NBER WORKING PAPER SERIES

ARE INVISIBLE HANDS GOOD HANDS?
MORAL HAZARD, COMPETITION, AND
THE SECOND BEST IN HEALTH CARE MARKETS

Martin Gaynor
Deborah Haas-Wilson
William B. Vogt

Working Paper 6865
http://www.nber.org/papers/w6865

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 1998

We are grateful to participants in seminars at Case Western Reserve University, the University of Pennsylvania, and at the Third Biennial Conference on the Industrial Organization of Health Care and the 1998 American Economic Association meetings for helpful comments. This research was supported in part by a grant from the Robert Wood Johnson Foundation. The usual caveat applies. The views expressed here are those of the author and do not reflect those of the National Bureau of Economic Research.

Are Invisible Hands Good Hands? Moral Hazard,
Competition, and the Second Best in Health Care Markets
Martin Gaynor, Deborah Haas-Wilson,
and William B. Vogt
NBER Working Paper No. 6865
December 1998

ABSTRACT

The nature, and normative properties, of competition in health care markets has long been the subject of much debate. In particular, policymakers have exhibited a great deal of reservation toward competition in health care markets, as demonstrated by the plethora of regulations governing the health care sector. Currently, as consolidation rapidly occurs in health care markets, concern about reduced competition has arisen. This concern, however, cannot be properly evaluated without a normative standard. In this paper we consider what the optimal benchmark is in the presence of moral hazard effects on consumption due to health insurance. Moral hazard is widely recognized as one of the most important distortions in health care markets. Moral hazard due to health insurance leads to excess consumption, therefore it is not obvious that competition is second best optimal given this distortion. Intuitively, it seems that imperfect competition in the health care market may constrain this moral hazard by increasing prices. We show that this intuition cannot be correct if insurance markets are competitive. A competitive insurance market will always produce a contract that leaves consumers at least as well off under lower prices as under higher prices. Thus, imperfect competition in health care markets can not have efficiency enhancing effects if the only distortion is due to moral hazard.

Martin Gaynor
H. John Heinz III School
of Public Policy and Management
Carnegie Mellon University
Pittsburgh, PA 15213-3890
and NBER
mgaynor@andrew.cmu.edu

Deborah Haas-Wilson
Department of Economics
Smith College
Northampton, MA

William B. Vogt
H. John Heinz III School
of Public Policy and Management
Carnegie Mellon University
Pittsburgh, PA 15213-3890
and NBER
wilibear@andrew.cmu.edu

I. Introduction

The nature and normative properties of competition in health care markets have long been the subject of debate. Many policymakers have exhibited reservation toward competition in health care markets.[1] Additionally, federal antitrust enforcement agencies were not vigorous in health care prior to the late 1970's. Currently, as consolidation rapidly occurs in health care markets, concern about reduced competition has arisen (Gaynor and Haas-Wilson, forthcoming). This concern, however, cannot be properly evaluated without a normative standard.

Antitrust enforcement policy in health has been based on a view of health care being like all other industries (e.g., Weller, 1983; Bingaman, 1995). Thus, competition serves as the benchmark.[2] In particular, distortions in health care markets, and their impacts on the socially optimal amount of competition, are not considered. Some have argued that particular distortions that characterize health care markets imply that competition is not optimal (Robinson and Luft, 1985; Crew, 1969; Lynk, 1995).

In this paper we consider what the optimal benchmark is in the presence of moral hazard effects on consumption due to health insurance. Moral hazard is widely recognized as one of the most important distortions in health care markets. In general, economic analysis suggests that marginal-cost pricing leads to static Pareto optimal allocations. In health care markets, however, moral hazard due to health insurance leads to excess consumption, in the sense that insured individuals will consume medical services past the point where the marginal utility of an additional service is equal to its marginal cost (Arrow, 1963; Pauly, 1968). Since health

---

[1] For example, hospital entry, investment, and service offerings have been regulated via state certificate of need regulations and health planning agencies, and pricing has been regulated by hospital "all-payer" regulation in some states and by Medicare and Medicaid.

[2] Note that actual antitrust enforcement is considerably more complicated than this simple representation. In particular, the standard employed is often whether a situation makes consumers worse off than they would be in its absence. This does not necessarily conform to enforcing a competitive outcome nor to maximizing social welfare.

insurance pays for part or all medical expenses, insured individuals face a price that is lower than the market price and consume more of the medical good than is optimal. Therefore it is not obvious that competition or marginal cost pricing is second best optimal given this distortion.

While the problem of optimal insurance in the presence of moral hazard has been extensively analyzed (e.g., Arrow, 1963; Borch, 1968; Pauly, 1968; Zeckhauser, 1970; Feldstein, 1973; Pauly, 1974; Feldman and Dowd, 1991; Arnott and Stiglitz, 1991; Ma and McGuire, 1997; Ma and Riordan, 1997), the issue of the optimal amount of competition in the presence of moral hazard has not. Intuitively, it seems that imperfect competition in the medical care market may constrain this moral hazard by increasing prices.[3] An early paper by Crew (1969) reaches this conclusion using graphical arguments. This intuition has been established as a kind of folk theorem in health economics (Frech, 1996; Pauly, 1998) and has influenced thinking about the optimality of competition in medical markets.[4]

An important limitation of this result is that the endogenous determination of the degree of consumer cost sharing by the insurance industry is not considered. These papers do not allow insurers to alter coinsurance rates in response to changes in medical prices.[5]

Nonetheless, it is clear that competition is the ideal, or the benchmark against which real world outcomes are measured.

[3] This intuition derives from the theory of the second-best. Put colloquially, in a second-best world, two distortions can be better than one. For example, it can be optimal for a producer with monopoly power to bear no product liability in the presence of consumer misperceptions (Polinsky and Rogerson, 1983). If consumers underestimate expected losses, it is desirable to shift liability from the producer to them, in order to counteract the producer's tendency to restrict output due to monopoly power. In another setting, Shapiro (1982) shows that imperfect information can be welfare improving in the presence of monopoly power. If consumers overestimate quality they will buy more, counteracting the monopoly restriction of output. Another example is that it may not be optimal to fully tax a monopolist for the cost of its pollution, again to counteract the incentive to restrict output (Buchanan, 1969).

[4] Beazoglou and Heffley (1995), Beazoglou, Heffley, Niakas, and Kyriopoulos (1993), and Dor and Rizzo (1995) are some papers which explicitly examine the effects of health insurance on the optimal market structure for medical markets.

