Assurances

Assurances

Insurance Economics

Georges Dionne et Scott E. Harrington

Volume 60, numéro 3, 1992

URI : https://id.erudit.org/iderudit/1104903ar
DOI : https://doi.org/10.7202/1104903ar

Aller au sommaire du numéro

Éditeur(s)

HEC Montréal

ISSN

0004-6027 (imprimé)
2817-3465 (numérique)

Citer ce document

Dionne, G. & Harrington, S. (1992). Insurance Economics. *Assurances*, *60*(3), 369–398. https://doi.org/10.7202/1104903ar

Résumé de l'article

Grâce à l’aimable autorisation des auteurs et éditeurs, nous sommes heureux d’offrir aux lecteurs les introductions extraites de deux ouvrages, parus conjointement, sur l’économie et l’assurance. Le premier ouvrage, « *Contributions to Insurance Economics* », regroupe des textes de recherches et des essais préparés par divers collaborateurs. Le professeur Dionne le décrit ainsi : « The purpose of this book is to fill this gap in literature... The contributions offer basic reference, new material and teaching supplement to graduate students and researchers in economics, finance and insurance. » Le deuxième ouvrage, « *Foundations of Insurance Economics* » collige également les recherches de divers collaborateurs sur la notion de risque, la demande d’assurance, les structures de marché, la tarification, la réglementation et autres aspects.

Insurance Economics

by

Georges Dionne

and

Scott E. Harrington*

Grâce à l'aimable autorisation des auteurs et éditeurs, nous sommes heureux d'offrir aux lecteurs les introductions extraites de deux ouvrages, parus conjointement, sur l'économie et l'assurance. Le premier ouvrage, « Contributions to Insurance Economics *», regroupe des textes de recherches et des essais préparés par divers collaborateurs. Le professeur Dionne le décrit ainsi : « The purpose of this book is to fill this gap in literature... The contributions offer basic reference, new material and teaching supplement to graduate students and researchers in economics, finance and insurance. »*

Le deuxième ouvrage, « Foundations of Insurance Economics *» collige également les recherches de divers collaborateurs sur la notion de risque, la demande d'assurance, les structures de marché, la tarification, la réglementation et autres aspects.*

~

Contributions to Insurance Economics**

In the Arrow-Debreu economy, different market arrangements are introduced to obtain efficient risk shifting. The role of insurance is to transfer individual risks to parties with comparative advantage in risk bearing. For example, risk-averse

---

* Georges Dionne est professeur titulaire au département des Sciences économiques et chercheur associé au Centre de recherche sur les transports de l'Université de Montréal. Scott E. Harrington est professeur de finance à l'Université de la Caroline du Sud.

** Introduction - *Contributions to Insurance Economics*, Edited by Georges Dionne, Kluwer Academic Publishers, 1992.

individuals pay a fixed price to a more diversified insurer who offers to bear the risk at that price. Under well known assumptions of basic insurance models, an equilibrium is characterized by full insurance coverage. The two standard theorems of welfare economics hold and market prices of insurance are equal to social opportunity costs.

However, risk is not completely shifted in any market. The contributions published in this book discuss many of the reasons that explain incomplete insurance coverage. C. Gollier's article

is concerned with transaction costs and insurers' risk aversion. Coinsurance is explained by either insurers' risk aversion or convexity of transaction costs. Transaction costs are also shown to be a motivation for deductibles. The role of the technical constraint that coverage must be nonnegative is also discussed. Another assumption in the standard models of optimum insurance concerns the implicit agreement of both parties on the probabilities of loss. J.M. Marshall relaxes that assumption and considers the case where the insured is more optimistic than the insurer. The results are functions of how *optimism* is defined. When optimism is defined broadly, the optimum risk bearing contracts can have almost any form and may not resemble insurance contracts. However, under a condition that restricts disagreement, the optimum contracts are insurance arrangements.

Externalities between individuals' losses are sufficient to justify the emergence of liability rules for injuries caused to third parties and risk aversion can explain the presence of liability insurance. P. Danzon and S. Harrington introduce the literature on the demand for and the supply of liability insurance. Emphasis is put on the relationships between liability law, liability insurance, and risk reduction. Asymmetric information on individuals' actions that affect other parties' welfare results in a tradeoff between care and risk bearing in liability insurance as in any form of insurance contracts. R. Winter synthesizes and extends the general theory of optimal insurance under moral hazard. He distinguishes ex-ante moral hazard on the probability of an accident from ex-ante moral hazard on the size of the loss.

The distinction generates predictions as to when deductibles, coinsurance, and coverage limits are observed. Moral hazard, however, does not only affect the nature of insurance contracts but also alters the positive and normative properties of competitive equilibrium. R. Arnott shows how moral hazard may give rise to nonconvex indifference curves, for example, and may alter the definition of competitive equilibrium. At one extreme, competitive equilibrium may not exist, while, at the other, there may be an infinity of equilibria. With moral hazard, neither the first nor the second theorem of welfare economics holds, and market prices do not correspond to social opportunity costs.

From the above theoretical contributions, we retain that insurance is not only an income transfer between different agents in the economy. Under moral hazard, it is predicted that insurance alters incentives for care. This prediction is empirically investigated by four papers in this volume. Moore and Viscusi analyze the net effect of job safety insurance programs where, on one hand, there are positive incentives for safety when the costs of the program are tied to the firm's safety records and, on the other hand, there are disincentives to maintain both safety (ex-ante moral hazard) and recovery periods (ex-post moral hazard) at levels corresponding to full information. From their data set, they found that net workers' compensation insurance provides incentives for safety to firms that outweigh the moral hazard effects. They also obtained that the insurance provided to workers on unsafe jobs reduces the net compensation paid to these workers.

The other three empirical contributions on moral hazard investigate the effects of no-fault automobile insurance. The main finding of D. Cummins and M. Weiss is that no-fault insurance induces drivers to shift property claims from property damage liability coverage into collision coverage. However, the effect of no-fault on total claims is less conclusive. Their results were obtained by using data from the United States. M. Gaudry and R. A. Devlin analyze the effects of the introduction of a pure no-fault system for all bodily-injury accidents in Quebec (1978).

