Health Insurance as a Two-Part Pricing Contract [*]

Darius Lakdawalla and Neeraj Sood
RAND Corporation and NBER

May 20, 2009

Abstract

Monopolies appear throughout health care, as a result of patents, limits to the extent of the market, or the presence of unique inputs and skills. Health insurance, often implemented as an ex ante premium coupled with an ex post co-payment per unit consumed, effectively operates as a two-part pricing contract that allows monopolists to extract profits from consumers without inefficiently constraining quantity. The efficiency of this downstream nonlinear pricing arrangement alters the welfare analysis of monopoly in the context of health care. At the limit, frictionless and competitive health insurance markets, even under incomplete or asymmetric information, perfectly eliminate deadweight losses from upstream monopoly in health care. Frictions limiting downstream "premium-discrimination" restore some of these upstream monopoly losses, which will then manifest as uninsurance for the healthy (or poor), rather than direct restriction of health care quantity. Empirical analysis of pharmaceutical patent expiration is consistent with the prediction that heavily insured health care markets experience little to no efficiency loss under monopoly, while less insured markets exhibit behavior more consistent with standard theories of monopoly and its associated deadweight loss.

---

[*] For helpful comments and suggestions, we wish to thank Alan Garber, Dana Goldman, Geoffrey Joyce, Tomas Philipson, Bob Town, Bill Vogt, participants in the NBER Summer Institute Health Economics Workshop and NBER Health Care Program Meetings, seminar participants at the University of Illinois, Federal Trade Commission, Indiana University Purdue University at Indianapolis, and the University of Chicago. We are grateful to the National Institute on Aging for funding.

A. Introduction

Optimal health insurance contracts balance risk-sharing against the need for efficient utilization incentives (Arrow, 1963; Pauly, 1968; Zeckhauser, 1970). This balance explains why such contracts charge an ex post unit price or co-payment. Co-payments reduce the degree of insurance, but in return produce fewer distortions in the goods market, because the consumer faces a private price that partially reflects social cost.

The trade-off between risk-sharing and incentives has been widely studied. However, health insurance contracts have another function that is less well-appreciated: the reduction of deadweight loss. Health insurance resembles a two-part pricing contract, in which a group of consumers pays an upfront fee in exchange for lower prices in the event of illness. Two-part pricing contracts allow a monopolist to sell goods at marginal cost, but extract consumer surplus in the form of an upfront payment (see the seminal paper by Oi, 1971). Typically, competition improves social welfare, because it minimizes deadweight loss. In the presence of two-part pricing, however, the monopolist has the same incentive to minimize deadweight loss, because such a strategy maximizes the social surplus available for extraction through higher premiums.

An example illustrates this mechanism in the insurance context. Setting a co-payment equal to marginal cost and a premium equal to expected consumer surplus would, for instance, allow a firm to extract the maximum possible surplus, while still ensuring efficient utilization and zero deadweight loss. The uncertainty of health care demand, coupled with ex ante or ex post asymmetric information, *creates* a contractual structure that better aligns the incentives of monopolists with those of society. At the limit, idealized, frictionless, and competitive health insurance markets will perfectly eliminate deadweight losses from market power in health care provision. This logic is robust to typical informational failures like moral hazard and adverse selection.

More generally, downstream nonlinear pricing mechanisms limit the deadweight loss from upstream monopoly. Therefore, frictions that interfere with the efficiency of the downstream nonlinear pricing instrument generate deadweight losses from upstream market power. This suggests an important complementarity between the efficiency of downstream health insurance markets and that of goods and service providers in health care. Specifically, deadweight losses from health care monopoly arise when regulations or other frictions compel health insurers to offer pooled and uniform contracts to heterogeneous consumers, and thus limit their ability to extract consumer surplus. Even if insurers are competitive, the resulting downward-sloping demand for insurance creates deadweight losses from "markups" on insurance premiums that are passed through from upstream providers. These societal losses take the form of higher rates of uninsurance for low willingness-to-pay customers, namely the healthy or the poor.

Our results have several important implications for evaluating monopoly loss in health care markets. First, the extent and even presence of deadweight losses from health care monopoly are determined by the structure of the insurance market. Therefore, the decision to regulate or allow monopoly in health care must be driven in large part by the structure of the health insurance market that mediates the provision of goods. Moreover, extending insurance coverage or promoting efficiency in the insurance market may be viewed as substitutes for regulating monopoly in health care markets.

Second, the price-cost margin is a less reliable measure of welfare loss in markets where health insurance plays a role. When insurance is widespread and reasonably complete, providers may be receiving very high monopoly prices and profits, even though consumers are paying prices near or even below marginal cost. The copayment-cost margin is a similarly inconsistent measure, which is driven primarily by the structure of information in the insurance market, and not by the monopoly power of upstream providers. The only reliable measures of market power are the change in quantity it induces, or, the change in rates of uninsurance.

Finally, health insurance lowers the static deadweight losses associated with patent monopolies. Health care markets may thus not be as subject to the usual societal trade-off between incentives to innovate and static inefficiency.

Empirical analysis of patent expiration in the pharmaceutical market provides evidence consistent with these results. First, the elimination of patent monopoly has little to no impact on quantity consumed for molecules that are heavily insured, but substantial quantity impact for molecules with less widespread insurance. Second, even though market quantity barely falls for heavily insured molecules, the profits earned by patent monopolists fall by more in these markets. Third, the standard theory of monopoly is quite consistent with the observed behavior of markets for uninsured health care products, but quite inconsistent with equilibrium behavior in highly insured markets.

Section B develops our argument that competitive and frictionless insurance markets eliminate deadweight loss from monopoly, even when information is incomplete, and market power is imperfect. Section C explores how contracting frictions and other incentives for insurance pooling ultimately drive deadweight losses from health care monopoly. Section D analyzes the implications for innovation and patent protection. Section E presents our empirical analysis. Finally, Section F summarizes our conclusions and implications for the analysis of market power in health care.

B. Two-Part Health Insurance and Surplus-Extraction

Any insurer who can charge both a premium ex ante and a co-payment ex post has enough tools to extract maximum consumer surplus and ensure efficient utilization. We first make this point in the context of a standard model with moral hazard. Our initial setup is very similar to that of Gaynor, Haas-Wilson, and Vogt (2000), who show that reductions in the price of medical care benefit consumers even in the presence of moral hazard.

Consumers face a risk of illness, and an uncertain demand for a medical remedy, produced at constant marginal cost MC. An insurance contract is an offer of an ex post co-payment (m), and a premium I. There are consumers of measure one, indexed by $h \in [0,1]$, and distributed uniformly over this interval. Patients with lower values of h are sicker. The fraction σ of consumers falls sick. Sick consumers place value on the medical remedy, while healthy consumers do not.

Consumers do not know their value of h ex ante, but learn it ex post. Insurers, however, cannot observe this value and thus cannot make indemnity payments conditional on the underlying health state. Payments can only be contingent on the consumer's observed decision to purchase the medical good or not. The necessity of tying payments to utilization, rather than the underlying source of risk, results in "over-utilization" relative to the first-best, full information case. This is a second-best means of delivering some additional insurance in the face of informational incompleteness.

B.1 The Typical Competitive Problem

We begin by characterizing the standard competitive equilibrium allocation in the presence of moral hazard. Consider a representative competitive insurer purchasing medical care from a competitive goods market selling at marginal cost, and providing insurance within the informational structure outlined above. The case of non-competitive insurers produces largely similar results, with a few important exceptions.[1] The firm chooses a co-payment and premium that maximizes consumer utility, subject to a break-even constraint, and incentive compatibility for the consumer. The insurer

[1] We take competitive insurance markets as the base case, since most privately insured beneficiaries are enrolled in self-insured plans, which by definition do not have mark-ups (except perhaps in administration) (Employee Benefit Research Institute, 2009). In addition, many believe that private insurance markets are competitive, due to the existence of several commercial insurers and the option to self-insure (Federal Trade Commission, 2004). However, since recent evidence suggests that insurer's might have market power in several important contexts (Dafny, 2008), we consider this case of bilateral market power in Appendix Section A.3.

knows the quantity of medical care demanded by consumer h, given the co-payment and income, according to $q(W-I,m,h)$.

The firm's optimization problem can be written as:

$$\begin{aligned} &\max_{I,m\le MC} \int_0^1 u(W-I-mq^*,q^*,h)dh \\ &s.t.\ I+(m-MC)E(q^*)\ge 0 \\ &and\ q^*=q(W-I,m,h) \end{aligned} \tag{1}$$

Associating the multiplier μ with the break-even constraint, the first-order conditions can be expressed as:

$$\begin{aligned} &[I]: Eu_W=\mu\left(1-(m-MC)E(q_W)\right) \\ &[m]: \int_0^1 \frac{q}{E(q)}u_W(W-I-mq,q,h)dh=\mu\left(1-(m-MC)\frac{-E(q_m)}{E(q)}\right) \end{aligned} \tag{2}$$

These first-order conditions illustrate the standard trade-off between risk-bearing and incentives in the presence of moral hazard. The left-hand side of the first order condition for m always exceeds the left-hand side of the condition for I, because u_W and q^* are decreasing in h.[2] This fact, coupled with the two first-order conditions, implies that

$$(m-MC)\left(\frac{E(q_m)}{E(q)}+E(q_W)\right)>0 \tag{3}$$

Observe that $E(q_m)+E(q)E(q_W)$ is the expected effect on q of a Hicksian-compensated increase in the co-payment m. Since the compensated demand for medical care is downward-sloping,[3] it

---

[2] $\int_0^1 \frac{q}{E(q)}u_W(W-I-mq,q,h)dh$ is a weighted average of u_W, where more weight is placed on its larger values.

[3] The consumer's first-order condition for medical care consumption is $E(-mu_W+u_q)=0$. The first-order effect of a compensated increase in the co-payment is an increase in m, but no change in u_W or u_q. Therefore, the consumer must decrease medical care consumption in response.

follows that $\left(\frac{E(q_m)}{E(q)} + E(q_W)\right) < 0$, and $m < MC$. In turn, this implies that the marginal utility of wealth will be higher than in the first-best, according to the first-order condition for insurance.

