Implications of Uncollectibles for Hospitalization Coinsurance Rates

Catherine S. Elliott

ABSTRACT

The aim of this article is to study the effects of hospital care uncollectibles on optimal coinsurance rates for group health plans. A model isolates the influence of uncollectibles from the influence of risk aversion. It is found that if hospitals are willing to negotiate price concessions with payer groups, in return for lower patient-copayment responsibilities, a representative risk neutral consumer may purchase actuarially fair insurance. In addition, because the benefits from lower bad debts and any risk pooling may outweigh the costs of moral hazard, the optimal insurance contract may specify full insurance. The theory helps to explain price concessions negotiated by numerous payer groups, such as health maintenance and preferred provider organizations, the contractual allowances given to insurers such as Blue Cross, the attraction of complete prepayment, and non-random health care coverage.

Hospital care uncollectibles represent a breakdown in the legal contract system. Several distinctive factors tend to foster uncollectibles in this market. First, the product for the most part is intangible. Once delivered, health care cannot be repossessed, and thus, the most effective safeguard against payment default is absent.

Second, unless convinced that patient treatment was entirely satisfactory, many lawyers advise against filing suit for unpaid bills or even turning bills over to collection agencies (see, for example, Alton, 1977, pp. 44-6; and Holder, 1975, pp. 319, 408). It is believed that the incidence of medical malpractice lawsuits may increase when health care providers institute stringent bill collection procedures.

Third, defaults on hospital bills may have little or no effect on consumer credit ratings. Gunther (1982) says that hospitals rarely report payments or non-payments to credit bureaus. If potential creditors are unaware of whether

---

Catherine S. Elliott is Assistant Professor of Economics at New College of the University of South Florida.

The author thanks John Cross, Paul Courant, and Theodore Bergstrom, Don Hayward, and Shirley Tewes. In addition, two anonymous referees are very much appreciated. This work was supported, in part, by the University of South Florida Research and Creative Scholarship Grant Program under Grant No. 36-24-930-R5.

medical debts have been paid, the ability of the patient to receive credit for other purchases is unimpaired.

Fourth, the failure to pay current debts may not affect a consumer's ability to obtain future hospital care. Hospitalization episodes are infrequent. In addition, community hospitals do not systematically compare billing records.

Finally, once started, treatment cannot be discontinued easily, even if it becomes apparent the patient will fail to pay the bill. If discharged before medically justified, patients have legal grounds to sue both hospitals and physicians for abandonment (see Holder, 1975, pp. 381, 386; and Aaron and Schwartz, 1990, p. 422).

By reducing direct-patient obligations for medical bills, public and private group insurance plans place limits on the magnitude of hospital care uncollectibles. Direct-patient bad debts are constrained by the high percentage of total charges paid by Medicare, Medicaid, Blue Cross, and commercial insurers. According to Aaron and Schwartz (1990, p. 418), over the past two decades, direct-patient obligations averaged just 10 percent of total hospital charges. In the only nationwide-statistical analysis of the determinants of uncompensated hospital care (which includes charity care as well as uncollectibles), Sloan, Valvona, and Mullner (1986) found a statistically significant and positive correlation between uncollectibles and patient-copayment responsibilities (portion of charges paid directly by patient).

Traditionally, analysis of health insurance has focused on the over-consumption of health care caused by full insurance or complete prepayment. Blue Cross and Medicare, historically the two major providers of full hospitalization insurance, have been increasing patient-copayment responsibilities in an attempt to reduce the overconsumption resulting from moral hazard – defined as the additional consumption of health care caused by prices that are lower than the marginal cost of care. Gibson, Waldo, and Levit (1983) reported that, although relatively constant in the previous decade, the percentage of third party coverage of hospitalization charges has been declining since 1975. However, the results in this article suggest that if direct-patient bad debts and bill collection costs are included in the analysis, it is no longer certain that coinsurance is optimal. Although moral hazard will exist with complete prepayment, the gains from eliminating uncollectibles may outweigh the costs of overconsumption.

Though nationwide data are not available, the existing figures suggest that the percentage of direct-patient obligations currently uncollectible is substantial. The Center for Health Policy Studies (1984) reported that, in 30 Massachusetts hospitals, defaults on insured patient coinsurance obligations accounted for 40 percent of 1982 bad debts. Again excluding charity care, an average of 20 percent of direct patient obligations at the University Hospital in Ann Arbor, Michigan was uncollectible (Biggs, 1981). Finally, for 1978 to 1980, Sloan, Valvona, and Mullner (1986) found that 28 percent of all uninsured patient obligations remained uncollected.

The remainder of this article is organized as follows. First, the article develops a model that isolates the influence of uncollectibles from the influence of risk aversion on a representative consumer's choice of an optimal coinsurance rate for a group hospitalization plan. Second, it investigates the characteristics of the optimal choice for risk neutral and risk averse consumers, and completes a comparative-static analysis of shifts in the behavioral relationships developed in the previous section. The third section argues that the role played by uncollectibles may explain other features of the health insurance market, including contractual allowances given to insurers such as Blue Cross and coverage of planned health care expenditures such as routine dental work and annual physicals. A summary and conclusions are presented at the end of the article.

The Model

The theoretical model is an extension of the standard model of insurance demand by a representative consumer.[1] It has been designed to analyze the effects of three factors on the choice of an optimal insurance contract – bad debts, risk aversion, and moral hazard. Specifically, the author wished: 1) to incorporate the influence of bad debts and subsequent hospital collection efforts on the optimal contract, 2) to provide for a separate analysis of the effects of bad debts from the effects of risk averse preferences, and 3) to account for the costs of moral hazard.

The first purpose is achieved by focusing on a representative consumer's choice of a group hospitalization plan, given the consumer's decision to pay or not to pay health care bills. The standard economic model of crime predicts that anyone can choose to default or delay payment on health care bills if the expected gains outweigh the expected costs. The question considered, then, is what level of coinsurance will be chosen given the effect of the default decision on hospital charges, insurance premiums, and demand for health care.

The remaining two purposes of the theoretical framework are achieved by use of the proposed specification of the expected utility function, introduced below. The moral hazard incorporated into the model is the type considered by Pauly (1968) in his classic comment on moral hazard, i.e., over-consumption based on the ability to change the amount of loss as opposed to the probability of hazard. The specification of the expected utility function is discussed next; the incorporation of bad debts into the model is presented in the following subsection.

Specification of the Expected Utility Function

Suppose consumers choose from among a variety of group hospitalization insurance plans. Each plan is differentiated solely by a coinsurance rate and an associated premium. Once a plan is chosen, premiums are deducted from individual paychecks, and thus, payment delays and defaults are associated

[1] See Raviv (1979) for a review of insurance models and Arrow (1976) for an example of the use of a representative consumer model.

only with coinsurance obligations. Assume further that all members receive care at the same community hospital.

A representative consumer, presented with group insurance plans offering a continuous range of coinsurance rates, chooses a plan that maximizes expected utility with respect to generalized hospital care in each state, x_i, and residual income for all other goods in each state, y_i. Let the von Neumann-Morgenstern expected utility function take the following form:

$$EU = \Sigma q_i W(y_i) + \Sigma q_i U_i(x_i), \tag{1}$$

where: $i \equiv 1, \ldots, n$ states of health, and the notation Σ denotes summation from $i = 1$ to n;

$q_i \equiv$ probability state of health i occurs, where $\Sigma q_i = 1$;

$x_i \equiv$ composite of health care consumption in state of health i;

$\Sigma q_i x_i \equiv$ total expected consumption of health care;

$y_i = I - r - (1-k)(1-s) \cdot px_i \equiv$ residual income;

$I \equiv$ gross income or wealth, constant over $i = 1, \ldots, n$;

$r \equiv$ insurance premium;

$k \equiv$ default rate on total expected patient-coinsurance obligations, such that $k \in [0,1]$;

$s \equiv$ insurance sharing rate (fraction of total health care bill reimbursed by insurer), such that $s \in [0,1]$;

$p \equiv$ total charge per unit of health care.

The probability distribution of q_i is assumed known to consumers as well as to health care and insurance providers, and fixed over $i = 1, \ldots, n$. Treatment consumption is determined by the state of health but is not assumed illness-specific nor fixed for each state. Each x_i is continuous and represents a composite of all treatment and diagnostic services available in a hospital setting. Therefore, treatments in each state of health are allowed to differ by continuous changes in the quantity of generalized health care consumed, but not in the composition of x_i. The consumer's legal coinsurance responsibility is $(1-s)$. An insurance contract without a provision for coinsurance is represented by a sharing rate such that $s = 1$, while no insurance is equivalent to $s = 0$.

The functional form for expected utility in expression (1) is state-dependent, multivariate, and separable in net income and health care. It has been chosen because of the implications of these three features.

State-dependent utility is a necessary assumption because preferences can be permanently altered when dramatic changes in health status prevent the return to an original state of health.[2] To incorporate the possibility of moral hazard, the function is multivariate in x_i and y_i. Because preferences determine health care expenditures, treatment must be specified as a separate decision variable from the composite good representing all remaining expenditures.

[2] See Viscusi and Evans (1990) for a summary and an empirical investigation of the use of the state-dependence assumption in health care modeling. However, the functional specification in their work is univariate.

Separability is assumed so that it is possible to specify definitions of risk preferences that are consistent with those derived under the more familiar assumptions of state-independent, univariate expected utility, i.e., the full-insurance definition of risk aversion and the constant marginal utility of income definition of risk neutrality.

Within the state-dependent, multivariate utility framework, risk aversion previously has been specified as the concavity of each state-specific utility function. But as Cook and Graham (1977) have shown, this definition of risk aversion can be consistent with less than full insurance against income. In addition, functional specifications of risk neutrality have been avoided in this framework.