Although they used different methodologies and data, they obtained similar results concerning the increases of accidents in the new system. However, their interpretations of the results differ. While R. A. Devlin attributes significant variations to the reduction in liability, M. Gaudry argues that very little significance is explained by that factor.

Adverse selection is another problem of information that explains partial insurance. G. Dionne and N. Doherty review some of the significant results in the literature. Particularly, they discuss the issues that multi-period contracting raises: time horizon, discounting, commitment of the parties, contract renegotiation, and underreporting of accidents. They also show that different predictions on the evolution of insurer profits over time can be obtained from different assumptions concerning the sharing of information between insurers in competitive markets. Commitment between the parties to the contract is another important factor. The role of risk categorization to improve resource allocation is also discussed and the last section of the paper presents models that consider moral hazard and adverse selection simultaneously. One of the predicted results from the literature on adverse selection is tested by B. Dalhby with data on collision insurance in Canada: if the market is subject to adverse selection, an increase in an insurance policy's premium holding the coverage constant should increase the average claim frequency. He obtains that, in general, the statistical results are consistent with the presence of adverse selection.

In the literature on adverse selection, partial coverage is generally interpreted as a monetary deductible. However, as pointed out by C. Fluet, in many insurance markets the insurance coverage is excluded during a probationary period, which can be interpreted as a sorting device. In fact, he demonstrates that contracts with time-dependent coverage provide a desirable screening mechanism. The presence of adverse selection is also a sufficient condition for risk categorization. In particular, costless imperfect categorization is known to enhance efficiency. However, when categorization entails some costs, the results are ambiguous. S. Rea shows that the gains from separation may be small and the market may give overinvestment in information. The author also obtains that it may be efficient to determine the individual's risk even when neither the insurer nor the insured knows the expected loss ex-ante.

In his survey, D. Cummins reviews the theory of financial pricing of insurance and proposes some extensions that include an option model of the insurance firm and an analysis of insurance company equity as a down-and-out option. By using a model of financial insurance pricing, H. Kunreuther and R. Hogarth show evidence of the importance of ambiguity on the insurance premium setting process. Their analysis is based on recently completed national surveys of both actuaries and underwriters. The paper also explores whether new institutional arrangements are required to replace traditional insurance mechanisms for providing protection that is currently unavailable.

The last survey article (Boyer, Dionne, Vanasse) deals with the econometrics of accident distributions with an application of the different models to automobile accidents. The authors have estimated four categorical models (linear probability, probit, logit, and multinomial logit) and four count data models (Poisson and negative binomial models with and without individual characteristics in the regression component). It is shown that the negative binomial model with a regression component produces a reasonable approximation of the true distribution of accidents. In the last section of this article, they apply the statistical results to a model of insurance rating in presence of moral hazard.

The linear expected utility model has been a very useful tool in the study of optimal insurance decisions. Until recently, it was the standard paradigm used to analyze economic behavior under uncertainty. The linearity in probabilities is directly associated with the independence axiom. This axiom has been challenged by many researchers and some of them have proposed alternative models. E. Karni reviews some of the recent developments in the theory of decision making under uncertainty and apply them to the choice of optimal insurance coverage and the associated comparative statics analysis with respect to the insured's risk aversion. The contribution highlights the methodology of local expected utility analysis.

However, for many contributors in this volume, the classical expected utility model remains a useful approach for applications in insurance. L. Eeckhoudt and M. Kimball show how the concept of absolute prudence can be useful to derive comparative statics results of optimal insurance in presence of uninsurable background risk. In particular, they show that

background independent risk may raise the optimal coinsurance rate and reduce the optimal level of the deductible. Generally, the presence of multiple risks is associated with incomplete markets. B. Ramaswami and T. Roe consider the effect of price risk on the crop insurance decision. They show that increasing demand uncertainty reduces optimal crop insurance whenever risk aversion is constant or decreasing. They also consider the special cases when either output risk or demand uncertainty is the sole cause of price risk. The optimal insurance coverage can also be affected by increases in the risk of the insurable asset. Intuitively, one may anticipate that a risk-averse agent faced with an exogenous mean preserving spread in the loss distribution will demand more insurance. Y. Alarie, G. Dionne and L. Eeckhoudt show that this widespread belief does not always turn out to be true. They then present subclasses of mean preserving increases in risk that make it possible to obtain intuitively acceptable results and apply their analysis to the coinsurance coverage.

One of the common conclusions of these contributions is the absence of perfect risk shifting in any market. Many limiting elements have been identified and analyzed in different chapters: transaction costs, moral hazard, adverse selection, imperfect information, ambiguity, externalities, deviant beliefs, absence of perfect diversification on the supply side, and nonindependence of individual risks. A next important step in the understanding of the functioning of different markets is to propose a general framework that would integrate these elements. As indicated by some authors, a promising way is the integration of economics and financial models of insurance decisions.

Foundations of Insurance Economics***

Although the prevalence of risk in economic activity has always been recognized (Green, 1984), deterministic models dominated economic explanations of observed phenomena for many years. As a result, the economics of insurance has a relatively short history. In early work that formally introduced risk and uncertainty in economic analysis (von Neumann and Morgenstern, 1947; Friedman and Savage, 1948; Allais, 1953; Arrow, 1953; Debreu, 1953), insurance was viewed either as a contingent good or was discussed in relation to gambling. Before 1960, economic literature was largely void of analyses of the nature of insurance markets or of the economic behavior of individual agents in these markets.

During the early 1960s, Kenneth Arrow and Karl Borch published several important articles (Arrow, 1963, 1965; Borch, 1960, 1961, 1962) that can be viewed as the beginning of modern economic analysis of insurance activity. Two of these papers are reprinted in this volume. Arrow was a leader in the development of insurance economics, and more generally, in the development of the economics of uncertainty, information, and communication. Arrow (1965) presented a framework of analysis that explains the role of different institutional arrangements for risk-shifting, such as insurance markets, stock markets, implicit contracts, cost-plus contracts, and futures markets. All of these institutions transfer risk to parties with comparative advantage in risk bearing. In the usual insurance example, risk averse individuals confronted with risk are willing to pay a fixed price to a less risk averse or more diversified insurer who offers to bear the risk at that price. Since both parties agree to the contract, they are both better off.