[5] Insurers do appear to alter their policies in response to medical prices. Frech (1979) and Phelps (1973) present evidence that the price of hospital care has a negative effect on the proportion of hospital expense paid by insurance, i.e., the coinsurance rate.

We model the optimal amount of health insurance and competition, given risk aversion and moral hazard. We show that if insurance markets are competitive and insurers set the degree of consumer cost sharing optimally, then contrary to the standard intuition, adding another distortion does not improve welfare. Since the insurer has already made the tradeoff between risk reduction and moral hazard, a price increase in the medical market cannot wring any further surplus out of the resulting decrease in moral hazard. Thus, imperfect competition in medical markets cannot have efficiency enhancing effects, even in the presence of moral hazard.

The principal claim of this paper is that most of economists' intuition regarding the welfare effects of price changes in markets not distorted by moral hazard applies quite well to markets where decision-making by consumers is distorted by moral hazard. In particular, lower prices are better for consumers than are higher prices. Furthermore, the gain to consumers from lowering price from supra-marginal cost levels to marginal costs outweighs the loss of profit to the medical industry. Finally, the usual method of computing consumer's surplus by integration under the demand curve is still appropriate in markets with moral hazard.

In what follows, the basic model and setup are presented in Section II. Section III contains an analysis of optimal competition when insurance markets do not adjust to medical prices. Section IV contains the main analysis in the paper: the normative properties of competition when the insurance market is competitive and free to respond optimally to price levels in the medical market. Section V demonstrates a method to quantify the change in consumers' surplus from a price change, an extension of the results to the case of managed care occupies Section VI, and we summarize and conclude in Section VII.

II. The Model

We use a standard model of insurance. There is a (measure 1) continuum of consumers, all of whom are identical except with regard to the realization of a random variable, ε, which is a shock to health. Consumers are uncertain *ex ante* with regard to the realization of ε, although they know its distribution. The size of the loss associated with a realization of ε is privately known to the consumer *ex post* (or at least is not verifiable to a court or other contract enforcer) and is remediable (at least in part) through the consumption of a good, which we will call the medical good. The medical good has a price, p, and is produced at a constant marginal cost, c. Insurance contracts take the form of a premium, assessed with certainty, and a partial cost reimbursement for consumption of the medical good. A contract is a pair (τ, m), where m is the premium paid by the consumer to the insurance company, and $\tau \in [0, p]$ is the price faced by the consumer for the medical good. The insurance company implements this price by facing the consumer with a coinsurance rate, $\theta = \frac{\tau}{p}$ so that the insurance company reimburses the consumer a fraction $1-\theta$ of his expenditures on the medical good.

Consumption of medical care is determined in the following way. Consider a consumer possessing an insurance contract, (τ, m). After the consumer's loss, ε, is realized, he solves:

$$\begin{aligned} &\max_x U\left(Y + \pi_i + \pi_m - m - \tau x, x, \varepsilon\right) \\ &s.t. \quad x \geq 0 \\ &\qquad Y - \tau x - m + \pi_i + \pi_m \geq 0, \end{aligned}$$

where Y is the consumer's income (except from medical or insurance industry profits), π_i, π_m are (this consumer's share of) the profits of the insurance and medical industries, respectively, p

is the price of the medical good, and x is the quantity of the medical good consumed.[6] We denote the solution to this problem $x^*(Y-m+\pi_i+\pi_m,\tau,\varepsilon)$. Substituting x^* into U yields the consumer's indirect utility function, $V(Y-m+\pi_i+\pi_m,\tau,\varepsilon)$. Taking expectations over ε defines the expected indirect utility function:

$$EV(Y-m+\pi_i+\pi_m,\tau)=E\{V(Y-m+\pi_i+\pi_m,\tau,\varepsilon)\}.$$

We assume throughout that $U_1>0$ and the insurance industry encounters no administrative costs.

The profits of the medical and insurance industries may be calculated as follows:[7]

$$\begin{aligned}\pi_i &= m-E\{(p-\tau)x^*\}\\ \pi_m &= E\{(p-c)x^*\}\end{aligned}$$

The consumer will consume the medical good to the point at which $\frac{U_2}{U_1}=\tau$. If the medical good is normal, there is a perfectly competitive medical market (so that $p=c$), and if $\tau<p$, then this consumption will not be optimal, *ex post*. Too much of the medical good will be consumed.

A graphical depiction of this standard analysis is presented in Figure 1. The *ex post* demand curve for the medical good is denoted $D(p)$. Insurance contracts distort demand by facing consumers with only a proportion θ of their expenses for medical care and this leads to

[6] Our setup differs slightly from the standard one in that we allow profits from the medical and insurance industries to be distributed to consumers. This is necessary since we consider the impact of monopoly power in the medical and insurance markets.

[7] Both the profits of the insurance and medical industry are expectations by our assumption of a continuum of consumers, with unit mass and indexed by ε.

the distorted *ex post* demand curve $D(\theta p)$. The *ex post* efficiency loss due to moral hazard at price P_1 is the area A. Since insurance lowers the price the consumer faces from P_1 to θP_1 consumption increases from the first-best quantity, X^*, to X_l. When price falls from P_1 to marginal cost at P_2, welfare loss *increases* to the area $A+B$. So, by decreasing prices from supra marginal cost levels to marginal costs, there is a welfare loss, the area B.

However, *ex post* efficiency is not a sensible welfare criterion here, as it ignores the benefits to obtaining insurance *ex ante* (Feldman & Dowd, 1991; Feldstein, 1973). Our purpose in the next section is to evaluate under what conditions a price increase (relative to marginal cost pricing) in the medical market can improve *ex ante* efficiency.

III. The Second Best With a Rigid Insurance Market

As a benchmark case consider now the problem of a benevolent social planner who cannot observe ε, cannot sell insurance, and who must allocate the medical good via the price mechanism. The 2nd best allocation under these constraints is achieved by setting a price, τ, for the medical good equal to the argmax of:

Social Planner's Program:

$$\begin{aligned}&\max_\tau EV(Y+\pi_m,\tau)\\ &\pi_m = E\{(\tau-c)x^*\}\end{aligned} \quad .$$

Let us ignore, for the moment, the problems of existence and uniqueness and suppose that τ^* is the solution to the social planner's problem. Let us further suppose that our social planner must attempt to implement this τ^* only by manipulating the price in the medical market, say via antitrust enforcement, certificate of need laws, or rate-setting. Finally, suppose that the insurance market sets a fixed, rigid coinsurance rate θ, which is insensitive to the medical price which the social planner sets.