A major purpose of this article, however, is to study the effects of uncollectibles on insurance demand separately from the effects of risk aversion. This requires investigating the insurance demand of a risk neutral consumer. Therefore, separability is assumed because it guarantees that: 1) risk neutrality can be defined as constant marginal utility of net income, and 2) risk aversion, specified as the strict concavity of the expected utility function in net income, y_i, is equivalent to equating risk aversion with a consumer's desire to equalize net income across states.[3]

Incorporation of Uncollectibles into the Model

The analysis focuses on a representative consumer's choice of an optimal coinsurance rate given a reduced form default-rate function. The default-rate function is analogous to the "supply of offenses" function first postulated by Becker (1968), and is consistent with most economic models of criminal choice developed since his seminal article.[4]

The default-rate function is represented as:

$$k = k(b,s), \tag{2}$$

where b ≡ per dollar bad-debt collection effort by the hospital, as a percent of total expected patient-coinsurance obligations, such that $b \in [0,1]$. It is assumed that k represents an expected delay or default on patient-coinsurance

[3] The proof is contained in Elliott (1987).

[4] See Heineke (1978) for a critical presentation of two predominant approaches to modeling criminal behaviour. However, the default choice of the hospital patient in this article differs from both approaches. First, unlike crime in the labor allocation models of Becker (1968), Ehrlich (1973), and Block and Heineke (1975), defaulting is not a time-constrained activity. Second, while the model shares a similar focus to the income tax evasion studies of Allingham and Sandmo (1972), Kolm (1973), and Singh (1973) — the decision not to pay a financial obligation, the consumer in the current model faces health risks and maximizes utility from health care consumption as well as from wealth.

The current model is also an extension of work on optimal theft protection, as both the structure of the insurance contract and the hospital's collection efforts affect the degree of theft (through the default rate). See Andreano and Siegfried (1980) for a review of the literature in this area, as well as of economic studies of other aspects of crime. More recently, Cornett (1988) published an article in this journal exploring the ability of financial institutions to safeguard wealth.

obligations over all states of health. However, it is possible that the consumer delays payments or defaults in one state of health, yet pays on time in another state of health.

The default-rate function enters the consumer's expected utility function as part of residual income, y_i. While an unpaid bill is a wealth transfer, not a social cost, the person who gets away with not paying a bill is better off because of it. The bill collection efforts are deadweight losses to society.

The default rate is a function of a default-cost variable — the hospital's collection efforts, and a default-benefit variable — the insurance sharing rate. Assume collection efforts are effective: $k_b(b,s) < 0$, but diminishing in effectiveness as the default rate tends toward zero: $k_{bb}(b,s) > 0$. These assumptions are consistent with evidence that punishment and police efforts are effective deterrents of crime. However, collection efforts do not represent imprisonment or fines as in other examples of punishment found in the literature on the economics of crime. Collection efforts are extra billings, phone calls, letters, and perhaps, threats the hospital might use to induce payment.[5] Even if all coinsurance obligations are collected eventually, collection costs will be positive if the default rate is positive.

Given a fixed level of collection effort by the hospital, it is assumed that the lower the consumer's total expected coinsurance obligation, the lower the average rate of default. This assumption is based on another well-accepted result — a decrease in the benefits or gain from crime will decrease the incidence of crime.

The benefit to the consumer of defaulting on the coinsurance obligation is an increase in residual income, which then can be spent on all other consumption activities. Thus, if the coinsurance rate, which is the price the consumer is legally obligated to pay for health care at the point of purchase, increases, the personal benefit of default also increases. The greater the coinsurance obligation for a given amount of health care consumption, the greater the residual income from defaulting on the bill. This assumption is expressed mathematically as $k_s(b,s) < 0$.

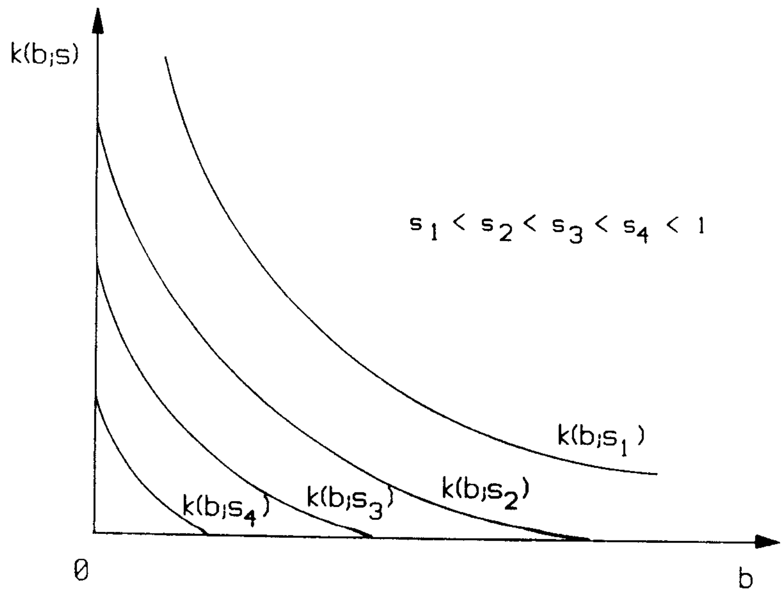
The relationship $k(b,s)$ can be represented as a system of non-intersecting, patient default-response functions: $k(b)$. Each separate function, $k(b)$, is generated by the consumer's response to a change in bill collection efforts by the hospital. By varying the coinsurance rate, such that $s \in [0,1)$, the complete system of curves is traced out. A representative system is drawn in Figure 1, where each curve in the system is defined by a different value of s. An increase in the sharing rate, i.e., a decrease in coinsurance responsibilities, shifts the default-response function toward the origin. However, if $s = 1$, then $k(b,s)|_{s=1} \equiv 0$.

---

[5] The costs of collecting insurance payments are ignored, although the collection efforts defined above can be considered to be in excess of those costs.

Figure 1

System of Default-Response Functions



Maximization of the Expected Utility Function

The consumer's choice of s is constrained by the pricing behaviors of the hospital care provider and the insurance firm. For analytical convenience, the consumer's decision process is modeled as taking place in two stages. First, utility-maximizing quantities of care within each state of health are chosen, contingent upon a given sharing rate, s, and the resulting default rate, k, total per unit charge for health care, p, and insurance premium, r. Second, the optimal coinsurance rate is found by maximizing the consumer's indirect expected utility function, which is shown to be a univariate function of the sharing rate.

The indirect expected utility function is used at this point because the sharing rate represents the consumer's price for hospital care, at the point of purchase. In addition, use of the indirect utility function allows direct differentiation with respect to the sharing rate. Both the interpretation of first order conditions and the comparative statics are more straightforward.

To determine treatment demand in each state of health, substitute the budget constraint into the expected utility function in expression (1) and differentiate with respect to each x_i. For all functions below, the notation $f'(\cdot)$ denotes a first order derivative, and $f''(\cdot)$ denotes one of second order. After rearranging, the first order conditions can be written as:

$$U'_i(x_i)/W'(y_i) = (1-k)(1-s)p, \quad i = 1, \ldots, n. \tag{3}$$

Each first order condition in (3) is state-specific; thus, demand for care in each state can be written separately as a function of the premium and the coinsurance obligation:

$$f_i(r,(1-s)p), \quad i = 1, \ldots, n. \tag{4}$$

The Hospital

Evidence suggests that hospitals, nonprofit or for-profit, cannot be considered either least-cost competitive firms or profit-maximizing monopolies.[6] Price concessions between hospitals and payer groups are becoming increasingly common. Of course, they may be involuntary: for example, the legislated changes in Medicare payment systems or all-payer rate regulation systems. Yet, significant price concessions also have been the result of voluntary negotiation between hospitals and payer groups such as insurance carriers, employer self-insurance plans, selective contracting plans, preferred provider organizations, and health maintenance organizations.

Discussing the incentive to negotiate, Cohodes argues: "The buyer gains, but so does the hospital, in that it receives the assurance of maintaining a certain volume of cases, or the prompt payment and handling of claims, or some other consideration" (1986, p. 266). Similarly, Ohsfeldt notes: "Some new payment arrangements, such as preferred provider organizations (PPOs), may reflect a willingness on the part of physicians to accept lower fees for their services in exchange for an increased probability of payment" (1985, p. 1338).

The following model of hospital behavior is designed to be consistent with both noncompetitive pricing behavior and negotiations between hospitals and payers. Therefore, it is assumed that the hospital negotiates price concessions with a payer group in return for a decrease in bad debts and collection costs, which is expected to accompany an increase in insurance coverage (and a decrease in direct-patient copayment responsibilities). A payer group is defined as any group that negotiates an insurance or payment contract with the hospital, and one that the hospital treats as a separate market segment.[7]

The purpose of this model is to explain the effect of price concessions by the hospital on the insurance purchasing decision of a representative consumer, with a given default-response function. The hospital chooses a level of bad-debt collection effort, b, in order to maximize net income from collection efforts, and a total per unit charge, p, such that a goal level of excess revenues is reached. The goal level of excess revenues is a value determined by the

[6] See Cohodes (1986) for a general discussion of cost inefficiencies, Feldman and Dowd (1986) for empirical support of the non- competitive nature of hospital pricing, and Schlesinger, Marmor, and Smithey (1987) for a review and analysis of nonprofit and for-profit health care providers. In addition, two health care journals recently published the following special issues: "Competition in the Health Care Sector: Ten Years Later" in *Journal of Health Politics, Policy and Law* (13: Summer 1988), and "The Managed Care Revolution" in *Health Affairs* (7: Summer 1988).

[7] See Feldman and Dowd (1986) for evidence of hospital price discrimination.

hospital management, and is used as a target in the process of negotiating prices or charges with payer groups.

The negotiation model is then straightforward. If the payer group offers an insurance plan that will lower bad debts and collection costs, thus increasing net income from collection efforts, some or all of the savings is passed onto the consumers as a price concession — a decrease in the total per unit charge for hospital care, p.

The goal level of excess revenues acts just like a binding-profit constraint for a revenue maximizing firm.[8] Later in this article, comparative static analysis will examine the effect on the results of the model of a change in the hospital's excess-revenue goal.

Goal Level of Excess Revenues: Let π^* represent the hospital's goal level of excess revenues. Then, the hospital is assumed to negotiate p so as to satisfy:

$$\pi^* = sp\overline{Q} + [1 - k(b^*,s)]\cdot(1-s)p\overline{Q} - C\overline{Q} - b^*\cdot(1-s)p\overline{Q}, \tag{5}$$

where $C \equiv$ constant per unit cost of health care, $\overline{Q} \equiv$ representative consumer's total expected demand for health care, and $b^* \equiv$ net-income maximizing level of collection efforts. The total charge per unit of care, p, is not equal to the cost of health care, but is a reflection of the hospital management's goals and negotiation skills. The marginal cost of health care, C, is assumed not to include provision for bad debts, profit or operating margin, or charity care.