Risk is seldom completely shifted in any market. Arrow (1963) discussed three of the main reasons that risk shifting is limited: moral hazard, adverse selection, and transaction costs. Arrow (1965) emphasized the problem of moral hazard and

*** Introduction - *Foundations of Insurance Economics*, Edited by Georges Dionne and Scott E. Harrington, Kluwer Academic Publishers, 1992.

suggested that coinsurance arrangements in insurance contracts can be explained by this information problem. Arrow (1963) showed in the absence of moral hazard that full insurance above a deductible is optimal when the premium contains a fixed-percentage loading. He also proved that risk aversion on the part of the insurer is another explanation for coinsurance. Both results were extended by Raviv (1979) and others.

Borch (1960, 1961, 1962) also made significant contributions to the theory of optimal insurance. He developed necessary and sufficient conditions for Pareto optimal exchange in risk pooling arrangements. He also showed, in a general framework, how risk aversion affects the optimal coverage (or optimal shares) of participants in the pool. Although his formal analysis was in terms of reinsurance contracts, it was shown by Moffet (1979) that the same result applies for contracts between policyholders and direct insurers. Borch's formulation of risk exchange influenced the development of principal-agent models (Ross, 1973; Holmstrom, 1979), and it has led to many other applications in the insurance literature. More generally, Borch made many contributions to the application of expected utility theory to insurance and influenced the development of portfolio theory and its applicability to the insurance industry. Finally, Borch's contributions established some important links between actuarial science and insurance economics (Loubergé, 1990).

Outline of this Volume.

The remainder of this introductory essay reviews the main developments of insurance economics subsequent to the pathbreaking work of Arrow and Borch. In the process, the articles included in this volume are introduced. The remaining eight sections include articles on (1) utility, risk, and risk aversion, (2) the demand for insurance, (3) insurance and resource allocation (in which we include Borch, 1962, and Arrow, 1965), (4) moral hazard, (5) adverse selection, (6) insurance market structure and organizational form, (7) insurance pricing, and (8) insurance regulation.

The selection of articles was based on several criteria including the significance of the contribution, the representativeness of the work, and the desire to include empirical as well as theoretical articles. The selection process also considered whether the level of mathematics employed was likely to be accessible to most readers. In a few instances, we showed a slight preference for articles in books that are not as readily available as those published in journals.

For the most part, neither this introductory essay nor the remainder of the volume attempts to cover the wide variety of applications of insurance economics in the areas of health insurance, life insurance and annuities, social insurance, and in the law and economics literature. Instead, we review significant applications and include several articles dealing with property-liability insurance. This approach is at least partially due to our taste (and expertise). However, these articles and our introductory discussion help to illustrate issues, concepts, and methods that are applicable in many areas of insurance.

Demand for Insurance

Basic Models of Coinsurance and Deductible Choice.

Mossin, (1968) and Smith (1968) proposed a simple model of insurance demand in which a risk averse decision maker has a total wealth (Y) equal to W - L where W is nonstochastic wealth and L is an insurable loss. To illustrate this model, first assume that the individual can buy coverage $\alpha L (0 \leq \alpha \leq 1)$ for a premium αP where α is the rate of insurance coverage (the coinsurance rate), $\lambda\ (\lambda \geq 1)$ is the premium loading factor, E(L) is the expected loss, and $P \equiv \lambda E(L)$. It can be shown that the optimal insurance coverage is such that $0 \leq \alpha^* \leq 1$ for $P \geq \overline{P} \geq E(L)$ where $\overline{P} \equiv \overline{\lambda}\, E(L)$ solves

$$E\left[U(Y + \alpha^*(L - \overline{\lambda}\, E(L))\right] = E\, U(Y)$$

and where U is a von Neumann-Morgenstern utility function $(U'(*) > 0, U''(*) < 0)$ and EU(Y) is the level of utility corresponding to no insurance. Hence, if the premium loading factor exceeds one but is less than $\bar{\lambda}$, partial coverage $(0 < \alpha^* < 1)$ is demanded.

When $\lambda = 1, \alpha^*$, is equal to one and the maximum premium that a risk averse individual is willing to pay over and above the actuarially fair value of full insurance is the Arrow-Pratt risk premium (Π^U) This premium solves

$$U(W - E(L) - \Pi^U) = EU(Y)$$

As shown by Pratt (1964), a more risk averse individual
378 with utility V such that $V = k(U), k' > 0$, and $k'' < 0$ will have a risk premium Π^V greater than Π^U.

Another important result in Mossin (1968) is that insurance coverage is an inferior good if the insured has decreasing absolute risk aversion. Under this assumption, there are two opposite effects on the demand for insurance when the loading factor (λ) increases : a negative substitution effect and a positive wealth effect. Hoy and Robson (1981) proposed an explicit theoretical condition under which insurance is a Giffen good for the class of constant relative risk aversion functions. More recently, Briys, Dionne, and Eeckhoudt (1989) generalized the Hoy and Robson (1981) analysis and provided a necessary and sufficient condition for insurance not to be a Giffen good. This condition bounds the variation of absolute risk aversion so that the wealth effect is always dominated by the substitution effect. Finally, Alarie, Dionne and Eeckhoudt (1990) present sufficient conditions to obtain the intuitive result that an insured will increase his demand for insurance when a mean preserving increase in risk is introduced in the initial loss distribution.

Another form of partial insurance is a policy with a deductible (Mossin. 1968; Gould 1969; Pashigian, Schkade, and Menefee, 1966; Schlesinger, 1981). For the above model, consider a general indemnity function $I(L)$ and premium $P = \lambda \int I(L)dF(L)$ where $\lambda(>1)$ is again a proportional loading factor. Then it can be shown under the constraint $I(L) \geq 0$ for all L, that for every P,

$$I^*(L) = \begin{cases} L - D^* \text{if } L - D^* \geq 0 \\ 0 \quad \text{if } L - D^* < 0 \end{cases}$$

where D^* is the optimal deductible. Since an insured bears some risk with the optimal contract it is reasonable to expect that a more risk averse insured would prefer a policy with a smaller deductible and higher premium. This result was proved by Schlesinger (1981) and Karni (1985). Moreover, under decreasing absolute risk aversion, $dD^* / dW > 0$ (Mossin, 1968). Also, it is possible to infer the degree of risk aversion of insurance buyers by observing their choices of deductibles (Drèze, 1981). The above results are generated under the assumption that the contract is free of default risk. With insolvency risk the above results do not in general hold but some qualitative results can be obtained with stronger utility assumptions (Doherty and Schlesinger, 1990).

