies, *Journal of Financial and Quantitative Analysis,* 20: 391-406.

Wade, J. W. 1987, An Evaluation of the Insurance Crisis and Existing Tort Law, *Houston Law Review,* 24: 81-96.

Figure 1

Value Maximizing Levels of Insurance