In expression (5), total revenues are the sum of the first two terms: insurance reimbursements plus actual patient-coinsurance payments. Total costs are the sum of the remaining two terms: total expected health care costs plus costs due to collection efforts. For a given coinsurance obligation, an increase in collection efforts affects the profit equation in two ways: an increase in revenues as a result of a decrease in the default rate (the second term), and an increase in costs (the last term).

Optimal Level of Collection Efforts: Consider the optimal solution for collection efforts, b^*. For the hospital, the insurance sharing rate is a parameter. Therefore, simplify notation by using $k(b) \equiv k(b,s)$, i.e., the notation for a single default-response function. The necessary and sufficient conditions for an optimal value b^* are found by differentiating expression (5) with respect to b. Letting π^* denote the relationship in expression (5), the first order condition for an extremum is:

$$d\pi^*/db = -k'(b)\cdot(1-s)p\overline{Q} - (1-s)p\overline{Q} = 0; \tag{6}$$

and second order condition for a maximum is:

$$d^2\pi^*/db^2 = -k''(b)\cdot(1-s)p\overline{Q} < 0. \tag{7}$$

After canceling, and rearranging terms in expressions (6) and (7), these conditions become:

[8] Thus, this model is consistent also with the frequent assumption that nonprofit organizations maximize excess revenues.

$$k'(b^*) = -1, \tag{8}$$

$$k''(b^*) > 0. \tag{9}$$

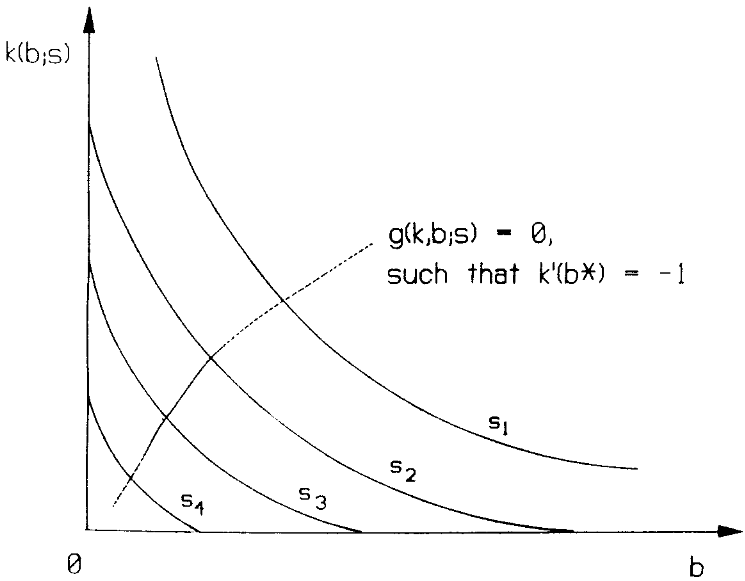
The second order condition in (9) is satisfied by previous assumption. Expression (8) implies that, for a given insurance sharing rate, the optimal level of collection efforts should be chosen so that the marginal bill collection effort equals the marginal increase in paid-up bills due to a lower value of k. In addition, it implies that the choice of an optimal level of collection efforts, b*, is independent of the choice of p.

Because each default-response function, k(b), is negatively sloped and monotonically increasing, expression (8) also implies that unique maximum values b* and k* exist for each value of s. Further, because k(b,s) represents a continuous system of non-intersecting curves, an "optimal collection-effort" expansion path can be traced out, as s varies from one to zero. An example of such an expansion path is drawn in Figure 2. Let the expansion path be represented by the implicit function:

$$g(k,b;s) = 0. \tag{10}$$

Figure 2

Optimal Collection-Effort Expansion Path



Thus, expression (10) is a complete description of each optimal pair (k^*,b^*) corresponding to each value of s.[9]

The expansion path is assumed to be positively sloped throughout: $dk^*/db^* > 0$. This assumption implies that increasing the sharing rate is an effective way of decreasing the optimal levels of both the default rate and the collection effort. It also implies that, for a given level of collection efforts, the marginal propensity to default will increase if the patient is obligated to pay a greater percentage of charges, i.e., if the sharing rate decreases.

The assumption $dk^*/db^* > 0$ allows the expansion path in (10) to be represented by the following two univariate functions:

$$k = k(s), \quad \text{where } k'(s) < 0 \tag{11}$$

$$\text{and } b = b(s), \quad \text{where } b'(s) < 0.\text{[10]} \tag{12}$$

The functions in expressions (11) and (12) will be used below to investigate the representative consumer's choice of an optimal insurance contract.

The Insurer

The possibility of premium loading is ignored initially. The consumer pays an actuarially fair premium, r, equal to the expected insurance outlay:

$$r = sp\Sigma q_i x_i. \tag{13}$$

Maximization of the Indirect Expected Utility Function

Because the choice of an optimal sharing rate is constrained by the pricing behaviors of the hospital care provider and the insurance firm, a simultaneous system of n + 2 equations must be solved in order to determine the optimal coinsurance rate: n treatment demand functions, represented by expression (4), plus equations (5) and (13). Appendix A establishes conditions under which the total expected demand for care, the insurance premium, and the hospital charge can be expressed as univariate, monotonic functions of the sharing rate — thus, yielding unique equilibrium solutions to the simultaneous equation system. Essentially, in addition to (11) and (12), the conditions are:

i) $\partial f_i(r,(1-s)p)/\partial(1-s)p < 0$, $i = 1, \ldots, n$;
ii) $r'(s) > 0$;
iii) $d\pi^*/dp|_{b=b^*} > 0$;
and iv) $\pi^* > 0$.

The first requirement states that if the consumer's price at the point of purchase, $(1-s)p$, increases, demand for care within each state of health decreases. The second requirement is necessary for uniqueness. The third

[9]See Henderson and Quandt (1980, pp. 78, 88-9) for a discussion of implicit functions and expansion paths.

[10]To simplify notation, the asterisks have been dropped. Throughout the article, k(s) and b(s) will denote the optimal values k^* and b^* along the expansion path traced out by $g(k^*,b^*;s)$.

requirement is equivalent to assuming that the hospital's behavior is consistent with constrained-revenue maximization, and the fourth, that the binding profit-constraint, π^*, is positive. An important implication of requirements iii) and iv) is that they allow writing a price-concession function, p(s), such that $p'(s) < 0$.

The demand for health care in each state of health now can be rewritten as a function of a single variable, the sharing rate:

$$g_i(s) = f_i(r(s),(1-s)p(s)), \quad i = 1, \ldots, n. \tag{14}$$

The consumer's choice of an optimal sharing rate will determine the values of the remaining three variables — the premium, the per unit charge for care, and the total expected demand for care. Because the sharing rate is essentially a price, this result also suggests the use of a univariate indirect-expected utility function for the remainder of the analysis.

The indirect expected utility function V(s) is constructed by substituting into expression (1) the definition of residual income for y_i, expression (11) for k, and $g_i(s)$ for x_i:

$$V(s) = \Sigma q_i W(I - r(s) - [1 - k(s)] \cdot (1-s)p(s)g_i(s)) + \Sigma q_i U_i(g_i(s)).$$

To determine the optimal sharing rate, the first step is to differentiate V(s) with respect to s:

$$\begin{aligned} V'(s) = {} & \Sigma q_i W'(y_i) \cdot \{-r'(s) + k'(s) \cdot (1-s)p(s)g_i(s) \\ & + [1 - k(s)] \cdot p(s)g_i(s) - [1 - k(s)] \cdot (1-s)p'(s)g_i(s) \\ & - [1 - k(s)] \cdot (1-s)p(s)g'_i(s)\} + \Sigma q_i U'_i(x_i) \cdot g'_i(s). \end{aligned} \tag{15}$$

Because the sign of the second derivative is indeterminate, expression (15) has not been set equal to zero. The specification of the model does not guarantee an interior solution. However, as shown in the next section, it is by investigating the characteristics of possible solutions to this optimization problem, including corner solutions, that the primary results of this article are found.[11]

Analysis of the Model

This section examines the implications of uncollectibles for the choice of an optimal sharing rate by both risk neutral and risk averse consumers. Five principal conclusions are reached. First, contrary to the standard no-insurance result, if decreases in uncollectibles and delayed payments lower the per unit charge for health care, a risk neutral consumer may demand a positive level of insurance.

Second, the level of potential savings from the elimination of bad debts and collection costs can be high enough, relative to the costs of moral hazard and

---

[11] If defaults and collection efforts are eliminated from the model, the standard results still hold with this functional specification: risk averse consumers insure against health care risks, while risk neutral consumers do not.

of honesty, to push the sharing rate to unity. This would imply full coverage is optimal. Therefore, in addition to minimizing risk, certain types of insurance programs could be functioning to minimize uncollectibles.

Third, a risk averse consumer chooses a higher amount of insurance coverage than a risk neutral consumer. If preferences exhibit risk aversion, it is therefore even more likely that the optimal insurance plan will provide complete prepayment.

Fourth, if a representative consumer becomes more likely to default on all possible insurance contracts, the optimal sharing rate will be lower. This result holds for both risk neutral and risk averse consumers. The benefits from dishonesty derive from defaults on coinsurance obligations. Thus, the greater the propensity to default, the greater the incentive to contract for a high coinsurance rate. This will increase the amount of "windfall" residual income obtained from the unpaid portion of any coinsurance obligation.

Fifth, the assumption of actuarially fair premiums is dropped. The results of this analysis are consistent with the well-known finding that positive loading fees have a negative effect on the amount of insurance coverage purchased (e.g., Mossin, 1968, and Smith, 1968).

