HEALTH INSURANCE AS A TWO-PART PRICING CONTRACT

Darius Lakdawalla
Neeraj Sood

ABSTRACT

Monopolies appear throughout health care markets, as a result of patents, limits to the extent of the market, or the presence of unique inputs and skills. In the health care industry, however, the deadweight costs of monopoly may be small or even absent. Health insurance, frequently implemented as an ex ante premium coupled with an ex post co-payment per unit consumed, effectively operates as a two-part pricing contract. This allows monopolists to extract consumer surplus without inefficiently constraining quantity. This view of health insurance contracts has several implications: (1) Low ex post copayments to insured consumers substantially reduce deadweight losses from medical care monopolies -- we calculate, for instance, that the presence of health insurance lowers monopoly loss in the US pharmaceutical market by 82 percent; (2) Price regulation or break-up of health care monopolies may be inferior to laissez-faire or simple redistribution of monopoly profits; and (3) Promoting efficiency in the health insurance market can reduce static losses in the goods market while improving the dynamic efficiency of innovation.

Darius Lakdawalla
The RAND Corporation
1776 Main Street
Santa Monica, CA 90407-2138
and NBER
darius@rand.org

Neeraj Sood
The RAND Corporation
1776 Main Street
Santa Monica, CA 90407
and NBER
neeraj_sood@rand.org

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 do not entitle policyholders to unlimited utilization, but instead charge an ex post unit price or co-payment. Co-payments reduce insurance, but in return produce fewer distortions in the goods market, because the consumer faces a private price that partially reflects social cost.

In addition to the trade-off between risk-sharing and incentives, health insurance contracts have another aspect that is less well-appreciated: consumer surplus-extraction. 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. It is well understood that such two-part pricing contracts allow a monopolist to sell goods at marginal cost, but extract consumer surplus in the form of an additional payment (see the seminal paper by Oi, 1971). In the context of health insurance, marginal cost co-payments allow a firm to extract the maximum possible consumer surplus, because there is no deadweight loss to consumers. Therefore, a firm with market power and access to a two-part health insurance contract has strong incentives to treat this contract as a two-part pricing scheme. In effect, the uncertainty of health care demand, coupled with ex ante or ex post asymmetric information, *creates* a contractual structure that facilitates the efficient extraction of consumer surplus. Moreover, while copayments exactly equal to marginal cost represent an ideal case, the actual practice of health insurance arrives fairly close to such an equilibrium.

This logic is robust to a wide variety of contexts. It applies directly when a monopolist or oligopolist health-care provider is integrated with a health insurer. In this case, the health-care provider directly uses the two-part insurance contract to extract surplus. Two prominent examples of this vertically integrated corporate form are: a staff-model Health-Maintenance Organization (HMO); and a pharmaceutical innovator integrated with a pharmacy benefit manager (PBM) that

manages prescription drug insurance. However, these incentives also operate on the relationship between a monopolist health-care provider and a *separate* health insurance industry. If insurance is competitively provided, a monopolist can use its market power to induce insurers to extract maximum consumer surplus on its behalf. This surplus-extraction is made possible by the two-part pricing tools at the insurers' disposal. Even if insurance is not competitive, both insurers and providers have strong incentives to maximize the consumer surplus available for extraction, which they then split amongst themselves. Finally, the core intuition is robust to common failures in the insurance market, like moral hazard and adverse selection. Regardless of the information structure, firms have incentives to maximize the consumer surplus available for extraction. The inefficiency of the insurance market lowers the total surplus available, not incentives to maximize extractible surplus.

Our results have several important and novel implications. First, monopolies in health care—whether due to patents, limited market size, or historical factors—may have smaller or even no deadweight costs in the goods market. As an important example of how the theoretical arguments influence actual practice, we calculate that the presence of health insurance lowers deadweight loss in the US pharmaceutical market (where patent monopolies are prevalent) by 82%. In the particular context of prescription drugs, patent protection stimulates innovation at much less static deadweight cost than in other markets.[1] Two-part health insurance has dynamic benefits, because it allows more efficient surplus-extraction by innovators.