Intuitively, when ex post information is asymmetric, the only way to provide insurance is to induce over-utilization by charging the consumer a price below marginal cost. Therefore, the benefits of insurance must be traded off against the cost of inducing distortion in the goods market. This leads to: (1) Over-utilization relative to the first-best, (2) Higher marginal utility of wealth relative to first-best, and (3) Incomplete insurance.

B.2 Two-Part Health Insurance with Monopoly

Two-part health insurance eliminates deadweight losses associated with monopoly, but it cannot solve the intrinsic informational problems that lead to moral hazard in this environment. As a result, a monopolist with access to two-part health insurance pricing will choose an allocation of resources identical to the competitive benchmark described above. For initial simplicity, we consider an insurer who is also the monopoly provider of the good with uncertain demand. Later, we separate the two functions. Moreover, while we consider the simpler case of pure uncontested monopoly in the text, Appendix A generalizes our results to the case of incomplete market power.

The insurer-provider maximizes profits subject to a reservation utility condition for the consumer. Define $\overline{U}$ as the level of utility the consumer would attain if he refused the insurance contract and failed to consume the medical care good. However, he may still have a claim on the firm's profits if he is a shareholder. The firm thus solves:

$$\begin{aligned} &\overline{\pi} \equiv \max_{I,m} I + (m - MC)E(q(W - I, m, h)) \\ &s.t. \int_0^1 u(W - I - mq + \pi, q, h)dh \geq \overline{U} \end{aligned} \tag{4}$$

It is straightforward to prove the mechanical equivalence between this problem and the competitive one. Since $\overline{\pi}$ is the optimal monopoly profit, the above problem is equivalent to:

$$\max_{I,m} \int u(W - I - mq + \bar{\pi}, q, h)dh$$
$$s.t. I + (m - MC)E(q) \geq \bar{\pi} \tag{5}$$

By displacement, this can be rewritten as:

$$\max_{m} \int_0^1 u(W - m(q - E(q)) - MC * E(q), q, h)dh, \tag{6}$$

which is exactly equivalent to the displaced version of the problem in 1. Intuitively, the monopolist can best maximize his own profits by first maximizing gross consumer surplus, and then extracting it in the form of an upfront payment.

B.3 Separating the Insurance- and Goods-Producers

If a downstream insurer has access to nonlinear pricing, in the form of a two-part tariff, a separate upstream monopolist goods-producer will be able to extract maximum surplus as long as the goods-producer can negotiate a different price-quantity pair with each insurer. A monopolist forced to name a single (linear) price for the entire market could not exploit the availability of nonlinear pricing downstream, and the standard deadweight loss from monopoly would accrue (Gaynor et al., 2000).[4]

Price-discrimination across insurers may occur in a number of ways, all of which are observed in health care markets, where quantities are either explicitly named (e.g., by a pharmaceutical wholesaler), tied to a nonlinear price schedule (e.g., in the form of quantity discounts, rebates, and the like), or where prices are separately negotiated with providers of different sizes. To offer a few examples, contracts between PBMs and pharmaceutical firms are of two types – non-

[4] They show that even in the presence of moral hazard, consumers are better off with competition (lower prices) than with monopoly (higher prices).

capitated and capitated.[5] Non-capitated contracts usually specify a list price or "wholesale acquisition cost" and terms for determining discounts or rebates. Rebates are usually tied to the dollar or unit sales of a particular drug product. Growth rebates offer PBMs a steeper discount if they achieve certain volume targets. Capitated contracts, on the other hand, specify a fixed payment from the PBM to the drug company per insured member per month, along with some risk-sharing arrangement that determines additional payments or concessions based on actual drug usage (Levy, 1999). The capitated rates combined with risk-sharing arrangements effectively render these equivalent to two-part pricing contracts. Outside the pharmaceutical industry, health insurers routinely negotiate different prices with different hospitals and providers, even for similar products and services. These negotiated prices depend on provider size, insurer size, the form of the insurer, and other market characteristics (Zwanziger and Melnick, 1988; Dranove et al., 1993; Keeler et al., 1999; Dor et al., 2004).

Suppose the monopolist provider is interacting with a representative insurer, standing in for a competitive insurance industry. The monopolist now offers a quantity Q and fixed-fee F that solve the following profit-maximization problem:

$$\begin{aligned} &\max_{Q,F} F - MC * Q \\ &s.t. F \leq G(Q) \end{aligned} \tag{7}$$

The monopolist cannot extract more than $G(Q)$, defined as the maximum amount of gross profit the insurer is able to earn on the quantity Q.

[5] Private-sector entities that offer prescription drug insurance coverage, such as employers, labor unions, and managed care companies, often hire pharmacy benefit managers (PBMs) to manage these insurance benefits. PBMs engage in many activities to manage their clients' prescription drug insurance coverage including assembling a network of retail pharmacies, designing the plan formulary and cost sharing arrangements (co-payments for different drugs) and negotiating with pharmaceutical companies.

Gross profits can be characterized as the maximum amount the insurer can extract from consumers at that expected quantity level, or:

$$\begin{aligned} &G(Q) = \max_{I,m} I + m\int_0^1 q(m, W-I, h)dh \\ &s.t. \int_0^1 u(W - I - mq(m, W-I, h) + \pi, q(m, W-I, h), h)dh \geq \overline{U} \\ &and \int_0^1 q(m, W-I, h)dh \leq Q \end{aligned} \tag{8}$$

Defining η as the Lagrange multiplier on the second constraint, the insurer's problem can be rewritten as:

$$\begin{aligned} &G(Q) = \max_{I,m} I + m\int_0^1 q(m, W-I, h)dh + \eta(Q - \int_0^1 q(m, W-I, h)dh) \\ &s.t. \int_0^1 u(W - I - mq(m, W-I, h) + \pi, q(m, W-I, h), h)dh \geq \overline{U} \end{aligned} \tag{9}$$

According to the envelope theorem, $G'(Q) = \eta$ in this problem. Moreover, the first-order condition for the monopolist implies that $G'(Q) = MC$. Since this implies $\eta = MC$ in equilibrium, we can write:

$$\begin{aligned} &G(Q) = \max_{I,m} I + (m - MC)\int_0^1 q(m, W-I, h)dh + MC * Q \\ &s.t. \int_0^1 u(W - I - mq(m, W-I, h) + \pi, q(m, W-I, h), h)dh \geq \overline{U} \end{aligned} \tag{10}$$

The equilibrium values of I and m are unaffected by the term $MC * Q$. Therefore, the problem above is equivalent the integrated insurer's problem in equation 4. This demonstrates the result that monopolists who can tie quantity to a fixed payment will induce insurers to behave as if they were goods-monopolists themselves.

These results have two important implications for measuring monopoly loss. First, the price-cost margin, or $\frac{F}{Q * MC}$, provides little insight into the extent of deadweight loss. F may be much higher than costs even when deadweight loss is zero. Even the copayment-cost margin, or $\frac{m}{MC}$ is an

inconsistent measure of monopoly loss *per se*, since the optimal copayment is determined by the extent of moral hazard (or, as we show later, of adverse selection). Upstream monopolists with market power have no incentive to request or promote high copayments downstream.

To keep things analytically simple, we removed all aggregate uncertainty from this problem, in the sense that utilization is certain at the insurer level. This is without loss of generality, because the provider can always offer a menu of quantity choices conditional on the health outcomes of the insured population. It would then extract the fixed-fee equal to the expected surplus associated with those offers. Alternatively, it can even attempt to harvest *actual* surplus associated with each outcome, provided it can vary prices with the realized outcomes of the insured population. This would be the case if, for instance, the provider could offer rebates tied to utilization.

B.4 Adverse Selection

The basic logic of health insurance as two-part pricing also holds up under another common failing of insurance markets — adverse selection.[6] Our analysis of moral hazard demonstrated that a monopolist with access to two-part health insurance can replicate the competitive equilibrium with moral hazard, or incomplete *ex post* information. Adverse selection adds incomplete *ex ante* information. In this case, the insurer can observe neither the severity of illness ex post, nor the ex ante differences in the propensity of consumers to fall ill.

As in the case of moral hazard, two-part pricing cannot remove the deadweight loss associated with asymmetric information, but it does remove all the incremental deadweight loss associated with monopoly. In other words, a monopolist with access to the two-part contract will do just as well as a competitive market.

---

[6] This analysis is related to Stiglitz's (1977) finding that, in the presence of adverse selection, monopoly within insurance markets can dominate competition.

We assume there are chronically ill patients (type C), and not chronically ill patients (type N). Firms cannot observe consumer types. Define $\mu^C(h)$ and $\mu^N(h)$ as the measures of chronically ill and not chronically ill people across the interval $h \in [0,1]$. The distribution for the chronically ill is assumed to dominate the other in the first-order stochastic sense. An insurance contract is an ex ante insurance premium (I), coupled with an ex post copayment (m). The appendix demonstrates that, under these circumstances, competition is Pareto-equivalent to monopoly, when two-part health insurance contracts are used.

C. Pooling Equilibria

The results for adverse selection rely on the canonical Rothschild-Stiglitz separating equilibrium response to asymmetric information. Indeed, Rothschild and Stiglitz proved the now widely known result that pooling equilibria are not generally possible in a frictionless, competitive, and unregulated insurance market. However, frictions in the labor market (cf, Crocker and Moran, 2003; Bhattacharya and Vogt, 2006; Fang and Gavazza, 2007), market power among insurers (Stiglitz, 1977), community-rating laws (cf, Adams, 2007), or restrictions on premium-differentiation (cf, Bhattacharya and Bundorf, 2005), can generate distortions that lead to pooling equilibria. In a pooling equilibrium, the premium is a less efficient instrument for surplus-extraction. This can lead to inefficiency and deadweight loss from market power in the provision of medical care, but not of the typical form. Quantity-restriction only occurs as a result of individuals dropping out of the insurance market, not for any individuals who remain insured.