Self-Insurance and Self-Protection.

Returning to the case of a single random variable L, market insurance can be analyzed in relation to other risk-mitigation activities. Ehrlich and Becker (1972) introduced the concepts of self-insurance and self-protection. Self-insurance refers to actions (y) that reduce the size (severity) of losses (i.e., $L'(y) < 0$ with $L''(y) > 0$) while self-protection refers to actions (x) that reduce the probability (frequency) of accidents ($p'(x) < 0$ with $p''(x) > 0$). Ehrlich and Becker gave conditions under which self-insurance and market insurance are substitutes and conditions under which self-protection and market insurance are complements. In both cases, self-protection and self-insurance activities were assumed to be observable by insurers.

While Ehrlich and Becker (1972) focused on the interaction between market insurance and activities involving either self-insurance or self-protection, they did not study in detail interactions between self-insurance and self-protection with and without the existence of market insurance. Boyer and Dionne (1983, 1989) and Chang and Ehrlich (1985) presented propositions concerning the choices among all three activities. When full insurance is not available, risk aversion affects the optimal choice of self-insurance and self-protection. While it seems intuitive that increased risk aversion should induce a risk averse decision maker to choose a higher level of both activities, Dionne and Eeckhoudt (1985) showed in a model with two states of the world that this is not always the case: more risk averse individuals may undertake less self-protection.

Corporate Demand for Insurance.

Portfolio decisions also have implications for the demand for insurance by corporations. When corporations are owned by shareholders who can reduce their investment risk at low cost through diversification of their own portfolios, risk aversion by owners is insufficient to generate demand for insurance. Specifically, if shareholders can costlessly eliminate the risk of corporate losses in their own portfolio's through portfolio diversification, the purchase of insurance by corporations can only increase shareholder wealth if it increases expected net cash flows by an amount that exceeds any loading in insurance premiums. Mayers and Smith (1982) analyzed the corporate demand for insurance from the perspective of modern finance theory (also see Main, 1982; Mayers and Smith, 1990, and MacMinn, 1990). They discussed how bankruptcy costs; risk aversion by managers, employees, customers, and suppliers; efficiencies in claims administration by insurers; and a number of other factors each can provide an incentive for the purchase of insurance even when shareholders can costlessly eliminate risk through portfolio diversification. In a later study, Mayers and Smith (1987) considered the possible ability of insurance to increase shareholder wealth by mitigating the underinvestment problem that was originally analyzed by Myers (1977).

Insurance and Resource Allocation

Allais (1953) and Arrow (1953) introduced general equilibrium models of resource allocation in the presence of uncertainty at a meeting on the subject in Paris during 1952. A year later, Debreu (1953) extended Arrow's (1953) contribution to a general framework of resource allocation under uncertainty. In this framework, physical goods are redefined as functions of states of the world and a consumption plan specifies the quantity of each good consumed in each state. Preferences among consumption plans reflect tastes, subjective beliefs about the likelihoods of states of the world, and attitudes towards risk. However, beliefs and attitudes towards risk do not affect producer behavior since for given contingent prices, there is no uncertainty about the present value of production plans. The existence of a competitive equilibrium that entails a Pareto optimal allocation of goods and services can be demonstrated for this economy.

Insurance markets can be viewed as markets for contingent goods. Borch (1962) proposed the first formal model of optimal insurance contracts. He presented a very elegant comparison between a general model of reinsurance and the Arrow-Debreu model with pure contingent goods and contingent prices for every state of the world. As noted earlier, Borch's insurance model can be reinterpreted in terms of standard insurance contracts. Two of his major contributions were to provide conditions for Pareto optimal exchange of risk and to show how risk aversion by insurers can explain partial coverage. Arrow (1963) used the same argument to introduce some element of coinsurance in optimal insurance contracts. Moreover, Arrow (1963) showed that if a risk neutral insurer offers a policy with a premium equal to the expected indemnity plus a proportional loading then the optimal contract provides full coverage of losses above a deductible. These forms of partial insurance limit the possibilities of risk shifting between economic agents (Arrow, 1965).

Raviv (1979) extended these results and showed that a Pareto optimal contract involves both a deductible and coinsurance of losses above the deductible. He also showed that the optimal contract does not have a deductible if the administrative cost of providing insurance does not depend on the amount of coverage. Coinsurance was explained either by insurer risk aversion or convexity of insurer costs. Conditions for an optimal contract with an upper limit of coverage also were presented. All these results were obtained under the constraint that coverage be nonnegative.

Kihlstrom and Roth (1982) studied the nature of negotiated insurance contracts in a non-competitive context in which there is bargaining over the amount and price of coverage. They showed that a risk neutral insurer obtains a higher expected income when bargaining against a more risk averse insured and that the competitive equilibrium allocation is not affected by the insured's risk aversion. Many of their results are represented in an Edgeworth Box diagram.

Moral Hazard

The concept of moral hazard was introduced in the economics literature by Arrow (1963), Drèze (1961), and Pauly (1968) (see also Kihlstrom and Pauly, 1971, and Spence and Zechauser, 1971). Two types of moral hazard have been defined according to the timing of an individual's actions in relation to the determination of the state of nature. They can be called ex ante and ex post moral hazard. In the first case the action is taken before the realization of the state of nature while in the second case the action is taken after.

Ex Ante Moral Hazard

Pauly (1974), Marshall (1976), and Shavell (1979) considered the case in which the occurrence of an accident (or the output of the consumption good) can be observed by the insurer and where neither the insured's actions nor the states of nature are observed. Under this structure of asymmetric information, the provision of insurance reduces (in general) the incentive to take care compared to the case of full information. Thus, there is a trade-off between risk sharing and incentives for care.