The social planner may then implement τ^* by setting the price in the medical market to $p^* = \frac{\tau^*}{\theta}$. Depending upon whether $\frac{\tau^*}{\theta}$ is greater than, less than, or equal to c, $p^* = \frac{\tau^*}{\theta}$ will be greater than, less than, or equal to c. In particular, for particularly generous insurance coverage it is likely that the optimal medical price will be in excess of marginal costs. In this setting, the standard intuition that adding a distortion in the form of supra-marginal cost pricing is optimal is correct.

IV. The Second Best With a Responsive Insurance Market

While the results in the preceding section are intuitive, the setup that generates them imposes an odd passivity upon the operation of the insurance market. We turn now to a social planner who can affect only the price in the medical market but who faces an insurance industry that responds competitively to whatever policy the social planner takes. The results above are now reversed: monopoly power is never optimal in the medical market, even in the presence of coinsurance.

For the propositions below, the medical market's price is set (by fiat or by an oligopoly or competitive equilibrium), and the medical producers supply whatever quantity is demanded at some constant marginal cost, c. The insurance industry is competitive, thus we assume that it chooses insurance policies (τ, m) that maximize consumer welfare, conditional on a break-even constraint.[8]

---

[8] We assume that consumers obtain insurance only from one insurer at the terms specified in the insurance contract, i.e., sellers control both the price and quantity of insurance. Arnott and Stiglitz (1991) have termed this an

A. Consumer Welfare

In Proposition 1, we show that, if consumers do not own the medical firms, they benefit from a price decline in the medical market. In the discussion afterwards, we show that the result of Proposition 1 holds even for a monopolist insurer. Proposition 2 then shows that consumers benefit (i.e. social welfare rises) even if consumers do own the medical firms. Again, the discussion afterwards shows that this result applies to a monopolist insurer as well.

Proposition 1:

Assume the following: the insurance market is competitive, consumers do not own the medical firms, and U is non-decreasing in its first argument. Then consumers are (weakly) better off if price decreases in the medical market.

Proof:

The problem of the competitive insurance industry is to:

$$\begin{aligned} &\max_{\tau,m} EV(Y-m,\tau) \\ &s.t. \quad m \geq E\{(p-\tau)x^{*}(\tau,\varepsilon)\} \\ &\qquad 0 \leq \tau \leq p \end{aligned}$$

We will consider a price decline from some price p^1 to p^2. We will show that, for each feasible contract under p^1, there exists a feasible contract under p^2 which leaves consumers at

"exclusive (quantity) contract." A competitive equilibrium exists in this case and is constrained efficient when there is only one consumer good (Pauly, 1974; Arnott and Stiglitz, 1991) least as well off. Let (τ^1, m^1) be a feasible contract under price p^1. There are two possibilities to be considered: $\tau^1 \le p^2$ and $\tau^1 > p^2$.

First, suppose $\tau^1 \le p^2$. Now consider a contract under price $p^2 < p^1$. Choose $m^1 = m^2$, $\tau^1 = \tau^2$. Observe that the consumer makes precisely the same choice of x, pays the same amount for it out of pocket, and pays the same premium as under the optimal contract at price p^1. Thus, he is as well off under this contract as he was under the contract at price p^1. Since $p^2 \le p^1$ this new contract is feasible (it does not violate the insurance industry's break-even constraint) at the new price.

Now suppose that $\tau^1 > p^2$. Consider moving the consumer to a situation of no insurance, $\tau^2 = p^2$, $m^2 = 0$. Clearly this is a feasible choice for the insurer. Now, we check that the consumer is better off at this contract and the lower price than he was at the old contract and price. Consider the consumer's optimization problem again: $\max_x U(Y - \tau x - m, x, \varepsilon)$. Under the new price and contract, the consumer faces a lower effective price ($\tau^1 > p^2$) and a higher effective income ($m^1 \ge 0$) for each realization of ε. Standard results from consumer theory imply that this will not decrease utility if U is non-decreasing in its first argument, as we have assumed. Q.E.D.

This proposition establishes that consumer welfare is decreasing in the price of the medical good. This implies that competition is indeed second-best optimal, from the consumer's point of view, in medical markets in the presence of moral hazard.

Another result worth mentioning is that consumers' welfare is decreasing in medical price even if the insurance industry is monopolized.[9] The insurance industry then solves:

$$
\begin{aligned}
&\max_{\tau,m} \{m-(p-\tau)E\{x^{*}(\tau,\varepsilon)\}\} \\
&s.t. \quad 0 \le \tau \le p \\
&\qquad EV(Y-m,\tau) \ge EV(Y,p)
\end{aligned}
$$

As long as utility is increasing in its first argument, the "participation constraint" will bind at the solution to this problem. Since a decrease in price clearly increases $EV(Y,p)$, a decrease in price will improve consumers' welfare.[10]

B. Social Welfare

A price decrease in the medical market may benefit consumers; however, it is also likely to harm producers in the medical industry, i.e. profits are likely to fall. To consider the impacts on social welfare we obviously want to see whether the harm to producers is smaller than the benefit to consumers of a price decline.

To avoid analytical difficulties, we substantially restrict the form of the utility function to eliminate income effects:

[9] We assume that the insurance industry has no power over price in the medical market, i.e., there is no exercise of monopsony power. Also notice that the insurer expropriates all consumer surplus, since consumers' demand for insurance is a step function in its price (m).

[10] We note that this result follows trivially from the nature of insurance demand in this model. Since consumers either buy a fixed quantity of insurance or none at all, monopoly has no effects on quantity and thus no efficiency effects. The monopolist simply raises its price to extract the entire risk premium from consumers.

$$U(Y-\tau x-m+\pi_i+\pi_m,x,\varepsilon)=v(Y-\tau x-m+\pi_i+\pi_m+g(x,\varepsilon))$$

$$v(\mathrm{x}) = -\exp(-\mathrm{rx})$$
$$g_1>0, g_2<0, g_{12}>0, g_{11}<0$$

The separability inside v guarantees that there are no income effects *ex post*. The assumption of exponential form for v guarantees that there are no income effects *ex ante*. The two assumptions together are the familiar "no income effects" formulation used frequently in partial equilibrium welfare analysis (Willig, 1976). As also is typical, we assume that income is great enough that consumers will never wish to spend all of their income on the medical good. The assumptions on g capture the fact that ε represents "medical need." Higher ε causes both lower utility and an increase in the efficacy of the medical good in producing utility.