Continue with our two types of consumers — C and N -- the chronically ill and not chronically ill. As before, the C types are sicker and derive more consumer surplus from any given health insurance offer. However, through some combination of legal, informational, or labor market constraints, firm either choose or are compelled to offer a single premium and copayment schedule.

Insurers cannot differentiate between the two types ex ante (adverse selection), nor can they distinguish ex post between patients with different levels of illness severity (moral hazard).

C.1 Competitive Outcomes

We first characterize the competitive pooling equilibrium that serves as a benchmark. Suppose there is some set of characteristics that enables a pooling equilibrium -- labor market frictions, or regulation – where no other distortions exist. Insurance and medical care are both provided by competitive firms. The representative medical care provider sells goods at marginal cost. All consumers have reservation utility levels $\overline{U}$. For completeness, we also endow the two types with s^C and s^N shares of the insurer's profits.

The representative insurer maximizes:

$$\begin{aligned}
&\max_{I,m} \int_0^1 u^N(W - I - mq^N, q^N, h)\mu^N(h)dh \\
&s.t.\ 2I + (m - MC)E(q^C(W - I, m, h)) + (m - MC)E(q^N(W - I, m, h)) \geq 0 \\
&\int_0^1 u^N(W - I - mq^N, q^N, h)\mu^N(h)dh \geq \overline{U} \\
&\int_0^1 u^C(W - I - mq^C, q^C, h)\mu^C(h)dh \geq \overline{U}
\end{aligned}$$

The participation constraint for C types will fail to bind. Suppose it does bind, and that the N types exit the market as a result. Since we assume a pooling equilibrium to be possible by construction, this cannot be an equilibrium generated by competitive firms, who could make both types better off by insuring everyone. Since the participation constraint for C types will fail to bind, and the constraint for the N types is analytically redundant, the benchmark equilibrium solves:

$$\begin{aligned}
&\max_{I,m} \int_0^1 u^N(W - I - mq^N, q^N, h)\mu^N(h)dh \\
&s.t.\ 2I + (m - MC)E(q^C(W - I, m, h)) + (m - MC)E(q^N(W - I, m, h)) \geq 0
\end{aligned} \tag{11}$$

By displacement, this assumes the familiar form:

$$\max_{I,m} \int_0^1 u^N \left[W - m\left(q^N - \frac{E(q^c) - E(q^N)}{2} \right) - MC\left(\frac{E(q^c) - E(q^N)}{2} \right), q^N, h \right] \mu^N(h) dh \quad (12)$$

C.2 Effects of Market Power on Uninsurance

Now consider an integrated monopolist that provides insurance and medical care. The integrated firm offers a single contract to both types. Each type is endowed with s^H and s^L shares of the firm, which solves the following problem:

$$\max_{I,m} 2I + (m - MC)E(q^C(W - I, m, h)) + (m - MC)E(q^N(W - I, m, h))$$
$$s.t. \int_0^1 u^C(W - I - mq^C + s^C\pi, q^C, h)\mu^C(h)dh \geq \overline{U}$$
$$\int_0^1 u^N(W - I - mq^N + s^N\pi, q^N, h)\mu^N(h)dh \geq \overline{U}$$

Define (I, m) as the solution to this problem. Since there is a single contract, the firm will not be able to capture consumer surplus from both these consumer types, and one of the participation constraints will fail to bind. Moreover, if the equilibrium contract extracts all available surplus from the C types, it will violate the participation constraint for the N types. Therefore, there are two cases to consider: (1) the participation constraint binds for the N types, and both types receive insurance; and (2) the participation constraint binds for the C types, and only they receive insurance.

Begin with the first case, where both types are insured. The problem simplifies to:

$$\max_{I,m} 2I + (m - MC)E(q^C(W - I, m, h)) + (m - MC)E(q^N(W - I, m, h))$$
$$s.t. \int_0^1 u^N(W - I - mq^N + s^N\pi, q^N, h)dh \geq \overline{U}$$

Define $\overline{\pi}$ as the equilibrium level of profit that solves this problem. It can then be equivalently rewritten as:

$$\max_{I,m} \int_0^1 u^N(W - I - mq^N + s^N\overline{\pi}, q^N, h)dh$$
$$s.t. \; 2I + (m - MC)E(q^C(W - I, m, h)) + (m - MC)E(q^N(W - I, m, h)) \geq \overline{\pi}$$

Substituting the reservation profit constraint into the objective function transforms this into an unconstrained maximization problem, as follows:

$$\max_{I,m} \int_0^1 u^N (W - MC\left(\frac{E(q^C)+E(q^N)}{2}\right) - m\left(q^N - \frac{E(q^C)+E(q^N)}{2}\right) + s^N \overline{\pi} - \frac{\overline{\pi}}{2}, q^N, h) dh$$

When both types own equal shares of the firm, this problem is equivalent to the formulation in equation 12. Clearly, monopoly may have distributional consequences, if the shares of the firm are not distributed in the prescribed manner, but it remains (second-best) Pareto-optimal in the usual sense.

However, efficiency losses can occur in the case where there are incentives for uninsurance. Define (I^{N*}, m^{N*}) as the optimal insurance contract that would be offered to type N in the absence of type C consumers. And define (I^{C*}, m^{C*}) similarly. When both types remain insured, the firm offers the contract (I^{N*}, m^{N*}). However, this *may* generate fewer profits than the firm could earn by offering (I^{C*}, m^{C*}) to the C types alone. In this case, the firm will choose to price the N types out of the market. This is the only efficiency consequence of distortionary pooling equilibria. The healthier (or poorer) consumers with less demand for insurance end up uninsured, where they face a linear monopoly price and all its corresponding deadweight losses. However, the C types continue to enjoy their second-best quantity levels. Deadweight loss from monopoly is equal to the welfare lost by those who are "priced out" of the insurance market. The empirical consequences of health care monopoly for uninsurance have been considered by Town et al (2006).[7]

[7] Several papers find that hospital mergers significantly increase prices and more concentrated hospital markets have higher prices (Gaynor and Vogt, 2000; Capps et al., 2003; Gaynor and Vogt, 2003). Despite this evidence of a substantial price effect, Town et al (2006) find that the wave of hospital mergers in the 1990's that dramatically increased concentration resulted in only a modest increase in rates of uninsurance (0.3 percentage points).

D. Innovation

A major reason for monopolies in health care is the use of patents to encourage innovation, which typically occurs at the expense of static inefficiency from monopoly loss. Increases in patent length or scope encourages greater innovation, but also leads to greater static inefficiency through the creation of monopoly loss (Shavell and van Ypersele, 1998).[8] The presence of health insurance changes this social welfare calculus for patent monopolies in health care markets. Expanding insurance coverage reduces static inefficiency from patent monopolies, by providing innovators with access to a two-part pricing instrument for a greater number of consumers. It also raises the profits they can earn on their innovations. The latter may increase dynamic efficiency when innovation is under-provided, but reduce dynamic efficiency when innovation is excessive. When innovation is excessive, insurance market expansions may need to be accompanied by tighter restrictions on patent length and scope, to restore efficient dynamic incentives.

D.1 Static Efficiency of Health Care Patents

Patent monopolies in health care are analytically identical to the monopolies examined earlier. To be specific, the static allocation under patent monopolies is identical to the perfectly competitive allocation of goods, so long as the innovator's profits are shared equally among consumers using the innovative goods. This result is proven more formally in Appendix C.

If patent monopolies – e.g., pharmaceutical companies – are able to set different unit prices for customers utilizing different quantities, they have enough tools to eliminate deadweight loss in the insured market. However, deadweight monopoly loss can still occur as a result of increased rates of uninsurance. The social welfare parameter of interest is the effect of increases in health care

[8] However, see the work by Boldrin and Levine (http://www.dklevine.com/general/intellectual/against.htm), who argue that patents are not essential for innovation.

patent length or scope on rates of uninsurance. Theoretically, this effect will be nonnegative, but it could be zero, at least in principle, if the health insurance market is frictionless.

D.2 Dynamic Incentives

When insurance and innovation markets are frictionless and efficient, the socially optimal level of compensation for the innovator is equal to the consumer surplus generated by the new innovation (see, for instance, the simple innovation model in Appendix C). However, there are a variety of plausible market failures that can move the optimal dynamic return in either direction.

First, employer-based health insurance premia are implicitly subsidized, because they are tax-exempt. If consumers face less than the full price of insurance, monopolists will be able to extract consumer surplus *plus* the value of the premium subsidy. However, monopolists will continue to have incentives to set the co-payment so as to maximize extractible consumer surplus. The result is that premium subsidies or taxes primarily affect dynamic efficiency, but not static inefficiency, which the monopolist has incentives to maintain. As Garber, Jones, and Romer (2006) have argued, this logic suggests that premium subsidies lead to over-innovation. If the innovator can extract total surplus, *in addition to* the value of the premium subsidy, the return on innovation is too high relative to first-best. The result is too much innovation, but efficient provision of the innovations that exist.

Second, a more detailed model of the innovation process could also lead to inefficiency. In the standard model used above, innovators ought to appropriate the full value of social surplus. Many analysts have pointed out that patent races, public subsidies, and other imperfections can alter this result, so that innovators ought to receive less than social surplus in the first-best allocation. Others, in contrast, have emphasized how little innovators are able to appropriate.[9] This is a difficult

[9] For contrasting views in the context of pharmaceuticals, see Garber, Jones, and Romer (2006), compared with Philipson and Jena (2006). In a broader context, see Shapiro (2007), compared with Nordhaus (2004). The analytics of patent races appear in Reinganum (1989).

question to resolve in our context, because — outside of the simple model presented above — there are a great many possible models of the innovation process, each with different implications. Depending on the first-best rate of appropriation, access to two-part health insurance pricing may result in inefficiently high profits. This affects the optimal tax-and-transfer policy or patent policy that should accompany a functioning market for health insurance — the social planner can undo incentives to over-innovate by taxing the profits of successful innovators or reducing patent length and scope. Regardless of dynamic incentives, two-part pricing through health insurance continues to ensure static efficiency, although it may require correctives to ensure dynamic efficiency as well.