Moral hazard may alter the nature of competitive equilibrium by, for example, introducing nonconvexities in indifference curves. A competitive equilibrium may not exist, and when it does, insurance markets for some risks may fail to exist. More importantly, neither the first nor second theorems of welfare economics hold under moral hazard. Since market prices will not reflect social opportunity costs, theory suggests that governmental intervention in some insurance markets possibly could improve welfare if government has superior information (Arnott and Stiglitz, 1990; Arnott 1990).

Moral hazard also can affect standard analyses of government responses to externalities. An important example involves liability rules and compulsory insurance. With strict liability and risk averse victims and injurers, Shavell (1982) showed with perfect information that both first-party and liability insurance produce an efficient allocation of risk between parties in a model of unilateral accidents (with pecuniary losses only). When insurers cannot observe defendants care, moral hazard results in a trade-off between care and risk sharing (as in the case of first-party coverage). Shavell (1982) noted that if the government has no better information than insurers, its intervention in liability insurance does not improve welfare. This conclusion assumed that defendants were not judgment proof (i.e., they had sufficient assets to fully satisfy a judgment). Otherwise, their incentives to purchase liability insurance are reduced (Keeton and Kwerel, 1984; Shavell. 1986). Under strict liability, Shavell (1986) showed that if insurers cannot observe care, insureds buy partial insurance and the level of care is not optimal. He also showed that making liability insurance compulsory under these conditions need not restore efficient incentives. In fact, compulsory insurance could reduce care, and it is even possible that prohibiting insurance coverage could improve the level of care.

Ex Post Moral Hazard

The second type of moral hazard was first suggested by Spence and Zeckhauser (1971) who showed that an optimal contract between a principal and agent depends on the principal's ability to monitor the state of nature, the ex ante action taken by the agent, and the nature of the accident. The previous discussion of ex ante moral hazard assumed that the principal knew the nature of the accident. Marshall (1976a), Dionne (1984), and Townsend (1979) investigated the case in which the nature of an accident is not perfectly observable by the principal. Townsend (1979) considered the case in which the nature of the accident is known by the agent and verification is costly to the principal. One interpretation of such costly verification is auditing.

Adverse Selection

Adverse selection occurs in insurance markets when information is asymmetric, i.e., when the insurer cannot observe an individual's risk at the time policies are issued and the individual has superior information about his or her risk. Akerloff (1970) proposed that if insurers have imperfect information about differences in risk for prospective insureds, then some insurance markets may fail to exist and others may be inefficient. Studies have analysed the ability of partial insurance coverage, experience rating, and risk categorization to reduce the negative effects of adverse selection.

Partial Insurance and Sorting

Partial insurance coverage can result from two types of insurance pricing: "price only" policies (Pauly, 1974) and "price-quantity" policies (Rothschild and Stiglitz, 1976; Stiglitz, 1977). In the first case, insurers charge a uniform premium rate per unit of coverage to all applicants. Pauly's model ruled out price-quantity competition by assuming that insurers could not observe the total amount of coverage purchased by a client. In the second case, insurers offer a menu of policies with different prices and quantities so that different risks choose different insurance policies. These pricing strategies have been studied for single vs. multi-period contracts, for competition vs. monopoly, and, when assuming competition, for several different equilibrium concepts.

Dahlby (1983) provided some empirical evidence of adverse selection in the Canadian automobile insurance market.

He suggested that his empirical results were consistent with the WMS model with cross-subsidization between individuals in each class of risk. However, Riley (1983) argued that Dahlby's results were also consistent with Wilson's (1977) anticipatory equilibrium and Riley's (1979) reactive equilibrium. Cross-subsidization is not feasible in either of these models.

Experience Rating

Experience rating can be viewed as either a substitute or a complement to both risk categorization and sorting contracts with self-selection constraints when adverse selection is present. One polar case is when infinite length contracts yield the same solution as with full information. In this case, ex ante risk categorization is useless. The other polar case is when costless risk categorization permits full observation of an individual risk so that information on past experience is irrelevant. While experience rating, risk categorization, and sorting contracts are used simultaneously in most markets, economic analysis to date has considered the three mechanisms independently (see Dionne and Doherty, 1990, for a more detailed review).

Dionne (1983), Dionne and Lasserre (1985), and Cooper and Hayes (1987) extended Stiglitz's monopoly model (1977) to multi-period contracts. Dionne (1983) considered infinite length contracts without discounting while Cooper and Hayes (1987) mainly dealt with a finite horizon model (without discounting). While findings in both cases suggested that experience rating induced sorting or risk disclosure, the analyses differ in many respects. In Dionne (1983), a simple statistical review strategy is proposed along with risk announcement in the first period. The insurer offers a buyer full coverage at the full information price unless the observed average loss is greater than the true expected loss plus a statistical margin of error. Otherwise, full coverage is offered at a premium that includes a penalty. Both elements — announcement of risk and penalties — are necessary to obtain the same solution as with full information. They have the same role as the self-selection constraint and the premium adjustment mechanism of Cooper and Hayes (1987). In their model, the premium adjustment mechanism served to relax the self-selection constraints and to increase the monopolist's profits. Finally, in both articles the monopolist commits to the terms of the contract.

Risk Categorization

In most types of insurance, insurers classify risks using many variables. In auto insurance, for example, evidence indicates that driver age and sex are significantly related to accident probabilities (Dionne and Vanasse, 1988). In particular, evidence suggests that young male drivers (less than age 25) have much higher accident probabilities than the average driver. Since age and sex can be observed at very low cost, competition will force insurers to charge higher premiums to young males. Categorization using particular variables is prohibited in many markets, and the efficiency of categorization is an important policy issue.

Is statistical classification efficient in the presence of asymmetric information and adverse selection? Crocker and Snow (1985, 1986; also see Hoy, 1982, and Rea, 1987, 1990) showed that costless imperfect categorization always enhances efficiency when efficiency is defined as in Harris and Townsend (1981): second-best efficiency given the self-selection constraints imposed by asymmetric information. However, if classification is costly, the efficiency implications were ambiguous. Crocker and Snow (1986) also considered the existence of a balanced-budget tax-subsidy policy that provides private incentives to use risk categorization. With appropriate taxes, they showed that no agent would lose from classification. In their 1986 article, the results were shown using a WMS equilibrium, but a tax system also may sustain an efficient allocation with a Nash equilibrium. Their results can also be applied to a Wilson (1977) anticipatory equilibrium, or to a Riley (1979) reactive equilibrium (see Crocker and Snow, 1985). These results suggest that prohibiting statistical discrimination will impose efficiency losses in insurance markets when classification is virtually costless (e.g., age and sex classification in auto insurance).