Demand for Hospitalization Insurance by a Risk Neutral Consumer

The expression for the first derivative of the indirect expected utility function in (15) can be simplified by noting that the first order conditions in (3) still must be satisfied. This reduces expression (15) to:

$$\begin{aligned} V'(s) = \; & \Sigma q_i W'(y_i)\cdot\{-r'(s) + k'(s)\cdot(1-s)p(s)g_i(s) \\ & + [1-k(s)]\cdot p(s)g_i(s) - [1-k(s)]\cdot(1-s)p'(s)g_i(s)\}. \end{aligned} \tag{16}$$

Let $W''(y_i) = 0$ define risk neutrality. This implies $W'(y_i)$ is a positive constant. Assume $W'(y_i) = \tau$, where $\tau > 0$. By substituting for $W'(y_i)$, and simplifying, the first derivative $V'(s)$ in (16) becomes:

$$\begin{aligned} V'(s) = \; & \tau\cdot\{-r'(s) + k'(s)\cdot(1-s)p(s)\overline{Q}(s) + [1-k(s)]\cdot p(s)\overline{Q}(s) \\ & - [1-k(s)]\cdot(1-s)p'(s)\overline{Q}(s)\}, \end{aligned} \tag{17}$$

where $\overline{Q}(s) \equiv \Sigma q_i g_i(s)$.

The Case of an Interior Maximum: $0 < s^* < 1$: To investigate the conditions under which risk neutral consumers will demand a positive level of insurance, rewrite expression (17). Define $\phi(s) \equiv [1-k(s)]\cdot(1-s)p(s)$ to be what the consumer perceives as the actual price per unit of hospital care, at the point of purchase. The first derivative of $\phi(s)$ is:

$$\phi'(s) = -k'(s)\cdot(1-s)p(s) - [1-k(s)]\cdot p(s) + [1-k(s)]\cdot(1-s)p'(s). \tag{18}$$

To simplify the necessary condition for an interior maximum, substitute $\phi'(s)$ into expression (17):

$$V'(s^*) = \tau\cdot\{-r'(s^*) - \phi'(s^*)\cdot\overline{Q}(s^*)\} = 0. \tag{19}$$

Therefore, if an interior solution s* exists, the consumer will set the marginal change in the insurance premium equal to the marginal change in the consumer's actual bill for a given level of care.

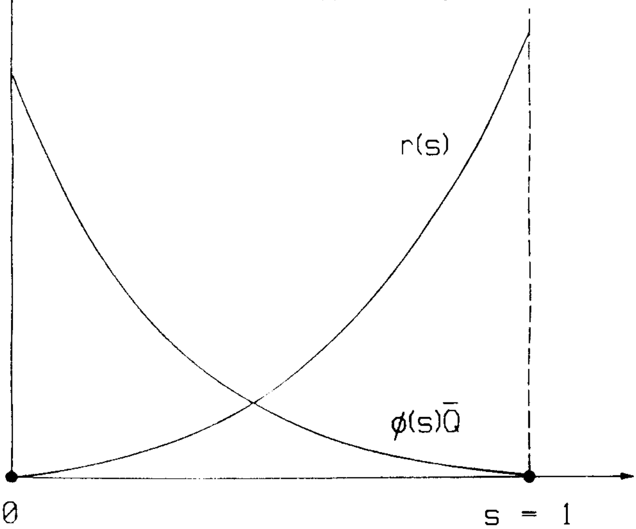
To determine necessary and sufficient conditions for an interior maximum, first note that the consumer's optimization problem is equivalent to minimizing the sum of two "cost" functions: $r(s) + \phi(s) \cdot \overline{Q}$, where $\overline{Q}$ acts as a positive constant. Given that $r(0) = 0$ and $\phi(1) = 0$, standard economic restrictions on the two functions are:

$$r'(s) > 0;\ r''(s) > 0, \text{ and } \phi'(s) < 0;\ \phi''(s) > 0.$$[12]

The restrictions on r(s) imply that the marginal change in the premium is positive and increasing. The restrictions on $\phi(s)$ indicate that, as the sharing rate increases, the consumer will benefit from a decline in the actual outlay for hospital care, but the benefit diminishes as $s \to 1$. The assumptions on the second-order derivatives guarantee that the sum of the two functions is strictly convex, i.e., $r''(s) + \phi''(s) \cdot \overline{Q} > 0$.[13] Figures 3a and 3b illustrate both r(s) and $\phi(s) \cdot \overline{Q}$, and their sum, while Figure 4 illustrates the first order condition for the interior solution: $r'(s^*) = -\phi'(s^*) \cdot \overline{Q}$. Figure 4 makes it clear that the first order condition is analogous to the consumer setting marginal cost equal to marginal benefit.

Figure 3a

Illustration of r(s) and $\phi(s)\overline{Q}$



[12] The requirement $\phi'(s) < 0$ implies that any increase in the purchase price due to a decrease in the default rate must be less than any decrease in the purchase price due to the increased sharing rate and to price concessions offered by the hospital. In a later section, it is shown that this requirement also guarantees that a risk averse consumer demands more insurance coverage than a risk neutral consumer.

[13] Strict convexity is assumed in order to derive, later in this article, the comparative static results for the behavioral relationships p(s), k(s), and $\overline{Q}(s)$.

Figure 3b

Illustration of $r(s) + \phi(s)\overline{Q}$

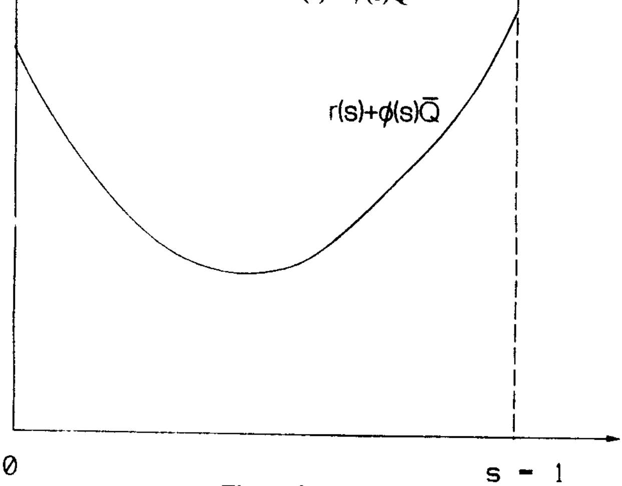
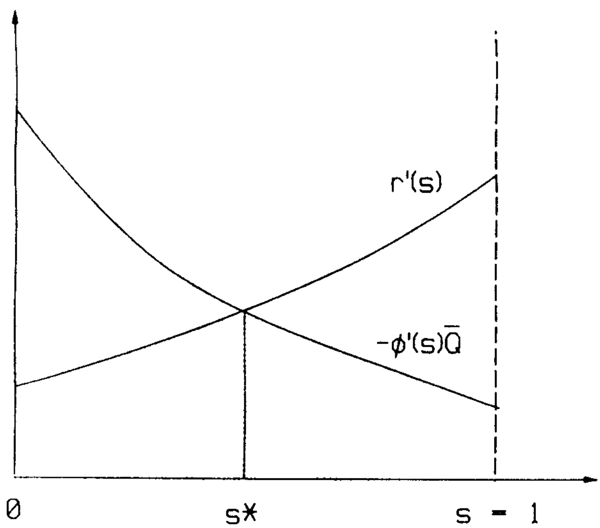


Figure 3c

Marginal Conditions for an Interior Maximum



Manipulation of expression (17) for $V'(s)$ provides additional insight into why an interior maximum might arise in the insurance optimization problem of a risk neutral consumer. First, differentiate equation (13) to find $r'(s)$:

$$r'(s) = p(s)\overline{Q}(s) + sp'(s)\overline{Q}(s) + sp(s)\overline{Q}'(s). \tag{20}$$

By substituting expression (20) for $r'(s)$, and combining terms where possible, the first derivative in (17) becomes:

$$V'(s) = \tau \cdot \{-sp(s)\overline{Q}'(s) + [(1-s)k'(s) - k(s)] \cdot p(s)\overline{Q}(s) - [1-(1-s)k(s)] \cdot p'(s)\overline{Q}(s)\}. \tag{21}$$

Set (21) equal to zero, and rearrange so that each term is positive:

$$s^*p(s^*)\overline{Q}'(s^*) - [(1-s^*)k'(s^*) - k(s^*)] \cdot p(s^*)\overline{Q}(s^*) = -[1-(1-s^*)k(s^*)] \cdot p'(s^*)\overline{Q}(s^*).$$[14]

The first term represents the marginal cost of moral hazard – the increase in the insurance premium because more health care is demanded when the sharing rate increases. The second term represents the "marginal cost of honesty." When, as the sharing rate rises and the default rate falls, the consumer's residual income also falls because a greater portion of any given hospital bill is paid. The term on the right-hand side of the equation represents the marginal benefit of an increase in the sharing rate. The hospital provides price concessions when it experiences a decrease in uncollectibles-related expenses. Thus, the consumer benefits because the total payment to the hospital (from the insurer as well as from the consumer) for a given quantity of care falls. Therefore, the first order condition for an interior maximum states that the consumer will balance the marginal costs of moral hazard and of honesty with the marginal benefit or savings on the price of hospital care.

The Case of Complete Prepayment: $s^* = 1$: The model does not guarantee a positive optimal coinsurance rate, $(1-s^*)$. A risk neutral consumer may choose an optimal insurance contract providing for complete prepayment. To examine the characteristics of this corner solution, assume expression (21) is non-negative at $s^* = 1$, where $k(1) \equiv 0$:

$$V'(s^*)|_{s^*=1} = \tau \cdot \{-p(s^*)\overline{Q}'(s^*) - p'(s^*)\overline{Q}(s^*)\} \geq 0,$$

or

$$-p'(s^*)\overline{Q}(s^*) \geq p(s^*)\overline{Q}'(s^*). \tag{22}$$

Thus, if a risk neutral consumer prefers $s^* = 1$, the marginal savings from hospital price-concessions when billing costs decrease (the left-hand term) must be greater than or equal to the marginal increase in the insurance premium due to the increase in the cost of moral hazard (the right-hand term). However, if the excess consumption of complete prepayment is "too great," the representative consumer will prefer a contract specifying coinsurance – thus, tolerating the hospital's price-surcharge, resulting from the uncollectibles and collection costs.