The competitive insurance industry's problem is as above:

$$\begin{aligned}&\max\nolimits_{\tau,m} EV(Y-m,\tau)\\ s.t.\quad & m\geq E\{(p-\tau)x^*(\tau,\varepsilon)\}\\ & 0\leq\tau\leq p\end{aligned}$$

This problem has a solution since the Theorem of the Maximum ensures that the insurance industry is optimizing a continuous function $\{E\{v(-m(\tau,p)+\max_x\{-\tau x+g(x,\varepsilon)\})\}\}$ over a compact set $[0,p]$. Furthermore, we show below in Lemma 1 that the optimal effective price is $\tau<p$.

The insurance industry's problem, in general, is not concave, nor is it quasi-concave. Furthermore, without additional assumptions, we cannot rule out $\tau=0$ as a solution. Since different choices of (consumer-optimal) τ lead to different levels of social welfare through their effects upon π_m, levels of welfare will be determined by which τ is chosen for each price. In the following, we will assume that $\tau(p)$ is a selection from the optimal correspondence.

For a particular price in the medical market, p, and profits in the medical market, π_m, welfare is:

$$W(p, Y, \pi_m) \quad = \exp(-r(Y + \pi_m)) \max_{\tau \in [0,p]} \{E\{v(-m(\tau, p) + \max_x \{-\tau x + g(x, \varepsilon)\})\}\}$$

Of course, π_m is determined in equilibrium, as $\pi_m = (p-c)x^*(\tau)$. By our no income effects assumption, the maximizing τ is not affected by the level of π_m. We write the (selection from the set of) optimal insurance contract $(\tau^*(p), m^*(p))$. Then welfare at a price p is defined to be:

$$W(p) \quad = W(p, Y, (p-c)x^*(\tau^*))$$

We now establish a series of lemmas we then use to prove our main proposition. We first adapt the following result of Zeckhauser (1970) to our setting:

Lemma 1: In any optimal insurance contract, $\tau \in [0, p)$.

Proof: See the Appendix.

For the next lemma, fix a price $p_0 > 0$ and (inductively) construct sequences $\{p_n\}$ and $\{\tau_n\}$ as follows. Choose $\tau_n \in \arg\max_{\tau \in [0,p_n]} \{E\{v(-m(\tau, p_n) + \max_x \{-\tau x + g(x, \varepsilon)\})\}\}$. Set $p_{n+1} = \tau_n$.

Lemma 2: $p_n \to 0$

Proof: See the Appendix.

Lemma 3: Any $\tau(p)$, a selection from $\arg\max_{\tau\in[0,p]} E\{U(Y-m(\tau,p)-\tau x(\tau,\varepsilon), x(\tau,\varepsilon), \varepsilon)\}$,

$m(\tau,p)=(p-\tau)E\{x(\tau,\varepsilon)\}$, with a utility function having no income effects, is non-decreasing in p.

Proof: See the Appendix.

Lemma 4: Fix $p^1 > c$. Let (τ^1, m^1) be an associated optimal contract and π_m^1 the medical industry profit. Let p^2 be a price satisfying: $p^2 > c$ and $\tau^1 \le p^2 \le p^1$ with associated optimal contract (τ^2, m^2) and medical industry profit π_m^2. Then, $W(p^1) \le W(p^2)$.

Proof: See the Appendix.

We now establish the main result of this section: that welfare is declining in price, for prices greater than marginal cost.

Proposition 2:

If $0 \le c \le p \le p'$, then $W(p) \ge W(p')$.

Proof:

Construct sequences, $\{p_n\}$ and $\{\tau_n\}$, as in the discussion before Lemma 2, using p' as the starting point. Since $p_n \to 0$, $p_n < p$ eventually. Find the least n, call it N, for which $p_n < p$. We claim that $W(p') = W(p_0) \le W(p_1) \le W(p_2) \le \ldots \le W(p_{N-1}) \le W(p)$. The first equality is by the definition of $\{p_n\}$. Consider any arbitrary inequality in the list, except the last one. Since $p_n = \tau_{n-1}$, we can simply apply Lemma 4 directly to prove the inequality. For the final inequality, we know by the choice of N and the construction of $\{p_n\}$ that $\tau_{N-1} = p_N \leq p \leq p_{N-1}$. The inequality follows by Lemma 4 again. Q.E.D.

This proposition establishes that, under the usual partial equilibrium assumptions, the usual result that welfare is decreasing in price for prices greater than marginal cost holds. This is the central result of the paper. It establishes that moral hazard in medical markets is not, *per se*, an argument for prices higher than marginal costs in the medical market; thus it is not an argument for laxity in antitrust enforcement or for blockading entry in medical markets.

Although we do not offer an explicit argument, this result follows for a monopolist insurer as well, again because the monopolist insurer captures all surplus. In addition, it can also be shown that marginal cost prices are socially optimal if the insurer is a managed care organization who sets quantity given to a consumer based upon a signal of ε.[11]

V. Consumers' Surplus

In the previous section we showed that consumers' welfare rises with a decline in medical prices in general and that social welfare rises with a decline in medical price when $p > c$. In partial equilibrium analysis without moral hazard, we normally "integrate under the demand curve" in order to assess the effect on consumers of a price change. It might seem, however, that this procedure could lead to an incorrect estimate of consumer surplus in the presence of moral hazard.

In this section we address the question of whether this procedure of "integrating under the demand curve" leads to an accurate assessment of the effect of price changes upon consumers in

[11] Proofs of both of these propositions are available on request from the authors.

the presence of moral hazard (and an insurance market which responds to price changes in the medical market). This is particularly relevant for the economic analysis of mergers. To balance the effect of any price increase against any beneficial effects of merger, a technique for analyzing consumers' surplus and its changes is needed.

The results we derive here will be local in nature. Since there may be multiple optima to the insurance industry's problem and since a change in price will affect each of these optima differently, our results apply only to welfare changes at local optima. Since we are maintaining our partial equilibrium focus, we maintain the "no income effects" specification introduced above.