E. An Empirical Test of Two-Part Pricing Through Insurance

The theory predicts higher deadweight loss from monopoly when rates of insurance are lower. Therefore, eliminating market power in less well-insured markets ought to have greater welfare benefits than in better-insured markets. This result obtains because health insurance allows upstream monopolists to price their products off the consumer demand curve. In this section, we empirically test these two key implications in the market for pharmaceuticals, where patent expiration provides a transparent change in market power for a particular molecular entity. In particular, we predict:

(H-1) Better insured molecules display smaller changes in quantity at patent expiration;

(H-2) Better insured molecules display larger declines in profits, and thus branded drug revenues, at patent expiration;

(H-3) For primarily uninsured drugs, the demand elasticity estimated from patent expiration will be less than -1.0, as the standard theory of monopoly predicts;

(H-4) For primarily insured drugs, the demand elasticity will be smaller than standard theory predicts, and possibly greater than -1.0.

Note that the theoretical predictions obtain, regardless of the source of variation in rates of insurance, which can occur due to insurance market imperfections, upstream market power, public health

insurance eligibility rules, or other institutional factors. Empirically, there are two factors of particular importance in driving pharmaceutical insurance coverage. The first is the proportion of users who are over 65 years of age, since (at the time of the data) Medicare coverage did not include a prescription drug component. For example, the correlation between percent of expense paid by the insurer and proportion of users above 65 years of age is 0.41. The second is hospital-administration (versus self-administration) of drugs. The correlation between percent of expense paid by the insurer and whether the drug is administered in a hospital is 0.40.

E.1 Data

We use the IMS Generic Spectra database for this analysis. These data represent 101 unique molecules, whose patents expired between 1992 and 2002. For each molecule, we use up to 5 years of monthly data, which span the interval from 3 years prior to 2 years after patent expiration. The monthly data include both revenues and quantities measured at the ex manufacturer level, which is the relevant point of measurement for our theory. Drug quantity data, in grams, are collected at both the wholesale and retail level (the latter includes hospital pharmacies). Revenue data are collected at the retail level exclusively. IMS then adjusts the revenue data, using proprietary estimates of drug mark-ups, to estimate the implied ex-manufacturer revenue. We drop 6 drugs with missing sales data. For the 95 remaining drugs in our sample, average monthly quantity was 3.77 million grams and average monthly revenues were $19.3 million. The complete list of molecules appears in Appendix C.

We use two sources of information to determine the insurance status of drugs. First, IMS identifies drugs that are primarily consumed and available in hospitals. These hospital drugs are typically covered by medical insurance rather than prescription drug insurance. Since medical

insurance is both more common and more generous than prescription drug insurance, hospital drugs are covered more generously than other drugs.[10] We had 16 such drugs in our data.

Second, we merged the IMS databases with the 1996 to 2002 Medical Expenditure Panel Survey (MEPS). MEPS contains detailed information on prescription drug use and insurance coverage for a nationally representative population. For each drug we calculate the share of total drug cost paid by the insurer, one year before patent expiration, as a measure of insurance coverage for the drug. However, since MEPS data is available only from 1996 onwards, this measure is only available for the 43 drugs whose patent expired in 1997 or later. The share of expenses borne by insurers varied significantly across these 43 drugs – the first quartile was 41%, the median was 57%, and the third quartile was 78%.

E.2 Empirical Framework

All the hypotheses listed above can be tested in the context of a model that estimates the elasticity of demand for pharmaceuticals, by the degree of insurance. The source of identifying variation for the demand curve is the change in market power induced by the expiration of the molecule's patent.

E.2.1 *Hypotheses 1 and 2: Quantity, and Branded Revenues*

The most direct way to test Hypothesis 1 is with the reduced-form version of this model that regresses patent expiration on quantity. Specifically, we estimate:

$$LnQ = a_1 + b_1 PatExp + c_1 PatExp * Ins + poly(month) + poly(month) * Ins + d + \varepsilon \qquad (13)$$

$$LnQ = a_2 + b_2 PatExp + c_2 PatExp * Hosp + poly(month) + poly(month) * Hosp + d + \varepsilon \qquad (14)$$

LnQ is the log of total grams of the drug sold in a given month. $PatExp$ is an indicator variable for the months after patent expiration. Ins measures the share of drug expenditures borne by insurers the

[10] For example, according to the data from MEPS 2006, 34.9% of expenditures on prescription drug were paid by consumers. In contrast only 3.8% of expenditures on hospital care and 16.3% of expenditures on office based visits were paid by consumers.

year before patent expiration. *Hosp* is an indicator variable for whether the drug is a hospital product. *poly*(*month*) is a cubic polynomial in months since patent expiration, and d is a drug fixed-effect.

The first equation models the increase in the quantity of the drug sold after patent expiration, and how this varies with the degree of insurance for the drug. The cubic in month absorbs a common life-cycle trend in drug sales. The drug fixed-effects absorb time-invariant differences across drugs. Hypothesis 1 predicts $c_1 < 0$. The last equation models whether the change in total quantity following patent expiration is different for hospital drugs. Relying on the fact that hospital drugs are better insured, Hypothesis 1 predicts $c_2 < 0$.

Similarly, to test Hypothesis 2, we run the reduced-form model of patent expiration on branded revenues, as follows:

$$LnR = \alpha_1 + \beta_1 PatExp + \gamma_1 PatExp * Ins + poly(month) + poly(month) * Ins + d + \varepsilon \quad (15)$$

$$LnR = \alpha_1 + \beta_1 PatExp + \gamma_1 PatExp * Hosp + poly(month) + poly(month) * Ins + d + \varepsilon \quad (16)$$

The hypothesis predicts that $\gamma_1 < 0$.

E.2.2 ***Hypotheses 3 and 4: Pricing on the Demand Curve***

To test the final two hypotheses, we estimate the full demand equations, as in:

$$LnQ = A_1 + B_1 LnP + C_1 LnP * Ins + poly(month) + poly(month) * Ins + d + \varepsilon$$
$$LnP = A_2 + D_1 PatExp + D_2 PatExp * Ins + poly(month) + poly(month) * Ins + d + \varepsilon \quad (17)$$

Hypothesis 3 predicts that $B_1 < -1$, since this represents the estimated elasticity of uninsured drugs. Hypothesis 4 predicts that $C_1 > 0$.

E.2.3 ***Identification***

The overall elasticity of demand is identified by the patent expiration, which we argue represents a shift in market power that is independent of contemporaneous and immediate shifts in

the demand curve. While demand may be changing over time, the patent expiration variable in our model captures the discontinuous effect at the month of expiration.

There is a further identification assumption needed in order to recover the interaction between patent expiration and the degree of coverage. We assume that the patent expiration has similar effects on the supply curve for insured and uninsured molecules. We present some evidence supporting this assumption in Section E.4.1. Note that our approach can accommodate differences in the shape of demand for uninsured and insured molecules, as explained below in Section E.4.2.

E.3 Empirical Results

E.3.1 *Market Power, Quantity, and Branded Revenues*

The results from our empirical tests of hypotheses 1 and 2 are presented in Table 1.

Table 1: Results of Empirical Tests

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	log total grams			log brand dollars		
Patent Expired	0.067	0.093	0.249	-0.175	-0.148	-0.018
	[0.020]***	[0.021]***	[0.105]**	[0.029]***	[0.029]***	[0.113]
Hospital Product*Patent Expired		-0.158			-0.165	
		[0.048]***			[0.087]*	
Share Insured*Patent Expired			-0.344			-0.492
			[0.135]**			[0.175]***
Constant	12.323	12.322	12.572	15.968	15.968	16.272
	[0.020]***	[0.020]***	[0.033]***	[0.024]***	[0.024]***	[0.040]***
Observations	5002	5002	2488	5002	5002	2488
Number of drugs	95	95	43	95	95	43

Notes: Robust standard errors in brackets. Standard errors are adjusted for clustering at the drug level. Other covariates in the model include drug fixed effects, cubic polynomial in months since patent expiration and cubic polynomial in months since patent expiration interacted with indicator for hospital drugs or percent of expenses borne by insurer.
* significant at 10%; ** significant at 5%; *** significant at 1%

Models 1 and 4 present benchmark results for the average drug. At patent expiration, the total quantity sold increases by 6.7% for the mean drug, while branded drug revenues decline by 17.5%.

Models 2 and 5 demonstrate that, for hospital-administered products, quantity rises by 16% points less, and branded revenues fall by 17% more. For hospital products, patent expiration is predicted to lower quantity by 6.5%, which is not statistically different from zero, and lower branded revenues by 31%. The significant differences between hospital and other products, and the insignificant change in quantity for hospital products, are all consistent with Hypotheses 1 and 2.

Models 3 and 6 repeat this analysis using share of expenses paid by insurers as an alternative measure of how well a molecule is insured. The coefficient estimates imply that, for molecules paid at the first quartile of insurance generosity (41% of expenses paid by insurers), patent expiration raises quantity by 11%, reduces branded revenues by 22%. In contrast, for molecules at the third quartile of insurance generosity (78% of expenses paid by insurers), patent expiration leaves quantity statistically unchanged (the point estimate is -2.0%), but lowers branded revenues by 40%. These results are also consistent with Hypotheses 1 and 2.

E.3.2 *Pricing on the Demand Curve*

The theory predicts standard monopoly pricing behavior for drugs without insurance, but pricing off the demand curve for drugs that are heavily insured. For uninsured drugs, we expect demand elasticities less than -1.0. For insured drugs, we expect demand elasticities substantially smaller in absolute value, or even positive.

We estimate the price elasticity of demand for drugs by generosity of insurance, using patent expiration as an instrument for Log (price). Other variables in the model are identical to those reported in equations 13 and 14. Results are displayed in Table 2.