Alternately, the necessary condition for complete prepayment can be expressed in terms of elasticities with respect to the sharing rate. By rearranging and multiplying both sides by s, expression (22) becomes:

---

[14] An equivalent version of this expression may be written in terms of elasticities with respect to the sharing rate:

$$s \cdot \epsilon_{Q,s} - k[(1-s) \cdot \epsilon_{k,s} + s] + \epsilon_{p,s} \cdot [1-(1-s)k] = 0,$$

where $\epsilon_{Q,s} \equiv \overline{Q}'(s) \cdot s/\overline{Q}$, $\epsilon_{k,s} \equiv k'(s) \cdot s/k$, and $\epsilon_{p,s} \equiv p'(s) \cdot s/p$.

$$|\epsilon_{p,s}| \geq |\epsilon_{Q,s}|,$$

where $\epsilon_{p,s} \equiv p'(s) \cdot s/p$ and $\epsilon_{Q,s} \equiv \overline{Q}'(s) \cdot s/\overline{Q}$.

Thus, for $s^* = 1$, the elasticity of expected demand is required to be less than or equal to the elasticity of hospital price-concessions.[15]

The Case of No Insurance: $s^* = 0$: The assumptions with respect to $r'(s)$ and $\phi'(s)$ do not rule out the possibility of a risk neutral consumer choosing not to insure. Consider expression (21) for $V'(s)$, evaluated at $s^* = 0$:

$$V'(s^*)|_{s^*=0} = \tau \cdot \{[k'(s^*) - k(s^*)] \cdot p(s^*)\overline{Q}(s^*) - [1 - k(s^*)] \cdot p'(s^*)\overline{Q}(s^*)\} \leq 0$$

or

$$-[k'(s^*) - k(s^*)] \cdot p(s^*)\overline{Q}(s^*) \geq -[1 - k(s^*)] \cdot p'(s^*)\overline{Q}(s^*). \quad (23)$$

Both terms are positive. When the sharing rate increases, the default rate falls. Thus, the consumer pays a higher portion of the total hospital charges. This is the cost of honesty. However, when the sharing rate increases, the hospital's price per unit falls and the total bill for a given level of care is less. This is the benefit of insurance. Therefore, the risk neutral consumer will not purchase insurance if the marginal cost of honesty is greater than or equal to the marginal benefit of price negotiations with the hospital.

Shifts in Behavioral Relationships: $p(s)$, $k(s)$, and $\overline{Q}(s)$

Comparative static investigations will center on the effects on the optimal coinsurance rate of shifts in the hospital's price-concession function, $p(s)$, and the consumer's default-response function, $k(s)$, and total expected demand for care function, $\overline{Q}(s)$. In order to focus on the case of an interior optimum for the risk neutral consumer, assume there exists s^* such that expression (21), and expression (19), identically equal zero. Let $D < 0$ denote the second-order sufficiency condition.

Consider first a positive, slope-preserving shift in the $p(s)$ function: $p(s) + \alpha$. To determine the effects of this shift, substitute $p(s) + \alpha$ everywhere for $p(s)$ in expression (21), and differentiate with respect to α. Then, use the implicit function rule to find $\partial s^*/\partial \alpha$:

$$\partial s^*/\partial \alpha = \{s\overline{Q}'(s) - [(1-s)k'(s) - k(s)] \cdot \overline{Q}(s)\} \div D. \quad (24)$$

Because the numerator is positive, the derivative in (24) is negative.

Expression (24) implies that a positive shift in the hospital's price-negotiation function will decrease the value of the optimal sharing rate chosen by a risk neutral consumer. If, for every sharing rate, the price demanded by the hospital increases (the equivalent of an increase in π^*), the benefits to the consumer from dishonesty increase, while the benefits from negotiating price-concessions decrease. The result: coinsurance and coinsurance obligations increase. In contrast, if the insurance provider were able to improve its negotiation position with the hospital, resulting in a negative shift in the $p(s)$

[15] Reversing this expression yields a condition under which complete prepayment would not be an optimum: $|\epsilon_{p,s}| < |\epsilon_{Q,s}|$.

function, then the optimal insurance contract would specify less coinsurance and a higher sharing rate.

Consider next the consumer's default function, k(s). Again, assume a positive shift-factor β. However, because $k(s)|_{s=1} \equiv 0$, for all s, one also must have $dk'(s)/d\beta < 0$. Then, $\partial s^*/\partial\beta$ is found by differentiating expression (21) with respect to β:

$$\partial s^*/\partial\beta = \{[-(1-s)\cdot dk'(s)/d\beta + 1]\cdot p(s)\overline{Q}(s) - (1-s)p'(s)\overline{Q}(s)\} \div D. \quad (25)$$

The numerator is positive, and thus, the derivative is negative.

Expression (25) shows that an increase in the propensity to default will decrease the optimal sharing rate s*. When the default rate increases for every sharing rate, the benefit of defaulting — an increase in residual income — increases with larger coinsurance obligations (of which a larger portion will be defaulted). Thus, the risk neutral consumer will demand less insurance coverage.

Finally, consider a positive, slope-preserving shift in the total expected health care demand function: $\overline{Q}(s)+\gamma$. Again, focusing on expression (21), application of the implicit function rule yields:

$$\partial s^*/\partial\gamma = -\{k'(s)\cdot(1-s)p(s) - k(s)p(s) - [1-(1-s)k(s)]\cdot p'(s)\} \div D.$$

In general, the sign of $\partial^*/\partial r$ is ambiguous; the first and second terms in the brackets are negative, but the third term is positive. However, by determining when the numerator is negative, insight is gained about the conditions under which a positive shift in total expected demand for care will increase the optimal sharing rate, i.e., when $\partial s^*/\partial\gamma > 0$. The numerator will be negative if the following inequality holds:

$$-[1-(1-s)k(s)]\cdot p'(s) > -k'(s)\cdot(1-s)p(s) + k(s)p(s). \quad (26)$$

Because each of the three terms is positive, the condition in (26) will be met only if:

$$-[1-(1-s)k(s)]\cdot p'(s) > -k'(s)\cdot(1-s)p(s). \quad (27)$$

Rewrite expression (27) in terms of elasticities with respect to the sharing rate. Take the absolute value of both sides to yield:

$$|[1-(1-s)k]\cdot\epsilon_{p,s}| > |(1-s)k\cdot\epsilon_{k,s}|. \quad (28)$$

Expression (28) illustrates that the sign of $\partial^*/\partial r$ depends upon the relative magnitudes of the elasticities, with respect to the sharing rate, of the hospital price-concession function and of the consumer default-response function. Because the legal coinsurance rate is $(1-s)$, $(1-s)k$ represents the portion of the coinsurance rate the consumer decides not to pay — a "dishonesty-gain rate." Equivalently, $(1-s)k$ is also the hospital's loss rate per patient-coinsurance obligation. The consumer's actual coinsurance rate is $[1-(1-s)k]$. This implies that, if the absolute value of the elasticity of price concessions, weighted by the actual coinsurance rate, is greater than the absolute value of the elasticity of the default rate, weighted by the dishonesty-gain rate, an increase in the total expected demand for care will increase the optimal sharing rate.

Demand for Hospitalization Insurance by a Risk Averse Consumer

A risk averse consumer will choose a higher sharing rate than a risk-neutral consumer. Risk averse preferences thus strengthen the result that the terms of the optimal insurance plan can stipulate complete prepayment.

Risk Aversion Increases the Demand for Insurance: The proof is accomplished by showing that an interior maximum for a risk neutral consumer cannot maximize the indirect expected utility function of a risk averse consumer. First, simplify the first order condition for the risk averse consumer by substituting $\phi'(s)$, as represented in expression (18), into expression (16):

$$V'(s) = \Sigma q_i W'(y_i) \cdot \{-r'(s) - \phi'(s) \cdot g_i(s)\}. \tag{29}$$

Let $\tilde{s}$ be an interior maximum when preferences are risk neutral. To show that $\tilde{s}$ cannot be a solution when the consumer is risk averse, assume on the contrary that $\tilde{s}$ is a solution. Then, $\tilde{s}$ must solve $V'(s) = 0$ as represented by expression (19):

$$\tau \cdot \{-r'(\tilde{s}) - \phi'(\tilde{s}) \cdot \overline{Q}(\tilde{s})\} = 0$$

or

$$-r'(\tilde{s}) = \phi'(\tilde{s}) \cdot \overline{Q}(\tilde{s}). \tag{30}$$

Evaluate the risk averse first-order condition in expression (29) at $s = \tilde{s}$. Substitute expression (30) for $-r'(\tilde{s})$, and simplify by moving $-\phi'(\tilde{s})$ in front of the summation sign:

$$-\phi'(\tilde{s}) \cdot \Sigma q_i W'(y_i) \cdot [g_i(\tilde{s}) - \overline{Q}(\tilde{s})]. \tag{31}$$

Given that $\phi'(s) < 0$, the final step of the proof is to show that the summed-term is positive. This is done in Appendix B, where expression (31) is shown to be positive if marginal utility of income is diminishing. Thus, the coinsurance rate that solves the maximization problem when the consumer is risk neutral no longer solves it when the consumer is risk averse.

Furthermore, because $V'(\tilde{s})$ is positive, it is also true that the optimal sharing rate for a risk averse consumer must be greater than $\tilde{s}$. Given that the corner solution $s^* = 1$ is possible when the consumer is risk neutral, the likelihood of complete prepayment in this market is even greater if preferences exhibit risk aversion.