Market Structure and Organizational Form

The seminal study by Joskow (1973) on market structure, conduct, and performance in the U.S. property-liability insurance industry considered market concentration and barriers to entry, estimated returns to scale, analyzed direct writer (exclusive agency/salaried employee) and independent agency (multiple insurer representation) distribution systems, and discussed possible effects of rate regulation on prices and availability of coverage. While written when rate regulation was predominant and when rating bureaus had a greater impact on the market than later in the 1970s and in the 1980s, this study nonetheless provided a basis for later work on a variety of subjects.

Concentration, Ease of Entry, and Consumer Search.

Joskow concluded that market concentration levels were low, especially for the national market, and that significant entry barriers did not exist. He estimated simple models of insurer operating expense ratios and concluded that the industry was characterized by constant returns to scale. He did find, however, that expense ratios were much lower for direct writers than for independent agency insurers. Cummins and VanDerhei (1979) estimated more elaborate models than those employed by Joskow using pooled cross-section and time-series data. Their results again indicated significantly lower expense ratios for direct writers, but they suggested increasing returns to scale throughout the range of output.

While the results of other studies that have estimated cost functions with cross-sectional accounting data also suggest increasing returns to scale (e.g., Doherty, 1981; Johnson, Flanigan, and Weisbart, 1981), the use of accounting data to infer returns to scale is problematic. Among other limitations, available data on insurance company operating expenses aggregate capital (e.g., product and market development) expenditures and current costs. Firm output also cannot be measured accurately. Appel, Worrall, and Butler (1985) analyzed changes in the size distribution of insurers over time. Their results were inconsistent with increasing returns for small insurers over time, Their results were inconsistent with increasing returns for small insurers and thus more in line with evidence on entry and levels of concentration.

Joskow argued that differences in operating costs between direct writers and independent agency insurers could not be explained by differences in service. In order to explain why direct writers had not grown more rapidly, he suggested that prior approval rate regulation had discouraged price cuts by direct writers, that difficulty in raising capital and obtaining consumer recognition slowed their expansion, and that it would be costly for independent agency insurers to become direct writers. As a result, he concluded that direct writers behaved as oligopolists subject to short-run capacity constraints and that constrained profit maximization involved selection of risks with lower than average expected claim costs.

Joskow also conjectured that costly consumer search for low prices impeded direct writer growth. Joskow and others (e.g., Kunreuther, Kleindorfer, and Pauly, 1983) have suggested that search for low prices is costly because of differences among insurers in risk selection criteria and because information provided by friends and neighbours that have different risk characteristics may convey little information. In an empirical analysis, Dahlby and West (1986) concluded that premium dispersion in Canadian auto insurance was consistent with a model of costly consumer search. This conclusion was contingent on their argument that risk classification could not account for premium variation. Berger, Kleindorfer, and Kunreuther (1989) modeled word of mouth transmission of price information in auto insurance in conjunction with consumer switch costs.

Returns to Scale and Underwriting Risk

The previously discussed studies of returns to scale and entry conditions focused primarily on insurer underwriting (risk selection), administrative, and commission expenses. Basic analysis of the relationship between insurer underwriting risk and scale of operations suggests that increasing returns to scale also could be associated with capital costs. If claim costs are not perfectly correlated across insured exposures, the standard deviation of an insurer's average claim cost will decline, ceteris paribus, as the number of insured exposures increases (e.g., Houston, 1964; Cummins, 1974; Venezian, 1983). If holding financial capital to reduce insurer default risk is costly (see below), this reduction in risk implies decreasing costs per insured exposure for any given probability of default because the required amount of capital per exposure will decline as the number of exposures increases. Low levels of market concentration and evidence on entry suggest that decreasing capital costs do not produce a large minimum efficient scale relative to market size. Underwriting risk declines at a decreasing rate with increases in scale, and the marginal reduction could be small relative to risk that cannot be reduced by writing more exposures (or by writing coverage in different lines of insurance).

Possible efficiency enhancing and anti-competitive aspects of institutional arrangements for pooling information among insurers have been analyzed in a number of studies (e.g., Danzon, 1983; Eisenach, 1985; also see Winter, 1988). Absent mechanisms for pooling data among insurers, claim cost forecasts might be expected to be more accurate for large firms due to their superior information. The costs of ratemaking and of complying with rate regulation also are likely to have a large fixed component. Hence, arrangements for pooling information and data analysis, some of which are made possible by the insurance industry's limited antitrust exemption under federal law and the laws of many states, are likely to reduce these costs and facilitate entry.

Alternative Organizational Forms.

In addition to significant variation in distribution methods, insurance markets generally are characterized by a variety of organizational forms. Most important, mutual organizations commonly have a significant market share. Mayers and Smith (1981) briefly considered the ability of alternative forms of insurance company ownership to minimize the cost of conflicts between owners, policyholders, and managers (also see Fama and Jensen, 1983). Mayers and Smith argued that while mutual organization eliminates owner-policyholder conflict, it can increase the cost of controlling manager-policyholder conflict compared to stock organization. They predicted that mutuals will specialize in lines of insurance where managers have limited discretion to pursue their own interests at the expense of policyholders.

Mayers and Smith (1988) provided further discussion of the ability of stock, mutual, and other organizational forms used in property-liability insurance to control conflict efficiently, and they developed and tested hypotheses concerning product specialization and geographic concentration across ownership types (also see Mayers and Smith, 1986). They obtained some evidence consistent with their predictions, including significant variation in product mix across ownership types. In other analysis, Hansmann (1985) provided detailed discussion of the possible role of mutual ownership in reducing conflicts between owners and policyholders over the level of insurer default risk (also see Garven, 1987). He also considered the possible ability of mutual ownership to facilitate risk selection during the formative years of U.S. insurance markets.