Table 2: Price Elasticity of Demand by Insurance Generosity

	Model 1 Log Total Grams	Model 2 Log Total Grams	Model 3 Log Total Grams
Log Price	-0.977	-1.449	-2.641
	[0.348]***	[0.427]***	[1.136]**
Hospital Product*Log Price		2.156	
		[0.605]***	
Share Insurance*Log Price			3.611
			[1.430]**
Constant	15.658	16.571	16.885
	[0.674]***	[0.826]***	[1.441]***
Observations	5002	5002	2488
Number of Drugs	95	95	43

Model 1 shows that the price elasticity of demand for the average drug is approximately -1.0, but there are substantial differences across insurance status. According to Model 2, products administered outside the hospital have demand elasticities of -1.4. In contrast, hospital-administered products have an elasticity of +0.7, which is significantly different from -1. Evidently, hospital-administered drugs display inelastic, or even upward-sloping demand at the equilibrium monopoly quantity.

Similarly, results from Model 3 suggest that drugs with a larger share of expenses borne by insurers have significantly less elastic demand. Drugs at the first quartile of insurance coverage (40% expenses paid by insurers) have a price elasticity of -1.2. In contrast, drugs at the third quartile of insurance coverage (80% expenses paid by insurers) have a dramatically different price elasticity of +0.25. The latter is significantly different from -1.

The price elasticities of drugs with relatively less insurance accord well with the standard theory of monopoly, which predicts that the demand elasticity should be approximately equal to the inverse of the monopoly markup, or that $-\frac{pD'(p)}{D(p)} = \frac{1}{1-\frac{MC}{p}}$. Based on this result, drugs administered outside the hospital would exhibit $\frac{MC}{p} \approx 0.28$, while drugs at the first quartile of

insurance coverage would exhibit $\frac{MC}{p} \approx 0.17$. Both these numbers are within the range of estimates for monopoly markups on pharmaceuticals (Caves et al., 1991). In contrast, we can reject the hypothesis that the elasticities for hospital-administered drugs or drugs in the upper two quartiles are less than or equal to -1.0. This is inconsistent with the standard theory of monopoly. It is similarly hard to reconcile these results with the hypothesis that heterogeneous demands alone can explain the data. If insurance functions as two-part pricing, manufacturers are not required to price on the consumer demand curve, and the standard implications for price and quantity movements no longer obtain.

E.4 Alternative Explanations

E.4.1 *Heterogeneous Entry Costs*

A threat to the validity of our results is a difference in entry costs between insured and uninsured drugs, due to underlying differences in production technology, regulatory hurdles, or differences in the strategic responses of incumbents. For example, if insured products have higher entry costs, patent expiration would lead to less generic entry, and a smaller increase in quantity. (By itself, however, this would not explain why patent expiration leads to larger reductions in branded revenues for these drugs.) To test this hypothesis, we estimate the effect of patent expiration on generic market share, by type of drug. Results from this analysis are presented in Table 3 below.

Table 3: Testing for Differences in Generic Share after Patent Expiration

	Model 1	Model 2	Model 3
	Generic share in quantity		
Patent Expired	0.196	0.202	0.222
	[0.013]***	[0.015]***	[0.049]***
Hospital Product*Patent Expired		-0.032	
		[0.035]	
Share Insured*Patent Expired			0.061
			[0.075]
Constant	0.087	0.087	0.087
	[0.006]***	[0.006]***	[0.009]***
Observations	5002	5002	2488
Number of drugs	95	95	43

Notes: Robust standard errors in brackets
* significant at 10%; ** significant at 5%; *** significant at 1%

Model 1 shows that patent expiration leads to a 19.6 percentage point increase in generic market share. Models 2 and 3 show that both hospital drugs and drugs with generous insurance coverage experience a similar increase in generic market share following patent expiration.

E.4.2 *Heterogeneous Demand Elasticities*

One might wonder if heterogeneous demand elasticities could also threaten the approach here. Our theory predicts that consumers exhibit less elastic demand for insured drugs, due to the presence of insurance. However, consumers may have systematically lower demand elasticities for insured drugs due to differences in their underlying preferences for those drugs, rather than insurance status per se[11].

From a theoretical point of view, however, heterogeneous demand elasticities do not compromise estimates of elasticities on the margin and cannot easily account for the results we observe. For instance, suppose that consumers exhibit systematically less elastic demand for insured drugs, due to differences in underlying preferences for such drugs. On the margin, standard theory

[11] As discussed earlier, the two primary drivers of insurance status are the proportion of elderly users, and hospital-administration of the drug. There is no empirical evidence that documents differences in demand elasticities for drugs either based on proportion of elderly users or site of administration.

predicts that the aggregate market elasticity is purely a function of the monopoly markup, and should be less than -1. This is inconsistent with the elasticity values we observe for highly insured molecules. The same result holds if insured consumers have systematically lower demand elasticities than uninsured consumers. In this case, demand elasticity at a given point would be lower for insured drugs, due to the presence of more insured consumers. However, on the margin and in equilibrium, aggregate elasticity will continue to depend on the markup, and to be less than -1.

F. Conclusions and Implications for Policy

The presence of health insurance alters the welfare analysis of monopoly. Price-cost margins and even copayment-cost margins become unreliable yardsticks of welfare loss, which is more reliably measured in terms of reductions in quantity, or marginal increases in rates of uninsurance. Evidence from pharmaceutical markets suggests that drug therapies with less insurer presence exhibit greater deadweight loss from monopoly and greater gains from the elimination of market power. On the other hand, drugs with greater insurer presence seem to be priced off the consumer demand curve in ways that are inconsistent with the standard theory of monopoly behavior and monopoly loss.

From a normative point of view, the observed structure of health insurance contracts provides a means for achieving the efficient outcomes associated with two-part pricing in goods markets. While it is not a panacea for informational problems in the insurance market, it can be a useful remedy to static deadweight losses from monopoly, as we have shown. The evidence suggests that the penetration of health insurance lowers the deadweight loss associated with market power, and facilitates the efficient extraction of rents by monopolists. In other words, a well-functioning insurance market transforms the problem of market power from one of efficiency into one of distribution.

A review of health care markets in the late 1990's highlights three interrelated trends: an increase in managed care as method of financing and delivering care; horizontal consolidation within

insurer, hospital and physician markets, and the blurring of the vertical distinctions between these markets (Gaynor and Haas-Wilson, 1999; 2002). Our analysis has important implications for analyzing the potential consequences of each of these trends.

First, our analysis suggests that the recent increase in horizontal consolidation and market power of health care providers may or may not reduce social welfare; the impacts depend on the structure and functioning of the relevant health insurance markets. Moreover, if welfare reductions do occur, they will tend to take the form of uninsurance, rather than simple and direct quantity-restrictions by monopolists.

Second, our analysis suggests that the rise in managed care and vertical integration of health care markets experienced in the 1990's provides unique benefits to society. From a positive point of view, our analysis suggests that vertical integration in health care may be motivated in part by the improved ability of an integrated firm to price-discriminate. This can help to explain why some pharmaceutical companies have chosen to invest in pharmacy benefit managers, and why health-maintenance organizations integrate health-care provision with insurance.

Innovation is of obvious importance in health care markets. Our analysis shows that two-part health insurance pricing has important implications for dynamic incentives. Well-functioning insurance markets may help patent monopolists extract the maximum amount of consumer surplus associated with their inventions, without restricting quantity. On the plus side, this limits the static costs associated with patent monopolies. However, this does not speak to the dynamically optimal level of profits and innovation, which may be affected by patent races, government subsidies to the insurance market, or other factors that can lead to excessive profits accruing to innovators. Health insurance may either mitigate or exacerbate dynamic inefficiencies in innovation.

The design of public health insurance often considers the trade-offs among optimal risk-bearing, moral hazard, and adverse selection. However, our analysis suggests that it ought to consider how the two-part health insurance contract can best maximize social surplus. An optimally

designed public health insurance scheme would set co-payments at or below marginal cost, depending on the extent of moral hazard. The division of resources among consumers can then be determined by the schedule of premia, which allows the government to extract as much or as little consumer surplus as it chooses. In innovative markets, two-part health insurance can also be configured to generate any desired change in the incentive for innovation, without compromising static efficiency (Lakdawalla and Sood, 2009).

Appendix

A. Incomplete Market Power

So far, we have considered the case of pure uncontested monopoly. Many health care markets are better approximated by monopolistic competition. For example, two drug companies might hold patents on different drugs that treat the same disease. Doctors may build unique relationships with their patients, who develop a preference for one physician over another. Patients may prefer to go to a hospital that is closest to their home. All these factors can create product differentiation in the minds of consumers. Market power results, but it is incomplete. In this section, we add monopolistic competition to the moral hazard information structure.

Monopolistic competition changes the distribution of resources relative to complete monopoly, but leaves intact the result that monopolistic competitors choose quantity so as to maximize extractible surplus. A monopolistic competitor must be mindful that her customers can defect to the other firm. This limits the amount of surplus available for extraction. However, conditional on consumer purchases from her, she will continue to set quantity so as to maximize their surplus.

To distill the key ideas, suppose we have two monopolistic competitors—A and B—and two kinds of consumers, with one strictly preferring A, but the other strictly preferring B. Both products have the same marginal cost of production. The firms are integrated in the sense that they both produce their goods and provide insurance contracts over them. Further, as with most spatial models of product differentiation, assume that consumers must choose to use one or the other of the products, but not both—these might be different drugs, physicians, or hospitals, which cannot be easily used with those of rivals. Define $u^A(c,q,h)$ as utility for consumers who prefer A and define $u^B(c,q,h)$ similarly. If a consumer uses the "wrong" good, she derives utility $u^i(c,\delta q,h)$,

where $\delta < 1$. Since each consumer can only consume one of the goods, we can assume without loss of generality that insurers provide two insurance contracts—one that provides good A and one that provides good B.

A.1 The Second-Best Efficient Allocation

Clearly, the efficient allocation provides each consumer with her preferred good, and its associated insurance contract. Goods are sold at marginal cost to the insurer. Each contract maximizes the utility of the consumer, subject to the break-even constraint of the insurer. As before, the insurer knows the quantity of good j demanded by a consumer of type j in health state h, given the co-payment and income, according to $q^j(W - I, m, h)$.