The Case of an Interior Maximum: $0 < s^* < 1$: Because the above analysis does not rule out an optimal contract with a positive coinsurance rate, the circumstances under which the corner solution $s = 1$ would not be an optimum will be considered next. First, it would require $V'(1) < 0$. Recalling that $k(s)|_{s=1} \equiv 0$, and substituting (20) for $r'(s)$, expression (16) evaluated at $s = 1$ becomes:

$$V'(s)|_{s=1} = \Sigma q_i W'(y_i) \cdot \{-p(s)\overline{Q}(s) - p'(s)\overline{Q}(s) - p(s)\overline{Q}'(s) + p(s)g_i(s)\}. \tag{32}$$

To attempt to separate the influence of risk aversion on the optimal contract from those of moral hazard and bad debts, rearrange expression (32):

$$V'(s)|_{s=1} = -[p'(s)\overline{Q}(s) + p(s)\overline{Q}'(s)] \cdot \Sigma q_i W'(y_i) + p(s) \cdot \Sigma q_i W'(y_i) \cdot [g_i(s) - \overline{Q}(s)]. \tag{33}$$

Risk aversion implies that the second term is positive.[16] Therefore, in order for an optimal contract to include coinsurance, i.e., for $V'(1) < 0$, the first term must be negative. Because marginal utility of income is positive, this implies that the following must be true: $-p'(s)\overline{Q}(s) < p(s)\overline{Q}'(s)$. Because this inequality is analogous to expression (22), a necessary condition for an interior maximum is:

$$|\epsilon_{Q,s}| > |\epsilon_{p,s}|. \tag{34}$$

Therefore, to guarantee that a provision for coinsurance is included in the optimal insurance contract, the elasticity of total expected demand for care must be greater than the elasticity of the hospital negotiation-price. This is a necessary condition for both risk averse and risk neutral consumers. However, unlike the case of risk neutrality, the inequality in expression (34) is not sufficient for an interior maximum. If it does hold, the first term in (33) will be negative. However, this term also must be greater than the second term in (33), which represents the influence of risk aversion. Thus, for an interior maximum, the marginal cost of moral hazard at s = 1 must be greater than both the marginal benefits derived from hospital price-concessions and from risk pooling.

Shifts in Behavioral Relationships: p(s) and k(s)

Comparative static investigations again will center on the effects on the optimal sharing rate of shifts in the behavioral functions. However, unlike the case of risk neutrality, the configuration of the first order condition for the risk averse consumer rules out conclusions with respect to shifts in the total expected demand for care function, $\overline{Q}(s)$. Therefore, consider shifts in the hospital's price-concession function, p(s), and in the consumer's default-response function, k(s). Assume an interior solution holds for the risk averse consumer, i.e., there exists s* such that expression (16), and (29), identically equals zero. Let $\overline{D} < 0$ represent the second-order sufficiency condition.

The effect of a positive, slope-preserving shift in the p(s) function is more transparent if the analysis is focused on the expression for the first order condition in (29). Again, first substitute $p(s) + \alpha$ everywhere for p(s) and differentiate with respect to α. Second, use the implicit function rule to determine $\partial s^*/\partial\alpha$:

$$\partial s^*/\partial\alpha = \Sigma q_i W'(y_i) \cdot [\overline{Q}(s) + s\overline{Q}'(s) + d\phi'(s)/d\alpha \cdot g_i(s)] \div \overline{D}. \tag{35}$$

[16] See Appendix B.

Because of the effect of a shift in the price-concession function on the marginal change in the consumer's price at the point of purchase, $d\phi'(s)/d\alpha$, the sign of $\partial s^*/\partial\alpha$ now is ambiguous. However, by requiring $d\phi'(s)/d\alpha > 0$ the same result as found in the case of risk neutrality can be guaranteed. Given $\phi'(s) < 0$, this implies that the increase in the hospital's negotiation demands also decreases the marginal benefit from a decline in the coinsurance rate (as reflected in the price at the point of purchase). The assumption $d\phi'(s)/d\alpha > 0$ is sufficient to guarantee expression (35) is negative — and that a positive shift in the hospital's overall price demands will decrease the optimal sharing rate.

Consider next the effect of a positive shift in the default-response function: $k(s)+\beta$, such that $dk'(s)/d\beta < 0$. Differentiate the first order condition for the risk averse consumer in expression (16) with respect to β:

$$\begin{aligned}\partial s^*/\partial\beta \;=\; & \Sigma q_i W'(y_i)\cdot\{-dk'(s)/d\beta\cdot(1-s)p(s)g_i(s) \\ & + p(s)g_i(s) - (1-s)p'(s)g_i(s)\} \div \overline{D}. \end{aligned} \tag{36}$$

The numerator of expression (36) is negative, therefore, $ds^*/d\beta < 0$. This is the same result as found for the risk neutral consumer. Regardless of risk preferences, the less honest the consumer, the higher the level of coinsurance demanded.

Actuarially Unfair Premiums

Assume premiums are actuarially unfair. If the positive loading fee is represented as a factor L such that $L \in (0,1)$, then the insurance premium in equation (13) becomes:

$$r(s) \;=\; (1+L)\cdot sp(s)\overline{Q}(s).$$

The first derivative of the premium is now:

$$r'(s) \;=\; (1+L)\cdot[p(s)\overline{Q}(s) + sp'(s)\overline{Q}(s) + sp(s)\overline{Q}'(s)].$$

Focusing on the optimal insurance contract for a risk averse consumer,[17] the first order condition in expression (16) becomes:

$$\begin{aligned} V'(s) \;=\; & \Sigma q_i W'(y_i)\cdot\{-(1+L)\cdot[p(s)\overline{Q}(s) + sp'(s)\overline{Q}(s) + sp(s)\overline{Q}'(s)] \\ & + k'(s)\cdot(1-s)p(s)g_i(s) + [1-k(s)]\cdot p(s)g_i(s) - [1-k(s)]\cdot(1-s)p'(s)g_i(s)\}. \end{aligned} \tag{37}$$

Expression (37) shows that positive loading fees exercise a downward pressure on the optimal sharing rate. To see this more clearly, assume $\hat{s}$ represents an interior solution to the consumer's optimization problem when premiums are actuarially fair. Let $r(\hat{s})$ represent the actuarially fair premium. By substituting the first order condition in expression (16) into (37) above, it can be shown that $\hat{s}$ is no longer a solution when $L > 0$, and that $V'(\hat{s}) < 0$. Because by assumption $\hat{s}$ solves expression (16) when it is set equal to zero, the following terms from (37) remain after the substitution:

[17] The following results also hold for risk neutrality.

$$V'(\hat{s}) = \Sigma q_i W'(y_i) \cdot [-L] \cdot [p(\hat{s})\overline{Q}(\hat{s}) + \hat{s}p'(\hat{s})\overline{Q}(\hat{s}) + \hat{s}p(\hat{s})\overline{Q}'(\hat{s})]. \tag{38}$$

Expression (38) represents the marginal increase in the cost of insurance coverage caused by the loading factor L. If the original requirements for an interior solution hold, specifically, if $r'(\hat{s}) > 0$, then $V'(\hat{s})$ is negative. Therefore, a positive loading factor decreases the optimal sharing rate.

What if the optimal contract under actuarially fair insurance did not specify coinsurance? Then, it is possible that, even with positive loading, complete prepayment still will be optimal. Expression (16), evaluated at $s = 1$, is required to be positive. This implies that, if $L > 0$, expression (37) will be non-negative. Rearrange (37) into the following four terms, where $k(s)|_{s=1} \equiv 0$:

$$\begin{aligned} V'(s)|_{s=1} = & -p'(s)\overline{Q}(s) \cdot \Sigma q_i W'(y_i) \\ & + p(s) \cdot \Sigma q_i W'(y_i) \cdot [g_i(s) - \overline{Q}(s)] \\ & - p(s)\overline{Q}'(s) \cdot \Sigma q_i W'(y_i) \\ & - L \cdot [p(s)\overline{Q}(s) + p'(s)\overline{Q}(s) + p(s)\overline{Q}'(s)] \cdot \Sigma q_i W'(y_i) \geq 0. \end{aligned} \tag{39}$$

The first and second terms are positive. They represent the marginal benefits of lower hospital prices and of risk pooling caused by an increase in the sharing rate to $s = 1$. The third and fourth terms are negative, representing the marginal cost of moral hazard and of actuarially unfair insurance. Thus, if the marginal benefits are greater than or equal to the marginal costs, complete prepayment will be optimal even when loading fees are positive.

In assessing the importance of positive loading, however, two points should be noted. First, the tax effect has been shown to reduce the effective loading fee for large group plans.[18] Second, insurance firms can recover administrative costs through investment returns earned with premium revenues, rather than through premium loading.

The Possibility of the Absence of Moral Hazard

Unless health care demand is perfectly inelastic, complete prepayment implies moral hazard. Zero coinsurance is equivalent to a zero price, at the point of purchase. On the other hand, if the optimal insurance contract includes coinsurance, and consumers never default on coinsurance obligations, it is possible for $(1-s^*)p(s^*) = C$. Thus, marginal cost pricing is not ruled out when health care is received at a hospital where $(1-s^*)C$ is not the purchase price paid by consumers. However, there is nothing in the insurance literature to indicate that marginal cost pricing is a likely outcome.

If default rates are positive, it is even more unlikely that health care will be priced to eliminate moral hazard. The consumer's purchase price becomes $(1-k^*)(1-s^*)p(s^*)$. Two variables, k^* and s^*, would have to be chosen "correctly" so as to equate this price to marginal cost. Thus, even if the

[18] See Pauly (1986) for comprehensive treatment of the effects of tax policy on health insurance and the health care market.

optimal insurance contract includes a provision for coinsurance, consumption of health care in this market doubtless will contain an element of the type of moral hazard considered in this article.

Additional Applications of the Model

The previous section has shown that unpaid hospital charges and collection costs can lead to a positive demand for insurance by risk neutral consumers, and an optimal insurance contract with no provision for coinsurance. Additional applications of the model in the health insurance market might be found in the contractual allowances negotiated with hospitals by insurers such as Blue Cross, the complete prepayment plans offered by university health services and health maintenance organizations, as well as Blue Cross service-benefit plans, and group insurance contracts with coverage for planned health care expenditures. These issues are discussed below.

Contractual Allowances Negotiated by Blue Cross

Uncollectibles may help explain a major puzzle in the health insurance market in the United States – the ability of Blue Cross service-benefit health plans to have survived so long in the marketplace, despite the seemingly inherent vulnerability of their contracts to moral hazard. Historically, members of Blue Cross plans received prepaid service benefits, hospitalization coverage that does not required consumer copayment. Thus, once admitted to a hospital, subscribers essentially received free care. In contrast, the private insurance firms, with which Blue Cross plans competed, protected themselves from the potential misuse of insurance benefits by limiting the dollar amount reimbursed with a system of deductibles and coinsurance.