Insurance Pricing

Economic and financial analysis of insurance pricing has largely focused on two issues: (1) the determinants of long-run equilibrium prices in view of modern financial theory, and (2) the existence and possible causes of temporal volatility in insurance prices and in the availability of coverage that cannot be explained by changes in expected costs. Both areas have important policy implications.

Determinants of Long-run Equilibrium Prices.

Using the equilibrium risk-return relation implied by the Capital Asset Pricing Model (CAPM), Biger and Kahane (1978) showed that equilibrium insurance underwriting profit margins (and thus premiums) were a linear function of the riskless rate of interest and the systematic risk (beta) of underwriting in the absence of income taxes. They also provided estimates of underwriting betas using accounting data for different lines of insurance (also see Cummins and Harrington, 1985). Fairley (1979) (also see Hill, 1979, and Hill and Modigliani, 1986) developed a similar model and showed that with income taxes equilibrium premiums also increased with the tax rate and the amount of financial capital invested to support the sale of insurance.

Myers and Cohn (1986) criticized the ad hoc approach used by Fairley and others to apply the CAPM to contracts with multiperiod cash flows. They proposed a discounted cash flow model that would leave insurance company owners indifferent between selling policies and operating as an investment company. Key variables affecting equilibrium premiums again included tax rates on investment income, the amount of capital invested, and the required compensation to owners for risk bearing. Kraus and Ross (1982) considered application of arbitrage pricing theory to insurance pricing using both discrete and continuous time models.

The preceding studies either ignored default risk or implicitly assumed unlimited liability for insurance company owners. Doherty and Garven (1986) analyzed long-run equilibrium premiums with limited liability using discrete time options pricing theory under conditions in which stochastic investment returns and claim costs could be valued using risk neutral valuation functions. They used numerical examples to illustrate the effects of changing various parameters. Among other implications, premiums increased and default risk declined as invested capital increased. Cummins (1988) illustrated the application of continuous time options pricing theory to calculation of risk-based premiums for insurance guaranty funds. Again, numerical examples were used to illustrate the sensitivity of premiums to changes in underlying parameters.

An important implication of research on long-run equilibrium prices is that variability in claim costs that cannot be eliminated by insurer diversification raises prices (premium loadings) for any given level of default risk and thus reduces the gains from trade in insurance markets (also see Danzon, 1984, 1985; Doherty and Dionne, 1989). Hence, undiversifiable risk provides a possible explanation of why some risks may be uninsurable in addition to the effects of adverse selection, moral hazard, and insurer sales and administrative costs.

Price Volatility and Underwriting Cycles

Many lines of insurance appear to be characterized by "soft" markets, in which prices are stable or falling and coverage is readily available, followed by "hard" markets, in which prices rise rapidly and the number of insurers offering coverage for some types of risk declines substantially. Popular wisdom holds that soft and hard markets occur cyclically with a period of about six years. Several studies have provided empirical evidence that reported underwriting and total operating profit margins follow a second-order autoregressive process that is consistent with a cycle (Venezian, 1985; Cummins and Outreville, 1987; Doherty and Kang, 1988; also see Smith, 1989). Interest in this area was stimulated by the liability insurance "crisis" of the mid-1980s, which was characterized by dramatic increases in premiums for many commercial liability risks and by reductions in the availability of coverage.

The traditional view of underwriting cycles by insurance industry analysts emphasizes fluctuations in capacity to write coverage. According to this view, which assumes an inelastic supply of capital, competition drives prices down until capital is depleted, insurers ultimately constrain supply in order to prevent default, and attendant increases in prices and retained earnings then replenish capital until price-cutting ensues again. Berger (1988) presented a simple model of this scenario that assumed that insurers were unable to add new capital and that pricing decisions were based on beginning of period surplus.

Several studies have questioned the existence of true cycles in prices. Cummins and Outreville (1987) considered whether cycles in reported underwriting results could simply reflect insurer financial reporting procedures in conjunction with information, policy renewal, and regulatory lags. They also provided evidence that reported operating margins follow a cyclical process for many lines of insurance in the United States and other countries. Doherty and Kang (1988) essentially argued that cycles in insurer operating results reflected slow adjustment of premiums to changes in the present value of expected future costs. However, the causes of slow adjustment and the influence of slow adjustment versus charges in costs were not clear in their analysis.

Harrington (1988) analyzed industry financial results surrounding the liability insurance crisis of the mid-1980s and discussed possible causes of the crisis including cyclical effects. This study also provided evidence that rapid premium growth in general liability insurance was associated with upward revisions in insurer loss reserves for prior years' business and rapid growth in reported losses for new business. The results suggested that much of the total growth in premiums during 1980-86 could be explained by growth in expected losses and changes in interest rates (i.e., by determinants of long-run equilibrium premiums). However, premiums grew slower than discounted reported losses during the early 1980s and faster than discounted reported losses during 1985-86, a result that is consistent with cyclical effects.

McGee (1986) suggested that heterogeneous expectations of future claim costs among insurers could lead to price-cutting that characterizes soft markets. Harrington (1988) questioned whether aggressive behavior by firms with little to lose in the event of default could lead to excessive price-cutting. Winter (1988 and 1989) developed a model in which undiversifiable risk and constraints on external capital flows (such as those that might arise from asymmetric information between insurer managers and investors or from income tax treatment of shareholder dividends) and solvency (which could be imposed by regulators or reflect policyholder preferences) could lead to periods of soft markets followed by sharp increases in prices. His model predicts a negative relation between price and capital. He reported (1989) some evidence consistent with this prediction

using aggregate industry data prior to the crisis of 1985-86, at which time the relationship became positive.

Volatility in the commercial liability insurance market during the 1980s has led to a number of recent working papers (several of which only contain preliminary analysis and results). Subjects covered include insurer responses to exogenous shocks to capital (Gron, 1989; Cummins and Danzon, 1990), the sensitivity of premiums to interest rates (Doherty and Garven, 1990), the possible effects of regulation (Winter, 1988a; Tennyson, 1989), and possible causes of price-cutting in soft markets (Harrington and Danzon, 1990).

Insurance Regulation

Most economic analyses of regulation of insurance markets have focused on solvency regulation and regulation of premium rates and the availability of coverage. Theoretical work has had both positive and normative aspects. Most empirical work has focused on estimating the effects of regulation.