A consumer with income Y and facing medical price p has utility:

$$W(p,Y) = \exp(-r(Y))\max_{\tau\in[0,p]}\{E\{v(-m(\tau,p)+\max_x\{-\tau x+g(x,\varepsilon)\})\}\}$$

$$W(p,Y) = \exp(-r(Y))W(p),$$

where the second equation serves to define $W(p)$.

We define consumers' surplus at a price p as the compensating variation which would have to be paid to the consumer in order to make him indifferent between purchasing insurance and consuming in the medical market when the price is p and taking the income and foregoing consumption in the medical market. So, consumers' surplus at a price p is the income Y' for which:

$$W(p,Y) = \exp(-r(Y+Y'))E\{v(g(0,\varepsilon))\}$$

$$CS(p) = Y' = \frac{1}{r}\Big(\ln\big(E\{-v(g(0,\varepsilon))\}\big)-\ln\big(-W(p)\big)\Big)$$

$$CS(p) = \frac{1}{r}\Big(\ln\big(-W(\infty)\big)-\ln\big(-W(p)\big)\Big)$$

where the last equality serves to define $W(\infty)$.

To begin with, fix a price, p , and let τ be an interior maximum to the insurance industry's problem at which second order conditions hold strictly. Now, we differentiate $CS(p)$:

$$\frac{\partial}{\partial p}CS(p) = \frac{1}{r}\frac{-1}{-W(p)}\left(-\frac{\partial}{\partial p}W(p)\right) = \frac{1}{r}\frac{1}{-W(p)}\frac{\partial}{\partial p}W(p)$$

$$W(p) = \max_{\tau\in[0,p]}\{E\{v(-m(\tau,p)+\max_x\{-\tau x+g(x,\varepsilon)\})\}\}$$

$$\frac{\partial}{\partial p}W(p) = -E\{v'(\bullet)E\{x^*\}$$

(applying the envelope theorem)

$$\frac{\partial}{\partial p}W(p) = -E\{v'(\bullet)\}E\{x\}$$

Returning to the change in consumer's surplus:

$$\frac{\partial}{\partial p}CS(p) = \frac{1}{r}\frac{1}{-W(p)}\frac{\partial}{\partial p}W(p)$$

$$= \frac{1}{r}\frac{1}{-W(p)}\left(-E\{v'(\bullet)\}E\{x\}\right)$$

$$= \frac{1}{r}\frac{1}{-E\{v(\bullet)\}}\left(-E\{v'(\bullet)\}E\{x\}\right)$$

$$= \frac{1}{r}\frac{1}{-E\{v(\bullet)\}}\left(rE\{v(\bullet)\}E\{x\}\right)$$

$$= -E\{x\}$$

So the change in consumers' surplus for a small change in price is equal to the expected level of demand. This is exactly as in the case of industries without moral hazard in consumption.

There is one thing to note about this result. If the result is to be used to "integrate under the demand curve" it is important to be clear what "demand curve" means here. The correct demand curve to integrate under is $x(\tau(p))$, which, in general, will be different than $x(\theta p)$ where $\theta = \frac{\tau}{p}$ is the current level of coinsurance. That is the demand curve to be integrated under is not the economic primitive which we normally call the demand curve. Rather the demand curve to be integrated under is the relationship between quantity demanded and price, taking into account the behavior of the insurance industry's response to the medical price.

This distinction affects the appropriate demand elasticity to use in calculations. If $\eta^x = \frac{\tau}{x(\tau)} \frac{\partial}{\partial \tau} x(\tau)$ is the conventional demand elasticity and $\eta^\tau = \frac{p}{\tau} \frac{\partial}{\partial p} \tau(p)$ is the elasticity of effective price with respect to medical industry price, then the demand elasticity to be used in evaluating changes in consumers' surplus is $\eta^x \eta^\tau$.

These results are summarized in Figure 2, where we analyze a price decline from P_1 to P_2. The "right" demand curve is $D(\tau(p))$. This represents the quantity demanded in the medical market, assuming that the insurance company sets its coinsurance rate optimally at each price. The curve $D(\theta p)$ is the quantity which would be demanded if the insurance company were to maintain its P_1-optimal coinsurance rate at all prices. Obviously, these two curves must intersect at P_1, by definition, and at 0, where coinsurance is irrelevant. At other prices, the optimal coinsurance is different from the P_1-optimal coinsurance rate, so that the curves are different.

On inspecting this figure, it becomes apparent how our results change the previous intuition. Since the undistorted demand curve, $D(p)$, is not relevant for evaluating welfare from an *ex ante* perspective, the area A is not the welfare loss caused by the moral hazard, it is only the cost of the information asymmetry --- a cost which has a countervailing benefit in the form of risk reduction. The area B is simply not relevant; it represents the cost that would accrue were the insurance industry not to change its coinsurance policy in response to the price decrease. Since the insurance industry will respond to the price change by changing its coinsurance level (in our example by decreasing it), B refers to an irrelevant counterfactual. Using $D(\tau(p))$, we see that the price decrease from P^1 to marginal costs results in a welfare gain corresponding to the Harberger triangle, C.[12]

VI. Summary and Conclusions

In this paper we have considered what the appropriate competitive benchmark is for medical markets in the presence of moral hazard. Moral hazard due to insurance introduces a distortion into the medical market that requires analysis of the second-best. In the presence of moral hazard due to health insurance, consumers will demand "too much" medical care *ex post.* However, contrary to the conventional wisdom, if insurance markets are competitive, or possibly even if they are monopolized, consumers benefit from reduced prices in the medical market. Furthermore, provided that price exceeds marginal cost in the medical market, the benefit to consumers of a price decrease outweighs the loss in profits suffered by the medical industry. So, presuming that competition causes prices to fall in the medical industry, the mere existence of moral hazard should not cast doubt on the general intuition that more competition is socially beneficial.

Our results have several implications for policy. First, supposing that the insurance market is competitive, the existence of moral hazard is not *per se* an argument in favor of lax antitrust enforcement or of erecting barriers to entry in the medical market. Second, if there are welfare gains to a merger on the producers' side arising, say, from cost savings, quality improvements, or increased benefits to indigents and if economic welfare is to be used to evaluate the merger, then the harm to consumers of any price increase can be measured in the usual way, by "integrating under the demand curve."