The competitiveness of Blue Cross insurance plans has been linked to their ability to negotiate discounted cost-reimbursement contracts with hospitals, amounting to 5 to 10 percent below the private charges paid by commercial insurers and uninsured patients (see Goldberg and Greenberg, 1982). It is possible that these contractual allowances reflect, in part, the lower collection expenses and bad debt write-offs hospitals experience with Blue Cross subscribers as opposed to privately-insured patients with larger copayment responsibilities.

This idea is not unfamiliar to the Blue Cross and Blue Shield Association. Headen, then a research economist with the organization, argued that "commercially insured patients . . . consume more credit and exhibit a higher risk of default for any level of credit . . . than the Blue Cross patients" (1982, p. 637). When first started, Blue Cross prepayment plans were credited with significantly reducing the high levels of uncollectibles experienced by community hospitals at that time.[19] Similarly, Cohodes writes that: "Medicare and Medicaid programs have greatly improved the financial position of most hospitals, by underwriting the expenses of those population groups who

[19]See, for example, Groseclose (1932), Smith (1933), and Hayes (1954).

formerly caused the largest portion of hospital bad debt and charity care" (1986, p. 263).[20]

If uncollectibles represent a failure in the legal contract system, i.e., the failure of patients to pay health care bills fast enough or at all, contractual allowances may exist because service benefit plans have been a successful way of dealing with the failure. This also can help explain the increase in the number of price negotiations between other third party payers and hospitals. According to the model presented in this article, two factors are important: noncompetitive hospital prices, providing the latitude for price concessions, and an expected decrease in the default rate and total uncollectibles as a result of a decrease in patient-copayment responsibilities.

Advantages of Complete Prepayment: The Case of University Health Services

The current analysis suggests that complete prepayment may be an optimal insurance contract. This could explain the trend toward complete prepayment at student health services in major universities throughout the country – especially since the health services themselves have offered as an explanation the subsequent reduction in bad debts.

Depending upon payment policies, student health services can provide an analogy to a hospital's experience both with commercial insurers and with Blue Cross service-benefit plans. If operated under a system of subsidy plus fees-for-service, a health service is comparable to a private insurance program. Collected from tuition payments, the subsidy is similar to an annual premium. The service fees act as coinsurance. On the other hand, completely prepaid health services operate like a Blue Cross prepayment plan or a health maintenance organization.

In early 1981, the University of Michigan Health Service (1981) completed a study of the feasibility of replacing its fee-for-service system with a full prepayment system. As most economists would have predicted, data showed that a 33 percent increase in demand could be expected if prepayment were substituted for fees-for-service. However, the study concluded that, even considering moral hazard, the savings in collections costs and uncollectibles would be large enough to warrant closing down the billing operations.

The University of Michigan Health Service switched to the prepaid system in 1982. This action was consistent with 22 other major universities in the east and midwest that also had opted for prepayment.[21] A 1985 follow-up report confirmed that, although increased utilization was experienced by the University of Michigan Health Service, the savings in bad debts and other billing-related costs outweighed the cost of the additional usage (Briefer, 1985).

---

[20] Cohodes (1986) also cites a 1983 Urban Institute study which found that, compared to Blue Cross patients, the level of bad debts experienced by nonprofit hospitals in 1980 were nearly twice as high for Medicare patients and nearly four times as high for commercial insurance patients.

[21] Survey of 27 major university health services conducted by the author, Summer 1986.

Health maintenance organizations also can take advantage of reduced uncollectibles. In the health maintenance organization, health care is consumed only after a joiner's fee is prepaid. Thus, health maintenance organizations (as well as any other prepaid health care organization) benefit from the cost savings obtained from reductions in collection operations and bad debts.

Non-Random Health Care Coverage

Non-random health care is defined as planned health expenditures such as routine dental work, annual physicals and pregnancy care — care which is considered to be elective by many insurance providers. Based upon the advantages of risk pooling, insurance theory would not predict the inclusion of coverage for non-random health care in group contracts. However, this type of coverage is common. Minimizing bad debts could be a motivating factor, and as a limiting case, complete prepayment would eliminate defaults and collection costs. If providers of non-random health care make price concessions available to group plans based on these types of savings, the results of the model presented in this article would apply.

Summary and Conclusions

The author has argued that uncollectibles are an important component to consider when analyzing the markets for health care and health insurance. The standard economic model of health insurance demand was extended 1) to incorporate the choice to default directly into the consumer's health care demand and payment decisions, and 2) to study the effects of uncollectibles and bill-collection efforts on health care prices, insurance rates, and demand. The behavior of three economic agents was modeled: a hospital, an insurance supplier, and a representative consumer. However, the primary decision-making agent was the consumer, who maximized expected utility subject to the pricing constraints of the hospital and the insurance firm.

It was found that if hospitals negotiate price concessions with third party payers in return for lower patient-copayment responsibilities, a representative risk neutral consumer may purchase actuarially fair insurance. In addition, because the benefits from lower bad debts and any risk pooling may outweigh the costs of moral hazard, the optimal insurance contract can specify complete prepayment. While the analysis focused on the hospital care market, the results can be generalized to other medical care providers and payer groups involved in payment negotiations based upon bad debt or billing cost reductions.

Many analysts have argued that health care costs could be reduced by increasing the patient's direct share of the bill, with higher deductibles and coinsurance leading to lower demand for health care. However, as patient-copayment responsibilities increase, the level of uncollectibles and default rates also may increase. Factors unique to the provision of health care make it unlikely that bad debts will disappear unless the possibility of default is eliminated. Therefore, it is important for studies of the implications of increases in patient-copayment responsibilities to account for the likelihood of higher levels of uncollectibles, and for the subsequent effects on consumer welfare and health care costs.

The advantages of minimizing bad debts may help explain the trend toward price concessions negotiated between health care providers and numerous payer groups, such as health maintenance and preferred provider organizations. The results of the model are also consistent with contractual allowances given to cost-based insurers such as Blue Cross, complete prepayment plans, and non-random health care coverage.

Further research is needed to determine which is greater – the downward pressure on coinsurance rates from the benefits of risk pooling and hospital price-concessions due to the containment of bad debts or the upward pressure from the costs of moral hazard and the benefits to consumers of defaulting on coinsurance obligations. The model does not rule out risk neutral consumers choosing against health insurance; neither does it rule out risk averse consumers choosing against complete prepayment.

Recently, Aaron and Schwartz (1990) and Fuchs (1990) cited billing and administrative costs as part of the explanation for the high level of health care expenditures in the United States compared to other developed countries.[22] This suggests that it also is important to identify both the effectiveness of various collection efforts on bad debts and default rates, and the magnitude and exact nature of the role bad debts play in the price concessions negotiated with health care providers.

Appendix A

Proof of Existence of a Unique Equilibrium

Represent the hospital care market as a system of three equations in four unknowns – r, p, s and $\overline{Q}$:

$$\text{Hospital:} \quad \pi^* = \{1-(1-s)[k(s)+b(s)]\}\cdot p\overline{Q} - C\overline{Q}; \tag{A1}$$

$$\text{Insurer:} \quad r = sp\overline{Q}; \tag{A2}$$

$$\text{Consumer:} \quad \overline{Q} = F(r,(1-s)p). \tag{A3}$$

Expression (A1) is found substituting expressions (11) and (12) into (5), and simplifying. Expression (A3) represents the sum over q_i of the demand functions in (4).

Because the system is composed of three equations in four unknowns, one variable will be held constant. Then, the conditions under which unique equilibrium solutions exist for the remaining variables will be determined.[23] Begin by fixing $r = \hat{r}$. Then, rewrite expression (A3) as

[22] According to Aaron and Schwartz: "Some estimates place the cost of administration at as much as 22 percent of national health care spending . . ." (1990, p. 420).

[23] Because $k = k(s)$, $k'(s) < 0$, and $b = b(s)$, $b'(s) < 0$, equilibrium values only need be determined for the four variables r, p, s, and $\overline{Q}$.

$$\overline{Q} = F((1-s)p). \tag{A4}$$

To simplify notation, define

$$\mu(s) \equiv \{1-(1-s)[k(s)+b(s)]\}.$$

Substitute both expressions into (A1):

$$\pi^* = \mu(s)\cdot pF((1-s)p) - C\cdot F((1-s)p). \tag{A5}$$

Expression (A5) depends only upon p and s. Therefore, use the implicit function rule to prove that if the hospital's behavior is consistent with the necessary and sufficient conditions for constrained-revenue maximization, if $\pi^* > 0$, and if $F'((1-s)p) < 0$, then $dp/ds < 0$. This allows writing the univariate function $p = p(s)$.

Application of the implicit function rule yields:

$$\begin{aligned} dp/ds = &-\{\mu'(s)\cdot pF((1-s)p) + \mu(s)\cdot pF'((1-s)p)\cdot(-p) - C\cdot F'((1-s)p)\cdot(-p)\} \\ &\div \{\mu(s)\cdot F((1-s)p) + \mu(s)\cdot pF'((1-s)p)\cdot(1-s) - C\cdot F'((1-s)p)\cdot(1-s)\}. \end{aligned} \tag{A6}$$

Consider the numerator in (A6) first. Given that $k'(s) < 0$ and $b'(s) < 0$, it follows that $\mu'(s) > 0$. Thus, the first term is positive. Combine the second and third terms in the numerator:

$$[\mu(s)\cdot p - C]\cdot F'((1-s)p)\cdot(-p).$$

Because $\pi^* > 0$, it can be concluded from (A1) that $[\mu(s)\cdot p - C] > 0$. Given that the slope of the demand function is constrained such that $F'((1-s)p) < 0$, all of the terms in brackets are positive and the numerator of dp/ds in (A6) is negative.

The denominator of expression (A6) is the derivative with respect to p of the excess-revenue function represented by (A3). If the Kuhn-Tucker sufficiency conditions for a constrained-revenue maximization are satisfied, then the denominator will be strictly positive.[24] This implies the function p(s) is monotonically decreasing in s.

Substitute p(s) for p in (A4):

$$\overline{Q} = F((1-s)p(s)). \tag{A7}$$

Since (A7) is a univariate function of s, it may be rewritten as $\overline{Q}(s)$. Take the first derivative of (A7) to find $\overline{Q}'(s)$:

$$\overline{Q}'(s) = F'((1-s)p(s))\cdot[(1-s)p'(s) - sp(s)].$$

Because $F'((1-s)p(s)) < 0$ and $p'(s) < 0$, $\overline{Q}'(s)$ is strictly positive.