The optimal contract for the type j consumer maximizes:

$$\begin{aligned} &\max_{I^j, m^j \le MC} \int_0^1 u^j(W - I^j - m^j q^{j*}, q^{j*}, h)dh \\ &s.t.\ I^j + (m^j - MC)E(q^{j*}) \ge 0 \\ &and\ q^{j*} = q^j(W - I^j, m^j, h) \end{aligned} \tag{18}$$

This problem is identical to the earlier case of moral hazard, and has a similar solution, characterized by:

$$\begin{aligned} &[I]: Eu_W^j = \mu\left(1 - (m^j - MC)E(q_W^j)\right) \\ &[m]: \int_0^1 \frac{q^j}{E(q^j)} u_W^j(W - I^j - m^j q^j, q^j, h)dh = \mu\left(1 - (m^j - MC)\frac{-E(q_m^j)}{E(q^j)}\right) \end{aligned} \tag{19}$$

The insurer sets a co-payment below marginal cost, in an effort to provide some insurance.

A.2 Equilibrium with Monopolistic Competition

The key difference between monopolistic competition and the earlier case of pure monopoly is in the consumer's reservation utility level. The pure monopolist had only to guarantee the consumer as much utility as she could derive without consuming any medical care goods. The

monopolistic competitor, on the other hand, has to guarantee the utility she could derive from the competitor's contract. As with most models of oligopoly, this reservation utility level depends on the absence, presence, and nature of strategic behavior between competitors. However, this does not affect the marginal valuation of goods, only the level of profit earned by the firm. The division of resources among the two firms and the set of consumers have no impact on efficiency. Indeed, if type j consumers own firm j, all profits extracted are returned to the consumers from which they were taken. The result is the same equilibrium observed under pure competition.

Without loss of generality, we will demonstrate this reasoning for firm A. Define $q^{BA}(W - I^B, m^B, h)$ as the amount of good B that consumer A will use when offered the good B insurance contract. Firm A then solves:

$$\begin{aligned} &\max_{I^A, m^A} I^A + (m^A - MC)E(q^A) \\ &s.t. \int_0^1 u^A(W - m^A q^{A*} - I^A + \pi^A, q^A, h)dh \geq \\ &\int_0^1 u^A(W - m^B q^{B0} - I^B + \pi^A, q^B, h)dh \\ &q^{A*} = q^A(W - I^A, m^A, h) \\ &q^{B0} = q^{BA}(W - I^B, m^B, h) \end{aligned} \tag{20}$$

The decisionmaking of the other firm only enters insofar as it affects the consumer's reservation utility level. If consumers own their respective firms, this will not even affect the distribution of resources.

Arguing as we did in the case of moral hazard, define $\bar{\pi}^A$ as the optimal level of profit that solves the firm's problem. The problem in 20 can be equivalently written as:

$$\begin{aligned} &\max_{I^A, m^A} \int_0^1 u^A(W - m^A q^{A*} - I^A + \bar{\pi}^A, q^A, h)dh \\ &s.t.\ I^A + (m^A - MC)E(q^A) \geq \bar{\pi}^A \\ &q^{A*} = q^A(W - I^A, m^A, h) \end{aligned} \tag{21}$$

The displaced version of this problem is identical to the displaced version of the competitive problem in 18. This demonstrates that monopolistic competition produces the same allocation as pure competition.

A.3 Two-Sided Market Power

Outcomes may not be perfectly efficient if both the insurance and provider sides are characterized by incomplete market power. In the case of bilateral monopoly with Nash bargaining, both sides continue to have incentives to maximize consumer surplus, and then divide it between themselves according to their respective bargaining power (Chipty and Snyder, 1999). However, imperfect competition on both sides can sometimes create unique distortions. For example, there may be strategic incentives for exclusive dealing, which is often observed in pharmaceutical and hospital markets (Ellison and Snyder, 2003; Ho, 2006). Such exclusive dealing would lower consumer welfare by restricting consumer choices to certain set of hospitals or drugs.

This point can be demonstrated with a simple example. Suppose there are ten therapies to treat a single disease, and two innovators — innovator A sells 9 of these therapies, while innovator B sells only one. There is a single insurer, and ten consumers. Each consumer derives $10 of surplus from the therapy she prefers: Nine consumers prefer one of A's therapies, while the tenth prefers B's therapy. Suppose innovator A demands an exclusive contract with the insurer. This is a credible demand if all $90 of consumer surplus is extracted, and if the innovator gives the insurer $15 of this surplus. Innovator B cannot match the offer. The result is that utilization of B's therapy is inefficient, because the patient preferring B can only buy it directly, and not through an insurance policy. This leads to the typical monopoly problem, and the under-provision it commonly implies.

This suggests that market power in the insurance industry may be the root cause of inefficient utilization, rather than market power on the provider/innovator side. This also suggests that inefficiency in the insurance industry will cut against the ability of insurance to produce perfectly

efficient outcomes. Nonetheless, in actual practice, private insurance contracts provide significant price reductions on a large number of therapies and treatments, even though the largest price reductions might be reserved for a few "preferred" drugs or providers. Even so, the actual price discounts observed lead to significant reductions in deadweight loss, and improvements in efficiency (Ho, 2006).

B. Adverse Selection

To model adverse selection, suppose that consumers are heterogeneous ex ante. There are chronically ill patients (type C), and not chronically ill patients (type N). Firms cannot observe consumer types. Define $\mu^C(h)$ and $\mu^N(h)$ as the measures of chronically ill and not chronically ill people, respectively. The distribution for the chronically ill is assumed to dominate the other in the first-order stochastic sense. An insurance contract is an ex ante insurance premium (I), coupled with an ex post copayment (m).

B.1 The Competitive Solution

A pooling equilibrium is not possible for the usual reasons (Rothschild and Stiglitz, 1976): given any putative pooling equilibrium, there is always a profitable contract that attracts only the low-risk insureds. Therefore, if an equilibrium exists, it must be a separating equilibrium. As such, the competitive insurance industry chooses two contracts that maximize the welfare of each type of agent, subject to incentive compatibility constraints (ensuring the contracts are chosen by the correct agents), and break-even constraints. The contract (m^C, I^C) for the chronically ill solves:

$$
\begin{aligned}
&\max_{m^C, I^C} \int_0^1 u(W - I^C - m^C q^C, q^C, h)\mu^C(h)dh \\
&s.t. \\
&[\gamma]: \int_0^1 u(W - I^N - m^N q^N, q^N, h)\mu^N(h)dh \\
&\geq \int_0^1 u(W - I^C - m^C q^C, q^C, h)\mu^N(h)dh \\
&[\beta]: I^C + \int_0^1 q^C \mu^C(h)dh(m^C - p) \geq 0 \\
&q^N \equiv q(W - I^N, m^N, h), q^C \equiv q(W - I^C, m^C, h)
\end{aligned}
\tag{22}
$$

This problem has the following first-order conditions:

$$
\begin{aligned}
&[I]: E_C(u_W) - \gamma E_N(u_W) = \beta(1 - E_C(q_W)(m^N - p)) \\
&[m]: E_C(u_W \frac{q^C}{E_C(q^C)}) - \gamma E_N(u_W \frac{q^C}{E_C(q^C)}) = \beta(1 - \frac{E_C(-q_m)}{E_C(q_C)}(m^C - p))
\end{aligned}
\tag{23}
$$

Notice that if the incentive constraint fails to bind, these first-order conditions are identical to the second-best equilibrium with moral hazard.

This observation reveals how the adverse selection equilibrium is affected by the introduction of moral hazard. In the absence of moral hazard, full insurance is the benchmark outcome. Full insurance is never incentive-compatible, because high-risk consumers always prefer the full insurance contract offered to the lower-risk, lower-cost consumers. This explains why, in the standard Rothschild-Stiglitz setting, adverse selection always impacts outcomes.

In the presence of moral hazard, however, the addition of adverse selection may be welfare-neutral, because the second-best moral hazard contracts may be incentive-compatible. Suppose, for example, that the second-best contract involves a very high copayment for the low-risks, because they have a highly elastic demand and relatively little insurable risk. If so, it is possible that the high-risk insureds would prefer their own second-best contract to that offered to the low-risks. In this event, adverse selection would have no impact, because incentive compatibility emerges of its own accord, due to moral hazard. This would leave us with the moral hazard equilibrium outlined above. If, however, the second-best contracts are not incentive-compatible, we obtain the typical Rothschild-

Stiglitz solution in which the high-risk consumers receive their second-best contract, but the low-risk consumers receive something worse than their second-best.

The indirect utility conferred by a specific contract is defined by $v^C(I,m)$ and $v^N(I,m)$ for the chronically ill and not chronically ill patients, respectively; these are defined as follows.

$$v(I,m) \equiv \max_q \int_0^1 u(W-I-mq(W-I,m,h),q(W-I,m,h),h)\mu(h)dh \tag{24}$$

We impose two assumptions that make this environment similar to the Rothschild-Stiglitz one. First, the chronically ill are willing to pay more for a given change in the copayment rate, in the sense that:

$$-\frac{dI}{dm}|_{v^C} > -\frac{dI}{dm}|_{v^N} \tag{25}$$

This is the typical "single-crossing" property from Rothschild and Stiglitz's (1976) analysis of adverse selection.[12] Second, a given change in the co-payment rate has a bigger impact on a firm's profits, so that:

$$\begin{aligned} -\frac{dI}{dm}|_{\pi=0} &= \frac{E(q)+(MC-m)E(q_m)}{E(q_W)(MC-m)} \\ -\frac{dI}{dm}|_{\pi^C=0} &> -\frac{dI}{dm}|_{\pi^N=0} \end{aligned} \tag{26}$$

Figure 1 illustrates the separating equilibrium in (I,m)-space. The curves Z^N and Z^C represent the zero-profit curves for the not chronically ill and chronically ill, respectively. v^C is the indifference curve for the chronically ill tangent to the zero-profit line — this represents the optimal (i.e., second-best) contract that is possible under moral hazard. Observe that if the second-best

---

[12] $-\frac{dI}{dm} = \frac{E(u_w q)}{E(u_w)}$. First-order stochastic dominance implies that the numerator is higher for the chronically ill. We assume this effect outweighs the fact that the marginal utility of wealth may also be higher for the chronically ill.

contract for the not chronically ill falls on the curve segment A, there is no adverse selection problem, because both second-best contracts are incentive-compatible.