[12] Note that the hypotenuse of C is the demand curve $D(\tau(p))$ between P_1 and P_2.

We must apply some caveats to these conclusions, however. In this paper we have only analyzed one of the distortions in medical markets: moral hazard. We have not considered other factors that are commonly cited in rendering competition in medical markets different: risk selection in insurance markets, agency problems in medical markets (i.e., induced demand), and the presence of not-for-profit firms. It remains for future research to consider the constellation of these imperfections in concert.

References

Arnott, Richard and Joseph E. Stiglitz (1991) "Equilibrium in Insurance Markets with Moral Hazard," Working Paper No. 3588, Cambridge, MA: National Bureau of Economic Research, Inc.

Arrow, Kenneth J. (1963) "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review*, December, 53: 941-973.

Beazoglou, Tryfon and Dennis Heffley, (1995) "Optimal Market Structures in an Insured Health Care Market: Some Implications for Health Care Reform," unpublished paper, University of Connecticut Health Center. Also published in Greek in *Society, Economy, and Health*, 4(1): 5-30, Jan-Mar 1995.

Beazoglou, Tryfon, Heffley, Dennis, Niakas, Demetrios, and John Kyriopoulos, (1993) "Efficiency in Insured Health Care Markets." unpublished paper, University of Connecticut Health Center. Also published in Greek in *Journal of Health Care Sciences*, 2(3): 3-18, September 1993.

Bingaman, Anne K. (1995) "The Importance of Antitrust in Health Care," *Utah Law Review*, 1995(2): 373-380.

Borch, Karl (1968) *The Economics of Uncertainty*, Princeton, N.J.: Princeton University Press.

Buchanan, James M. (1969) "External Diseconomies, Corrective Taxes, and Market Structure," *American Economic Review*, 59: 174-177.

Crew, Michael (1969) "Coinsurance and the Welfare Economics of Medical Care," *American Economic Review*, December, 59(5): 906-908.

Dor, Avi and John A. Rizzo, (1995) "The Welfare Effects of Mergers in a Service Industry." unpublished paper, Yale University.

Feldman, Roger, and Bryan Dowd (1991) "A New Estimate of the Welfare Loss of Excess Health Insurance," *American Economic Review*, 81(1): 297-301.

Feldstein, Martin S. (1973) "The Welfare Loss of Excess Health Insurance," *Journal of Political Economy*, 81: 251-80.

Frech, H.E. III (1979) "Market Power in Health Insurance, Effects on Insurance and Medical Markets," *Journal of Industrial Economics*, September, 28(1): 55-72.

Frech, H.E. III (1996) *Competition and Monopoly in Medical Care*, Washington, DC: AEI Press.

Gaynor, Martin and Deborah Haas-Wilson (forthcoming) "Change, Consolidation, and Competition in Health Care Markets," *Journal of Economic Perspectives*.

Lynk, William J. (1995) "Nonprofit Hospital Mergers and the Exercise of Market Power," *Journal of Law and Economics*, 38: 437-459.

Ma, Ching-to Albert and Thomas G. McGuire (1997) "Optimal Health Insurance and Provider Payment," *American Economic Review*, 87: 685-704.

Ma, Ching-to Albert and Michael H. Riordan (1997) "Health Insurance, Moral Hazard, and Managed Care," unpublished manuscript, Boston University.

Milgrom, Paul and Chris Shannon, (1994) "Monotone Comparative Statics," *Econometrica*, 62(1): 157-180.

Pauly, Mark V. (1968) "The Economics of Moral Hazard: Comment," *American Economic Review*, 58: 531-537.

Pauly, Mark V. (1974) "Overinsurance and Public Provision of Insurance: The Roles of Moral Hazard and Adverse Selection," *Quarterly Journal of Economics*, 88: 44-62.

Pauly, Mark V. (1998) "Managed Care, Markets, and Monopsony," unpublished manuscript, University of Pennsylvania.

Phelps, Charles E. (1973) *Demand for Health Insurance: A Theoretical and Empirical Investigation*, R-1054-OEO, Santa Monica, CA: Rand Corporation, July 1973.

Polinsky, A. Mitchell and William P. Rogerson (1983) "Products Liability, Consumer Misperceptions, and Market Power," *Bell Journal of Economics*, Autumn, 14(2): 581-589.

Robinson, James C. and Harold S. Luft (1985) "The Impact of Hospital Market Structure on Patient Volume, Average Length of Stay, and the Cost of Care," *Journal of Health Economics*, 4: 333-356.

Shapiro, Carl (1982) "Consumer Information, Product Quality, and Seller Reputation," *Bell Journal of Economics*, Spring, 13(1): 20-35.

Varian, Hal, R. (1984) *Microeconomic Analysis*, 2nd Edition, New York, NY: W.W. Norton.

Weller, Charles D. (1983) "The Primacy of Standard Antitrust Analysis in Health Care," *Toledo Law Review*, Spring, 14: 609-637.

Willig, Robert (1976) "Consumer's Surplus without Apology," *American Economic Review*, 66: 589-597.

Zeckhauser, Richard (1970) "Medical Insurance: A Case Study of the Tradeoff between Risk Spreading and Appropriate Incentives," *Journal of Economic Theory*, 2: 10-26.

Appendix: Proofs

Proof of Lemma 1:

Since the constraint on the insurance company is $\tau \in [0, p]$, we show that $\tau = p$ cannot be optimal.

Since $EV(Y - m(\tau), \tau)$ is differentiable by the functional form assumptions and the assumptions on

g, it will suffice to show that $\frac{\partial}{\partial \tau} EV(y - m(\tau), \tau)$ is negative at $\tau = p$:

$$\frac{\partial}{\partial \tau} EV(y - m(\tau), \tau) = E\left\{\frac{\partial}{\partial \tau} v(Y - m(\tau) + \max_x \{-\tau x + g(x, \varepsilon)\})\right\}$$

$$E\left\{v'(\bullet)\left[-\frac{\partial}{\partial \tau} m(\tau) - x^*(\tau, \varepsilon)\right]\right\}$$

(using the envelope theorem)

Now, we evaluate $\frac{\partial}{\partial \tau} m(\tau)$ at $\tau = p$.