Default Risk and Solvency Regulation

Solvency regulation in the United States has three major facets: (1) direct controls on certain activities and financial reporting, (2) monitoring of insurer behavior, and (3) a system for paying claims of insolvent insurers (see Harrington and Danzon, 1986, for details). Direct controls include minimum capital requirements and limitations on investment activities. The principal monitoring system is administered by the National Association of Insurance Commissioners. Guaranty funds exist to pay claims of insolvent property-liability insurers in all states; many states have similar arrangements for other types of insurance. The traditional rationale for solvency regulation is that consumers are unable to monitor the risk of insurer default.

Actuarial literature (see Kastelijn and Remmerswaal, 1986, for a survey) has analyzed default risk as a function of various operating and financial decisions or analyzed decisions necessary to achieve a given probability of default (which generally is presumed to be chosen by regulators or management). Portfolio models of property-liability insurance company behavior (e.g., Michaelsen and Goshay, 1967; Kahane and Nye, 1975; Hammond and Shilling, 1978) have either treated default risk as exogenously determined or subject to insurer choice. Economic factors that could influence this choice have not been the focus of this literature.

More recently, economic analysis of insurer default risk has focused on factors that influence insurer capital decisions under default risk. Building on the work of Borch (e.g., Borch, 1982; also see DeFinetti, 1957), Munch and Smallwood (1982) and Finsinger and Pauly (1984) model insurer default risk assuming that insurers maximize value to shareholders, that demand is inelastic with respect to default risk, and that investing financial capital to support insurance operations is costly. The principal implication is that optimal capital is positively related to the amount of loss that shareholders would suffer if claim costs were to exceed the firm's financial assets. Munch and Smallwood (1982) considered possible loss of goodwill in the event of default; Finsinger and Pauly (1984) assumed that an entry cost would be forfeited that otherwise would allow the firm to continue operating (also see Tapiero, Zuckerman, and Kahane, 1978). If shareholders have nothing to lose, they will not commit any capital. If they are exposed to loss, and if it is assumed that firms cannot add capital after claims are realized, firms will commit some capital ex ante.

In an empirical analysis of the effects of solvency regulation using cross-state data, Munch and Smallwood (1980) estimated the impact of minimum capital requirements and other forms of solvency regulation on the number of insurers selling coverage and the number of insolvencies. While subject to significant data limitations, their results provided some evidence that minimum capital requirements reduced insolvencies by reducing the number of small domestic insurers in the market. They also compared characteristics of solvent and insolvent firms and concluded that the results were consistent with selection of default risk to maximize firm value.

Other empirical studies generally have focused on predicting insurer defaults using financial data without closely relating the variables chosen to the theory of default risk (e.g., Pinches and Trieschmann, 1973; Harrington and Nelson, 1986; McDonald, 1988). Not much is presently known about the magnitude of the effects of regulatory monitoring and guaranty funds on default risk.

Rate Regulation

Regulation of rates, which is used primarily in property-liability insurance, can affect an insurer's average rate level or overall percentage change in its rates during a given period. It also can affect rate differentials between groups of consumers by imposing limits on voluntary or involuntary market rates for particular groups or by restricting risk classification.

Voluntary market rates for most U.S. property-liability lines presently are subject to prior approval regulation in about half of the states. Most states had prior approval regulation during the 1950s and 1960s, and rate regulation was likely to have encouraged insurers to use rates developed by rating bureaus (Joskow, 1973; Harrington, 1984; also see Danzon, 1983). A trend towards deregulation began in the late 1960s and continued until the early 1980s. A number of states reregulated commercial liability insurance rates following the liability insurance crisis of 1985-86. California adopted prior approval regulation for property-liability insurance with the enactment of Proposition 103 in 1988. Several additional states either have reenacted or are considering reenactment of prior approval regulation.

Most studies of rate regulation have estimated the impact of voluntary market rate regulation in auto insurance on average rate levels for the overall (voluntary and involuntary) market. Major hypotheses have been that regulation raises rates due to capture by industry, that regulation has short-run effects due to regulatory lag, and that regulation persistently reduces rates due to consumer pressure (see Harrington, 1984). Most studies have regressed either the statewide ratio of premiums to losses (or of losses to premiums) on a rate regulation dummy variable and on a variety of control variables. Harrington (1987) used this procedure and maximum likelihood estimation to provide evidence of cross-state variation in the impact of regulation. The results of this and other studies using data from the late 1970s and early 1980s (e.g., Pauly, Kleindorfer, and Kunreuther, 1986; Grabowski, Viscusi, and Evans, 1989) suggested that on average prior approval regulation reduced the ratio of premiums to losses.

Some evidence of variation in the impact of prior approval regulation across states was provided in Harrington (1987) and several other studies, but causes of such variation generally were not addressed. A large amount of anecdotal evidence suggests that substantial regulatory intervention in insurance pricing tends to occur in states where the unregulated cost of coverage would be relatively high, that regulation favors high risk groups, and that exits eventually have occurred in response to restrictive regulation. Pauly, Kleindorfer, and Kunreuther (1986) provided evidence that direct writer market share was significantly lower in states with prior approval regulation. Building on the work of Ippolito (1979), they also provided evidence that restrictive rate regulation was associated with lower operating expenses (and presumably lower quality; also see Braeutigam and Pauly, 1986).

Involuntary markets in auto insurance have been found to be significantly larger in states with prior approval regulation of voluntary market rates (e.g., Ippolito, 1979; Grabowski, Viscusi, and Evans, 1989). Involuntary market rate regulation and state restrictions on risk classification (e.g, unisex rating rules) also will affect involuntary market size (as was implied by Joskow, 1973). The relative effects of these influences and of voluntary market rate regulation would be difficult to sort out. Voluntary and involuntary market regulation of auto liability insurance rates could reduce the number of uninsured drivers by lowering rates to drivers who otherwise would fail to buy coverage (see Kunreuther, Kleindorfer, and Pauly, 1983; Keeton and Kwerel, 1984). If so, the efficiency loss that otherwise would be expected from rate regulation would be mitigated. Not much is known about the magnitude of these effects or the effects of insurance rate regulation on decisions to drive and the frequency and severity of accidents.