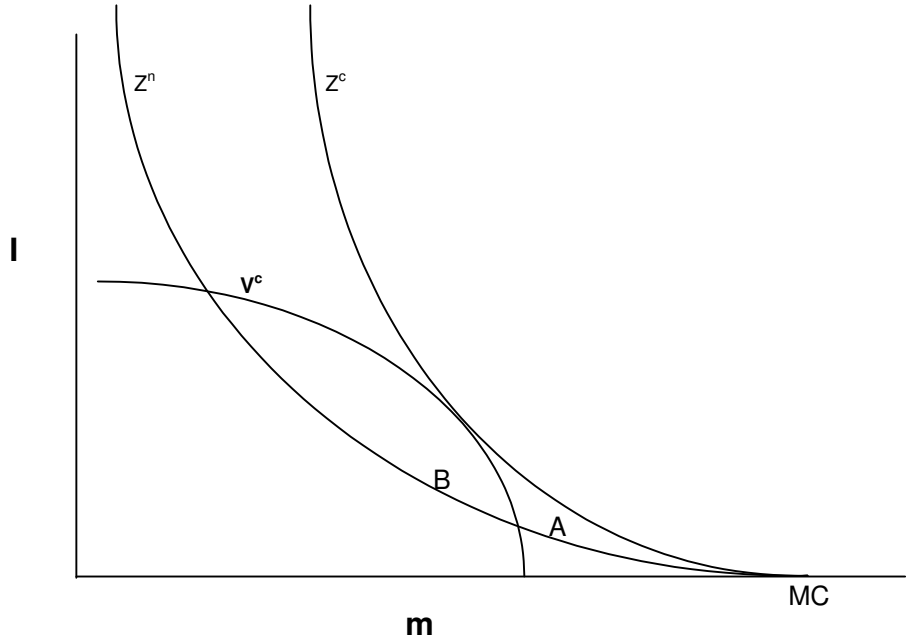

Figure 1: Equilibrium with adverse selection and moral hazard.

Now consider the case where adverse selection has an impact: if the second-best contract for type N falls on the curve segment B. In this case, the chronically ill will receive their second-best contract, while the other type will receive the contract at the intersection of v^C and Z^N.

B.2 Equilibrium with Two-Part Monopoly Pricing

A monopolist who charges an upfront premium and an ex post copayment maximizes profits subject to reservation utility conditions (i.e., participation constraints) and incentive constraints.

$$
\begin{aligned}
&\max_{m^C, I^C, m^N, I^N} I^C + (m^C - MC)\int_0^1 q^C \mu^C(h)dh + I^N + (m^N - MC)\int_0^1 q^N \mu^N(h)dh \\
&s.t. \\
&\int_0^1 u(W - I^N - m^N q^N, q^N, h)\mu^N(h)dh \\
&\geq \int_0^1 u(W - I^C - m^C q^C, q^C, h)\mu^N(h)dh \\
&\int_0^1 u(W - I^C - m^C q^C, q^C, h)\mu^C(h)dh \\
&\geq \int_0^1 u(W - I^N - m^N q^N, q^N, h)\mu^C(h)dh \\
&\int_0^1 u(W - I^C - m^C q^C, q^C, h)\mu^C(h)dh \geq \overline{u}^C \\
&\int_0^1 u(W - I^N - m^N q^N, q^N, h)\mu^N(h)dh \geq \overline{u}^N \\
&q^N \equiv q(W - I^N, m^N, h), q^C \equiv q(W - I^C, m^C, h)
\end{aligned}
\tag{27}
$$

Since this problem is additively separable in (I^C, m^C) and (I^N, m^N), the joint profit-maximization problem is identical to two separate problems, in which the monopolist maximizes profits over each contract. Specifically, the maximization problem in 27 is equivalent to the pair of maximization problems below:

$$
\begin{aligned}
&\max_{m^C, I^C} I^C + (m^C - MC)\int_0^1 q^C \mu^C(h)dh \\
&s.t. \\
&\int_0^1 u(W - I^{N^*} - m^{N^*} q^{N^*}, q^{N^*}, h)\mu^N(h)dh \\
&\geq \int_0^1 u(W - I^C - m^C q^C, q^C, h)\mu^N(h)dh \\
&\int_0^1 u(W - I^C - m^C q^C, q^C, h)\mu^C(h)dh \geq \overline{u}^C \\
&q^N \equiv q(W - I^{N^*}, m^{N^*}, h), q^C \equiv q(W - I^C, m^C, h)
\end{aligned}
\tag{28}
$$

$$
\begin{aligned}
&\max_{m^N, I^N} I^N + (m^N - MC)\int_0^1 q^N \mu^N(h)dh \\
&s.t. \\
&\int_0^1 u(W - I^{C*} - m^{C*}q^C, q^C, h)\mu^C(h)dh \\
&\geq \int_0^1 u(W - I^N - m^N q^N, q^N, h)\mu^C(h)dh \\
&\int_0^1 u(W - I^N - m^N q^N, q^N, h)\mu^N(h)dh \geq \bar{u}^N \\
&q^N \equiv q(W - I^N, m^N, h), q^C \equiv q(W - I^{C*}, m^{C*}, h)
\end{aligned}
\tag{29}
$$

As in the moral hazard case, it is straightforward to show that these problems yield Pareto-equivalent allocations to the competitive problems.

Without loss of generality, we show this for the type N contract. To net out distributional effects, we assume that the representative type N consumer holds a claim on all profits that flow from contracts with type N consumers. There may not be a well-defined equilibrium in the case of adverse selection, but for our purposes, it suffices to consider the case where an equilibrium exists. If no equilibrium exists, deadweight loss from monopoly is undefined. Define $\bar{\pi}^N$ as the equilibrium profit associated with the solution to 29. If so, then 29 is identical to a problem in which the firm maximizes consumer utility subject to a reservation profit constraint, and the incentive constraint. This problem will also yield profits equal to $\bar{\pi}$, incentive-compatibility, and utility at least equal to $\bar{u}^N$:

$$
\begin{aligned}
&\max_{m^N, I^N} \int_0^1 u(W - I^N - m^N q^N + \bar{\pi}^N, q^N, h)\mu^N(h)dh \\
&s.t. \\
&\int_0^1 u(W - I^{C*} - m^{C*}q^C, q^C, h)\mu^C(h)dh \\
&\geq \int_0^1 u(W - I^N - m^N q^N, q^N, h)\mu^C(h)dh \\
&I^N + (m^N - MC)\int_0^1 q^N \mu^N(h)dh \geq \bar{\pi}^N \\
&q^N \equiv q(W - I^N, m^N, h), q^C \equiv q(W - I^{C*}, m^{C*}, h)
\end{aligned}
\tag{30}
$$

Substituting the reservation profit constraint into the consumer's objective function yields:

$$\max_{m^N, I^N} \int_0^1 u(W - m^N(q^N - E(q^N)) - MC * E(q^N), q^N, h)\mu^N(h)dh$$

s.t.

$$\int_0^1 u(W - I^{C*} - m^{C*}q^C, q^C, h)\mu^C(h)dh \tag{31}$$

$$\geq \int_0^1 u(W - I^N - m^N q^N, q^N, h)\mu^C(h)dh$$

$$q^N \equiv q(W - I^N, m^N, h), q^C \equiv q(W - I^{C*}, m^{C*}, h)$$

This problem is identical to the displaced version of the competitive problem in 22.[13] Therefore, the monopoly allocation is identical to the competitive one.

[13] Under competition, $p = MC$, and $I^C = -(m^C - p)\int_0^1 q^C \mu^C(h)dh$.

C. Static Efficiency of Health Care Patent Monopolies

C.1 The Efficient Static Allocation

To calculate the efficient allocation we first solve the Pareto problem. We augment the model in Section B.1 by considering an economy researching a continuum of innovations, indexed over $i \in [0,1]$. For simplicity, suppose that any given consumer benefits from at most one particular innovation. For each value of i, there exists a continuum of consumers indexed over their ex post health status, $h \in [0,1]$. Effectively, this economy consists of consumers populating the unit square, $(h,i) \in [0,1]x[0,1]$, defined over health status and demand for a particular innovation.

If innovation i is discovered, consumers of type i enjoy expected utility $\int_0^1 u^i(W - I^0 - mq^*, q^*, h)dh$. Society has access to an insurance technology for each of the discovered innovations: each consumer of type i pays the unconditional insurance premium I^0, and receives the copayment m on the newly discovered technology. This insurance scheme is required to break even.

If innovation i is not discovered, consumers of type i consume zero units of the medical remedy and have no demand for insurance as a result. To compartmentalize the risks from innovation and health shocks, we assume that innovations are discovered before the consumer has to make an insurance purchase decision. There is no way to insure against the failure to discover innovations for those consumers who need them.

The more society invests in research and development, the more innovations are discovered. Define $0 \le \rho(r) \le 1$ as the fraction of innovations discovered; in particular, all innovations for which $i \le \rho(r)$ are discovered. The Pareto problem can thus be written as:

$$\begin{aligned} &\max_{I,m,r} \rho(r)\int_0^1 u(W - I^0 - mq, q, h)dh + (1-\rho(r))\int_0^1 u(W,0,h)dh \\ &s.t.\ \rho(r)\left[I^0 + (m - MC)E(q)\right] \geq r \end{aligned} \tag{32}$$

By displacement, we can write this as:

$$\max_{I,m,r} \rho(r)\int_0^1 u(W - m\left(q - E(q)\right) - \frac{r}{\rho(r)}, q, h)dh + (1-\rho(r))\int_0^1 u(W,0,h)dh \tag{33}$$

It is straightforward to show that innovations, when discovered, are utilized (second-best) efficiently, where $m < MC$.