$$m(\tau) = E\{(p - \tau)x^*(\tau, \varepsilon)$$

$$\frac{\partial}{\partial \tau} m(\tau) = E\left\{-x^*(\tau, \varepsilon) + (p - \tau)\frac{\partial}{\partial \tau} x^*(\tau, \varepsilon)\right\}$$

$$\left.\frac{\partial}{\partial \tau} m(\tau)\right|_{\tau = p} = -E\{x^*(\tau, \varepsilon)$$

Substituting . . .

$$\left.\frac{\partial}{\partial \tau} EV(y - m(\tau), \tau)\right|_{\tau = p} = E\{v'(\bullet)[E\{x^*(p, \varepsilon)\} - x^*(p, \varepsilon)]$$

$$= -Cov(v'(\bullet), x^*)$$

$$< 0$$

The last inequality follows from the assumptions on g and v. Since $g_{12} > 0, g_{11} < 0$, x^* must be increasing in ε. However, since $g_2 < 0$, $\max_x\{-px + g(x,\varepsilon)\}$ must be decreasing in ε. Thus, the argument of v is decreasing in ε. Since v is concave, v' is increasing in ε. Thus both v' and x^* are increasing in ε, establishing the inequality. Q.E.D.

Proof of Lemma 2:

Suppose the contrary. Then, since $p_n \geq 0$ and p_n non-increasing (since $\tau_n \in [0, p_n]$), $p_n \to p_* > 0$. But this implies that $\tau_n \to p_*$, for $p_{n+1} = \tau_n$. The Theorem of the Maximum then establishes that $p_* \in \arg\max_{\tau \in [0,p_*]}\{E\{v(-m(\tau, p_*) + \max_x\{-\tau x + g(x,\varepsilon)\})\}\}$. This contradicts Lemma 1. Thus $p_n \to 0$. Q.E.D.

Proof of Lemma 3:

Notice that $E\{U(Y - m(\tau, p) - \tau x(\tau,\varepsilon), x(\tau,\varepsilon), \varepsilon)\} = \exp(-r(Y - m(\tau, p)))E\{U(-\tau x(\tau,\varepsilon), x(\tau,\varepsilon), \varepsilon)\}$, by our assumptions on the utility function, and that the values of this function are negative. Since $-\frac{1}{r}\ln(-x)$ is a strictly increasing function we can define

$\hat{U}(Y - m(\tau, p) - \tau x(\tau,\varepsilon), x(\tau,\varepsilon), \varepsilon) = -\frac{1}{r}\ln(-E\{U(Y - m(\tau, p) - \tau x(\tau,\varepsilon), x(\tau,\varepsilon), \varepsilon)\})$, and maximize it instead. The solutions to $\max_{\tau \in [0,p]} E\{U(Y - m(\tau, p) - \tau x(\tau,\varepsilon), x(\tau,\varepsilon), \varepsilon)\}$ and $\max_{\tau \in [0,p]} Y - m(\tau, p) - \frac{1}{r}\ln(-E\{U(-\tau x(\tau,\varepsilon), x(\tau,\varepsilon), \varepsilon)\})$ are the same. The cross partial of $\hat{U}$ with respect to p and τ is non-negative, at $-\frac{\partial^2 m}{\partial p \partial \tau} = -\frac{\partial x(\tau,\varepsilon)}{\partial \tau}$.

Let τ be the largest maximizer at price p, then $\forall \tilde{\tau} \le \tau, \hat{U}(\tau, p) - \hat{U}(\tilde{\tau}, p) \ge 0$. Consider, now, price $p' > p$. The cross-partial between p and τ positive implies that $\forall \tilde{\tau} \le \tau, \hat{U}(\tau, p') - \hat{U}(\tilde{\tau}, p') \ge \hat{U}(\tau, p) - \hat{U}(\tilde{\tau}, p) \ge 0$. Also, τ continues to be feasible at p'. Thus, any maximizer under $p' > p$ must be greater than or equal to τ. Q.E.D.

Proof of Lemma 4:

We seek to show that:

$$
\begin{aligned}
W(p^1) &= E\{v(Y + \pi_m^1 - m^1 + \max_x\{-\tau^1 x + g(x,\varepsilon)\})\} \\
&\le \max_{\tau \in [0,p^1]} E\{v(Y + \pi_m^1 - (p^1 - p^2)E\{x^*(\tau^1,\varepsilon)\} - m(\tau, p^2) + \max_x\{-\tau x + g(x,\varepsilon)\})\} \\
&= E\{v(Y + \pi_m^1 - (p^1 - p^2)E\{x^*(\tau^1,\varepsilon)\} - m^2 + \max_x\{-\tau^2 x + g(x,\varepsilon)\})\} \\
&\le E\{v(Y + \pi_m^2 - m^2 + \max_x\{-\tau^2 x + g(x,\varepsilon)\})\} \\
&= W(p^2)
\end{aligned}
$$

The first inequality follows since, by setting $\tau = \tau^1$, the objective function in the second line is equal to the first line. The maximization assumption then guarantees the inequality. To show the second inequality, it will suffice to show that $\pi_m^1 - (p^1 - p^2)Ex^*(\tau^1, \varepsilon) \le \pi_m^2$. Consider the following:

$$
\begin{aligned}
\pi_m^1 - (p^1 - p^2)Ex^*(\tau^1,\varepsilon) &= (p^1 - c)Ex^*(\tau^1,\varepsilon) - (p^1 - p^2)Ex^*(\tau^1,\varepsilon) \\
&= (p^2 - c)Ex^*(\tau^1,\varepsilon) \\
&\le (p^2 - c)Ex^*(\tau^2,\varepsilon) \\
&= \pi_m^2
\end{aligned}
$$

The inequality follows since $\tau(p)$ is non-decreasing and demand is downward-sloping (no income effects). Q.E.D.

Note that the argument in Lemma 4 is a simple supermodularity argument (see Milgrom & Shannon, 1994, Theorem 4′).