Thus far, it has been shown that both p and $\overline{Q}$ can be expressed as univariate functions of s. By substituting p(s) and $\overline{Q}(s)$ into (A2), it is clear that the premium also can be expressed as a function of s:

---

[24] See Chaig (1984, p. 750-752) for the derivative of the equivalent condition when quantity, Q, is the choice variable: $R'(Q) - C'(Q) < 0$. To derive the condition used in this proof, substitute the demand function, Q(p), for Q, where $Q'(p) < 0$, and totally differentiate with respect to p. This yields: $R'(p) - C'(p) > 0$.

$$r(s) = sp(s)\overline{Q}(s).$$

Relax the assumption that r is fixed in order to investigate the derivative of r with respect to s:

$$r'(s) = p(s)\overline{Q}(s) + sp'(s)\overline{Q}(s) + sp(s)\overline{Q}'(s). \tag{A8}$$

The first and third terms in (A8) are positive, however, the second term is negative. Because $r(0) = 0$, the assumption $r'(s) > 0$ is required to guarantee that a unique value of r exists for each value of s.[25]

To summarize, the following three functional relationships have been derived:

$$p = p(s), \text{ such that } p'(s) < 0;$$
$$\overline{Q} = \overline{Q}(s), \text{ such that } \overline{Q}'(s) > 0;$$

and $$r = r(s), \text{ such that } r'(s) > 0.$$

Given that each function is univariate in s, continuous, and monotonic, for a given value s*, there exist unique values p*, $\overline{Q}$*, and r* that solve the system of equations (A1), (A2), and (A3).

Appendix B

Proof that $\Sigma q_i W'(y_i) \cdot [g_i(\bar{s}) - \overline{Q}(\bar{s})]$ *is Positive*

In the following proof, expression (31) is shown to be positive when consumers are risk averse, i.e., if $W(y_i)$ is strictly concave. Ignoring the first term, rewrite (31) as follows: partition the $g(\bar{s})$'s into the set of those that are greater than the expected level of consumption, $\overline{Q}(\bar{s})$, and the set that are less than or equal to $\overline{Q}(\bar{s})$. Let $g_j(\bar{s})$ represent all $g_i(\bar{s}) > \overline{Q}(\bar{s})$, (J of them), and $g_m(\bar{s})$ represent all $g_i(\bar{s}) \leq \overline{Q}(\bar{s})$, (M of them). This yields the following two sums:

$$\sum_J q_j W'_j \cdot [g_j(\bar{s}) - \overline{Q}(\bar{s})] + \sum_M q_m W'_m \cdot [g_m(\bar{s}) - \overline{Q}(\bar{s})] \tag{B1}$$

where $W'_j \equiv W'(y_i)|_{g_j(\bar{s})}$ and $W'_m \equiv W'(y_i)|_{g_m(\bar{s})}$.

Next, it is shown that the first term in (B1) is positive, while the second is negative. If $W'_j = W'_m$, the sum would equal zero. However, the marginal utilities with respect to income in states j and m are not equal, because each is evaluated at a different level of health care consumption. Moreover, $W'_m \leq W'_j$, for all m and j, because in state of health j, where actual treatment demand is greater than expected demand, the corresponding marginal utility of income is larger than in states where the opposite is true. The more income spent on health services, the less available to be spent on other consumption goods. Therefore, if $g_j(\bar{s}) > \overline{Q}(\bar{s}) \geq g_m(\bar{s})$, it is also true that $W'_j \geq W'_m$. Thus, the positive terms in (B1) are multiplied by a series of factors that are each

---

[25] The requirement $r'(s) > 0$ implies that any decrease in the premium due to price concessions offered by the hospital must be swamped by the increases in the premium due to the increased sharing rate for a given level of care and to the increased comsumption from moral hazard.

larger than those multiplied times the negative terms. The summation term in expression (31) must be positive if the consumer is risk averse.

REFERENCES

Aaron, Henry and William B. Schwartz, 1990, Rationing Health Care: The Choice Before Us, *Science*, 247: 418–22.

Allingham, Michael G. and Agnar Sandmo, 1972, Income Tax Evasion: A Theoretical Analysis, *The Journal of Public Economics*, 1: 323–38.

Alton, Walter G. Jr., 1977, *Malpractice: A Trial Lawyer's Advice for Physicians (How to Avoid, How to Win)*, (Boston: Little, Brown and Company).

Andreano, Ralph and John J. Siegfried, eds., 1980, *The Economics of Crime*, (New York: John Wiley and Sons).

Arrow, Kenneth J., 1976, Welfare Analysis of Changes in Health Coinsurance Rates, in: Richard N. Rosett, ed., *The Role of Health Insurance in the Health Services Sector* (New York: National Bureau of Economic Research) 3–23.

Becker, Gary S., 1968, Crime and Punishment: An Economic Approach, *The Journal of Political Economy*, 76: 169-217.

Biggs, Thomas D., 1981, Assistant Director of Finance, University Hospital, Ann Arbor, Michigan. Interview, August 1981.

Block, M. K. and J. M. Heineke, 1975, A Labor Theoretic Analysis of the Criminal Choice, *The American Economic Review*, 65: 314–25.

Briefer, Dr. Caeser, 1985, Director of Financial Services, University of Michigan Health Services, Ann Arbor, Michigan. Telephone Interview, June 1985.

Center for Health Policy Studies, 1984, *Massachusetts Payer Differential Study*, (Columbia, Maryland: Center for Health Policy Studies).

Cohodes, Donald R., 1984, Cost-Shifting: A Modern Myth, *Journal of Health Politics, Policy and Law*, 9: 261–67.

Cook, Phillip J. and Daniel A. Graham, 1977, The Demand for Insurance and Protection: The Case of Irreplaceable Commodities, *The Quarterly Journal of Economics*, 91: 143–56.

Cornett, Marcia Millon, 1988, Undetected Theft as a Social Hazard: The Role of Financial Institutions in the Choice of Protection Mechanisms, *The Journal of Risk and Insurance*, 55: 723–33.

Ehrlich, Isaac, 1973, Participation in Illegitimate Activities: A Theoretical and Empirical Investigation, *The Journal of Political Economy*, 81: 521–65.

Elliott, Catherine S., 1987, A Note on Risk Aversion and Multivariate, State-Dependent Preferences, *Economics Letters*, 24: 113–16.

Feldman, Roger and Bryan Dowd, 1986, Is There a Competitive Market for Hospital Services? *Journal of Health Economics*, 5: 277–92.

Fuchs, Victor R., 1990, The Health Sector's Share of the Gross National Product, *Science*, 247: 534–38.

Gibson, Robert M., Daniel R. Waldo, and Katherine R. Levit, 1983, National Health Expenditures, 1982, *Health Care Financing Review*, 5: 1-31.

Goldberg, Lawrence G. and Warren Greenberg, 1982, The Dominant Firm in Health Insurance, Working Paper, presented to the American Economic Association Meetings, New York, December 30, 1982.

Groseclose, J. H., 1932, Hospital Insurance is Difference between Solvency, Insolvency, *Hospital Management*, 34: 19.

Gunther, Marilyn, 1982, Credit Counseling Center, Southfield, Michigan. Telephone interview, March 1982.

Hayes, John H., ed., 1954, *Financing Hospital Care in the United States, Volume 1: Factors Affecting the Costs of Hospital Care*, (Commission on the Financing of Hospital Care, New York: Blakiston Company).

Headen, Alvin E. Jr., 1982, The Relationship Between the Blue Cross Market Share and the Blue Cross 'Discount' on Hospital Charges: Comment, *The Journal of Risk and Insurance*, 49: 634-42.

Heineke, J. M., 1978, Economic Models of Criminal Behavior: An Overview, in: J. M. Heineke, ed., *Economic Models of Criminal Behavior* (New York: North-Holland Publishing Company) 1-34.

Henderson, James M. and Richard E. Quandt, 1980, *Microeconomic Theory: A Mathematical Approach*, (New York: McGraw-Hill Book Company).

Holder, A. R., 1975, *Medical Malpractice Law*, (New York: John Wiley and Sons).

Kolm, Serge-Christophe, 1973, A Note on Optimum Tax Evasion, *The Journal of Public Economics*, 2: 265-70.

Mossin, Jan, 1968, Aspects of Rational Insurance Purchasing, *The Journal of Political Economy*, 76: 553-68.

Ohsfeldt, Robert L., 1985, Uncompensated Medical Services Provided by Physicians and Hospitals, *Medical Care*, 23: 1338-44.

Pauly, Mark V., 1968, The Economics of Moral Hazard: Comment, *The American Economic Review*, 58: 531-37.

———, 1986, Taxation, Health Insurance, and Market Failure in the Medical Economy, *The Journal of Economic Literature*, 24: 629-75.

Raviv, Artur, 1979, The Design of an Optimal Insurance Policy, *The American Economic Review*, 69: 84-96.

Schlesinger, Mark, Theodore R. Marmor, and Richard Smithey, 1987, Nonprofit and For-Profit Medical Care: Shifting Roles and Implications for Health Policy, *Journal of Health Politics, Policy and Law*, 12: 427-57.

Singh, Balbir, 1973, Making Honesty the Best Policy, *The Journal of Public Economics*, 2: 257-63.

Sloan, Frank A., Joseph Valvona, and Ross Mullner, 1986, Identifying the Issues: A Statistical Profile, in: Frank A. Sloan, James F. Blumstein, and James M. Perrin, eds., *Uncompensated Hospital Care: Rights and Responsibilities* (Baltimore: The Johns Hopkins University Press) 16-53.

Smith, Herman, 1933, Group Payment Plans, *Bulletin of the American Hospital Association*, 7: 68-76.

Smith, Vernon L., 1968, Optimal Insurance Coverage, *The Journal of Political Economy*, 76: 68-77.

University of Michigan Health Service, 1981, University of Michigan Health Service Funding Proposal, unpublished report to the Regents, April 1981.

Viscusi, W. Kip and William N. Evans, 1990, Utility Functions That Depend on Health Status: Estimates and Economic Implications, *The American Economic Review*, 80: 353-74.