C.2 The Monopoly Allocation with Two-Part Health Insurance

Now suppose that innovation is performed by an integrated insurer-producer-innovator, who is a monopolist. This firm has access to two-part health insurance pricing. Define π^i as the profits earned by consumer type i . Profits are distributed as follows: consumers for whom $i \geq \rho(r)$ receive net distributions of zero; those for whom $i \leq \rho(r)$ share equally in the firm's positive profits. The integrated innovator solves the problem:

$$\begin{aligned} &\max_{I,m,r} \rho(r)\left(I + \left(m - MC\right)E(q)\right) - r \\ &s.t.\ \int_0^1 u(W - I - mq + \pi^i, q, h)dh \geq \int_0^1 u(W,0,h)dh,\ \forall i \leq \rho(r) \end{aligned} \tag{34}$$

Defining Π as the innovator's total profits in equilibrium and adding a constant term to the objective function, we can rewrite this as:

$$\begin{aligned} &\max_{I,m,r} \rho(r)\int_0^1 u(W - I - mq + \pi^i, q, h)dh + (1-\rho(r))\int_0^1 u(W,0,h)dh \\ &s.t.\ \rho(r)\left(I + \left(m - MC\right)E(q)\right) - r \geq \Pi \end{aligned} \tag{35}$$

In equilibrium, π^i is equal for all $i \leq \rho(r)$, and $\Pi = \rho(r)\pi^i$. Therefore, we can rewrite this as:

$$\max_{I,m,r} \rho(r)\int_0^1 u(W - m\left(q - E(q)\right) - \frac{r}{\rho(r)}, q, h)dh + (1-\rho(r))\int_0^1 u(W,0,h)dh \tag{36}$$

This is identical to the planner's problem in equation 33.

D. Pharmaceutical Data

Table D-1: List of drugs in the analysis

Product Name	Product Name	Product Name
ACTIGALL	GLUCOPHAGE	PROZAC
AMIDATE	GLUCOTROL	QUESTRAN
AMIKIN	HALCION	RELAFEN
ANAFRANIL	HYDREA	RIFADIN
ANSAID	HYTRIN	RYTHMOL
AREDIA	IMURAN	SECTRAL
ATROVENT	INTAL	SELDANE
BETAGAN	KLONOPIN	SINEMET
BETAPACE	LODINE	STADOL
BUMEX	LOPID	TAGAMET
BUSPAR	LOPRESSOR	TAMBOCOR
CAPOTEN	LOTRIMIN	TAXOL
CAPOZIDE	LOZOL	TEMOVATE
CARAFATE	LUVOX	TENEX
CARDENE	MEFOXIN	TENORETIC
CARDIZEM	MEVACOR	TORADOL
CARDURA	MEXITIL	TRACRIUM
CECLOR	MUTAMYCIN	TRENTAL
CEFTIN	NAPROSYN	VASOTEC
CLOZARIL	NEORAL	VEPESID
CORGARD	NEPTAZANE	VERSED
COUMADIN	NIZORAL TA	VISKEN
CYLERT	NOLVADEX	VIVACTIL
DAYPRO	OGEN	VOLTAREN
DIABETA	ORUDIS	WELLBUTRIN
DOBUTREX	PARLODEL	WYTENSIN
DOLOBID	PEPCID	XANAX
ELDEPRYL	PLAQUENIL	ZANTAC
ESTRACE	PRIMACOR	ZARONTIN
EULEXIN	PROCARDIA	ZIAC
FLUMADINE	PROPINE	ZINACEF
	PROSOM	ZOVIRAX

References

Adams, S. (2007). "Health Insurance Market Reform and Employee Compensation: The Case of Pure Community Rating in New York." Journal of Public Economics **91**(5-6): 1119-33.

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

Bhattacharya, J. and M. K. Bundorf (2005). The Incidence of the Healthcare Costs of Obesity, National Bureau of Economic Research, Inc, NBER Working Papers: 11303.

Bhattacharya, J. and W. B. Vogt (2006). "Employment and Adverse Selection in Health Insurance." National Bureau of Economic Research Working paper 12430. Cambridge, MA.

Capps, C., D. Dranove and M. Satterthwaite (2003). "Competition and Market Power in Option Demand Markets." Rand Journal of Economics **34**(4): 737-63.

Caves, R. E., M. D. Whinston and M. A. Hurwitz (1991). "Patent Expiration, Entry, and Competition in the U S Pharmaceutical Industry." Brookings Papers on Economic Activity: 1-48.

Chipty, T. and C. M. Snyder (1999). "The Role of Firm Size in Bilateral Bargaining: A Study of the Cable Television Industry." Review of Economics and Statistics **81**(2): 326-40.

Crocker, K. J. and J. R. Moran (2003). "Contracting with Limited Commitment: Evidence from Employment-Based Health Insurance Contracts." Rand Journal of Economics **34**(4): 694-718.

Dafny, L. (2008). "Are Health Insurance Markets Competitive?" National Bureau of Economic Research Working Paper 14572. Cambridge, MA.

Dor, A., S. M. Koroukian and M. Grossman (2004). "Managed Care Discounting: Evidence from the MarketScan Database." Inquiry **41**(2): 159-69.

Dranove, D., M. Shanley and W. D. White (1993). "Price and Concentration in Hospital Markets: The Switch from Patient-Driven to Payer-Driven Competition." Journal of Law and Economics **36**(1): 179-204.

Ellison, S. F. and C. Snyder (2003). "Countervailing Power in Wholesale Pharmaceuticals." MIT Working Paper. Cambridge, MA.

Employee Benefit Research Institute (2009). "Health Plan Differences: Fully Insured vs. Self-Insured." Employee Benefit Research Institute Fast Facts 114. Washington, DC.

Fang, H. and A. Gavazza (2007). Dynamic Inefficiencies in Employment-Based Health Insurance System Theory and Evidence. National Bureau of Economic Research, Inc, NBER Working Papers.

Federal Trade Commission (2004). Improving Health Care: A Dose of Competition. Washington, DC.

Garber, A., C. Jones and P. Romer (2006). "Insurance and Incentives for Medical Innovation." Forum for Health Economics and Policy **9**(2): Article 4.

Gaynor, M. and D. Haas-Wilson (1999). "Change, consolidation, and competition in health care markets." Journal of Economic Perspectives **13**(1): 141-64.

Gaynor, M. and D. Haas-Wilson (2002). "Change, Consolidation, and Competition in Health Care Markets." The Heinz School, Carnegie-Mellon Working Paper 2002-19. Pittsburgh, PA.

Gaynor, M., D. Haas-Wilson and W. B. Vogt (2000). "Are Invisible Hands Good Hands? Moral Hazard, Competition, and the Second-Best in Health Care Markets." Journal of Political Economy **108**(5): 992-1005.

Gaynor, M. and W. B. Vogt (2000). Antitrust and Competition in Health Care Markets. Handbook of health economics. A. J. Culyer, and J. P. Newhouse, Eds. Amsterdam; New York and Oxford, Elsevier Science, North-Holland. **1B:** 1405-87.

Gaynor, M. and W. B. Vogt (2003). "Competition among Hospitals." Rand Journal of Economics **34**(4): 764-85.

Ho, K. (2006). "The Welfare Effects of Restricted Hospital Choice in the US Medical Care Market." Journal of Applied Econometrics **21**(7): 1039-79.

Keeler, E. B., G. Melnick and J. Zwanziger (1999). "The Changing Effects of Competition on Non-profit and For-Profit Hospital Pricing Behavior." Journal of Health Economics **18**(1): 69-86.
Lakdawalla, D. N. and N. Sood (2009). "Innovation and the Welfare Effects of Public Drug Insurance." Journal of Public Economics **93**: 541-548.
Levy, R. (1999). "The pharmaceutical industry. A discussion of competitive and antitrust issues in an environment of change." Bureau of Economics/Federal Trade Commission staff report.
Nordhaus, W. D. (2004). "Schumpeterian Profits in the American Economy: Theory and Measurement." National Bureau of Economic Research Working Paper 10433. Cambridge, MA.
Oi, W. Y. (1971). "A Disneyland Dilemma: Two-part Tariffs for a Mickey Mouse Monopoly." Quarterly Journal of Economics **85**: 77-96.
Pauly, M. (1968). "The Economics of Moral Hazard: Comment." American Economic Review **58**(3): 531-537.
Philipson, T. J. and A. B. Jena (2006). "Surplus Appropriation from R&D and Health Care Technology Assessment Procedures." National Bureau of Economic Research Working Paper 12016. Cambridge, MA.
Reinganum, J. (1989). The Timing of Innovation: Research, Development, and Diffusion. Handbook of Industrial Organization. R. Schmalensee, and R. D. Willig, Eds. New York, Elsevier Science. **1:** 850-908.
Rothschild, M. and J. E. Stiglitz (1976). "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." Quarterly Journal of Economics **90**(4): 630-49.
Shapiro, C. (2007). Patent Reform: Aligning Reward and Contribution. National Bureau of Economic Research, Inc, NBER Working Papers.
Shavell, S. and T. van Ypersele. (1998). "Rewards versus Intellectual Property Rights." Retrieved December 1998, from SSRN: <http://ssrn.com/abstract=145292> or DOI: 10.2139/ssrn.145292 <http://dx.doi.org/10.2139/ssrn.145292>
Stiglitz, J. E. (1977). "Monopoly, Non-Linear Pricing and Imperfect Information: The Insurance Market." The Review of Economic Studies **44**(3): 407-430.
Town, R., D. Wholey, R. Feldman, et al. (2006). "The Welfare Consequences of Hospital Mergers." University of Minnesota Working Paper. St. Paul.
Zeckhauser, R. (1970). "Medical Insurance: A Case Study of the Tradeoff Between Risk Spreading and Appropriate Incentives." Journal of Economic Theory **2**(1): 10-26.
Zwanziger, J. and G. A. Melnick (1988). "The Effects of Hospital Competition and the Medicare PPS Program on Hospital Cost Behavior in California." Journal of Health Economics **7**(4): 301-20.