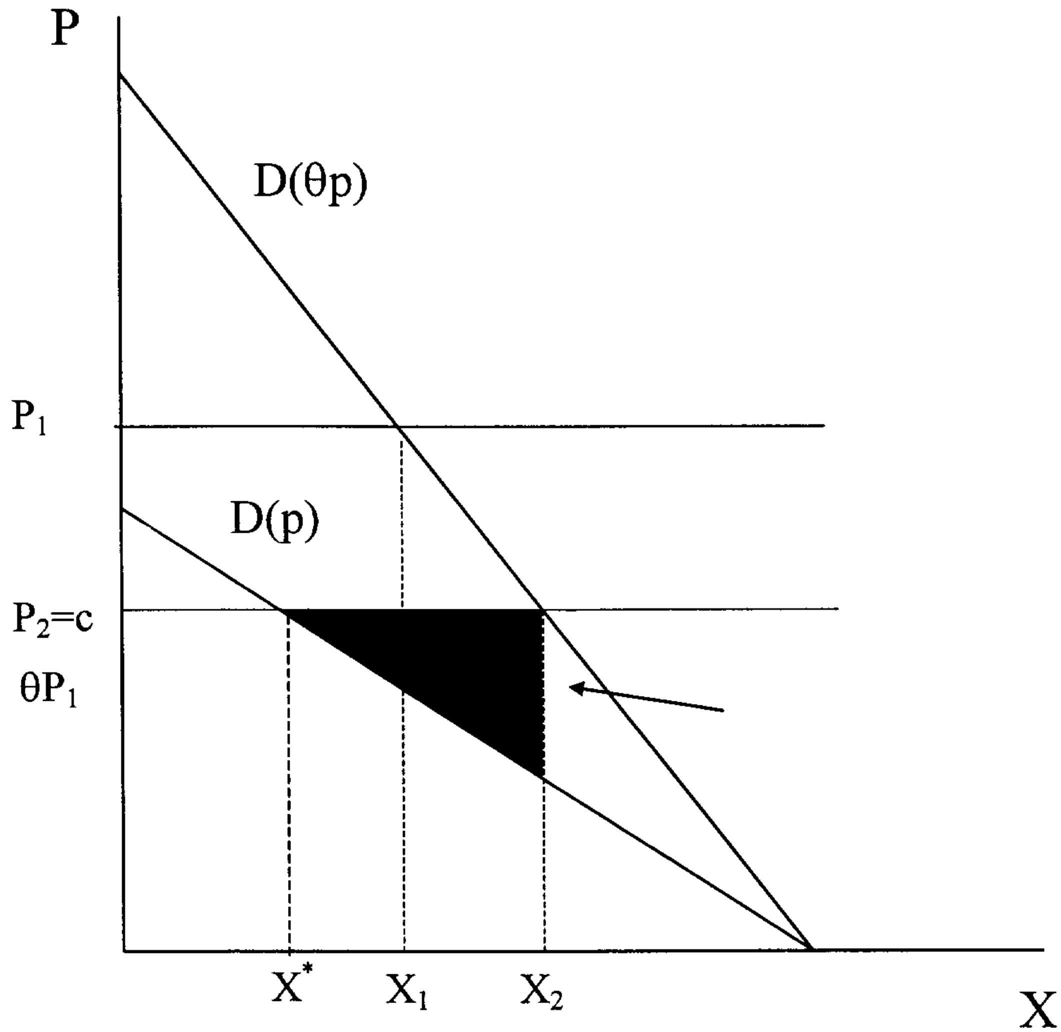

Figure 1: Standard Analysis

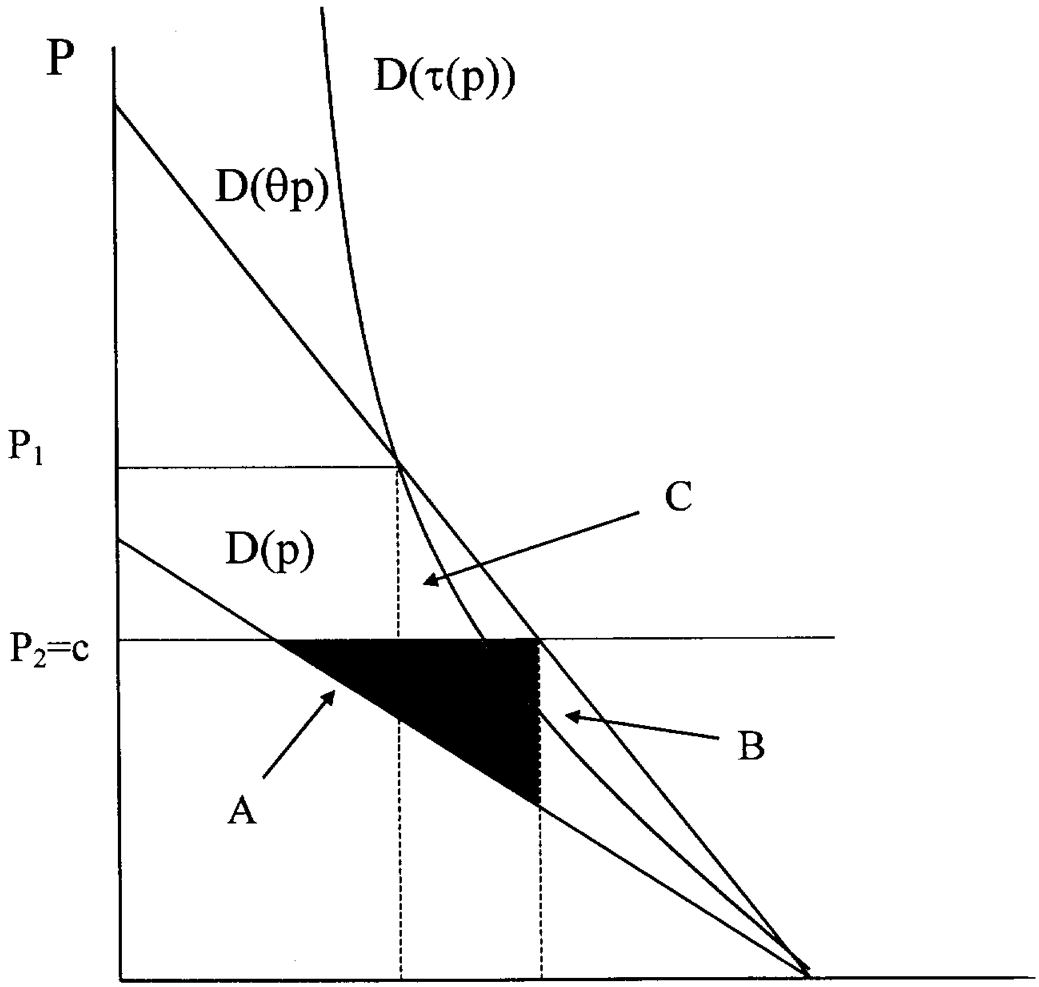

Figure 2: Flexible Insurance