Second, market power may lead to higher than competitive insurance premia, but it should not affect co-payments and utilization, which are predicted to be driven by marginal costs, and

---

[1] The need for patents and the difficulties of encouraging innovation are well-understood (Nordhaus, 1969; Wright, 1983). The efficiency of paying innovators consumer surplus has implications for cost-effectiveness analysis, which should account for the need to reward innovation (Pauly, 2005; Philipson and Jena, 2006).

possibly the incompleteness of information. When coupled with efficient co-payments, higher premia merely redistribute, without affecting the allocation of goods. Society could undo this redistribution, if it desired, by taxing the profits of the monopolist and redistributing to consumers. This scheme would neither improve nor detract from efficiency. In contrast, directly regulating monopolists cannot improve efficiency and can sometimes harm consumers.

Finally, our analysis provides some guidance for the optimal design of public health insurance benefits, which ought to set co-payments at or below marginal cost, and set insurance premia according to society's particular redistributive goals.

We develop our argument by analyzing four progressively less ideal contexts. As a benchmark, we begin with first-best efficiency in Section B, where all consumers are identical ex ante, and all ex post heterogeneity is fully observable. We consider both an integrated insurer-provider, and the separation between providers and insurers. Section C extends the argument to the context of innovation, where two-part health insurance pricing yields first-best utilization in the goods market *and* (in the presence of competitive insurers)[2] first-best innovation. Sections D, E, and F show how the argument generalizes even in the presence of moral hazard, monopolistic competition among health-care providers, and adverse selection in insurance. Section G presents our estimate of deadweight loss reduction in the US pharmaceutical market. Section H concludes with several implications for health care policy.

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 of the good. This point can be

---

[2] Garber, Jones, and Romer (2006) analyze the impacts of subsidizing the insurance market, which leads to more than first-best innovation.

made most simply in the context of a full information model, where there is neither moral hazard nor adverse selection. Our initial setup is very similar to that of Gaynor, Haas-Wilson, and Vogt (2000), who show that imperfect competition in healthcare markets does not reduce deadweight losses from moral hazard in insurance.

We begin with the most standard setup of full information and indemnity insurance. From the ex ante perspective, consumers face a risk of illness, and an uncertain demand for a medical remedy. The medical remedy is produced at constant marginal cost equal to MC. An insurance contract is an offer of an ex post co-payment (m), coupled with indemnity transfers (τ). In this simplest full information case, the indemnity transfers can be conditional on the consumer's health state. As such, ex post co-payments are not strictly necessary, because the insurer could write contracts characterized entirely by a set of indemnity payments and pre-specified medical care quantities. Co-payments in this case may result in a simpler contractual form, but they are equivalent to contracting on quantity. However, the analysis of this simple case helps set the analytical stage for our discussion of incomplete information.

Suppose there are consumers of measure one, indexed by $h \in [0,1]$, and distributed uniformly over this interval. Consumer health is represented by this index h, which is a random variable unknown ex ante, but revealed to the consumer after the insurance contract is purchased. Ex post consumer utility depends on non-medical consumption, the quantity of medical care consumed (q), and the revealed health state, according to $u(c,q,h)$.

Since information is complete here, it sacrifices no generality to assume that there are just two states: sickness and health. The consumer is sick with probability σ. Utility in each state is given by u^s and u^h. The marginal utility of medical care is positive when sick, but zero when healthy. The marginal utility of consumption is higher in the healthy state.

Ex post, a sick consumer with wealth W and the health insurance contract (m, τ^s, τ^h) solves the following problem:

$$\max_q u^s (W - I - mq + \tau^s, q) \tag{1}$$

This is characterized by the first-order condition:

$$u^s_W m = u^s_q \tag{2}$$

This first-order condition implicitly defines the ex post demand for medical care as a function of ex post disposable income, the co-payment, and health, according to $q^*(W - I + \tau^s, m)$.

B.1   Competitive Outcomes

If a consumer faces a competitive insurance industry and a competitive goods market, the outcome is first-best. This is our benchmark case; we will show that a monopolist with access to a two-part health insurance contract will replicate its outcomes. The medical care sector sells its goods at the constant marginal cost of production, and (since information is complete), there is also a competitive insurance market that offers indemnity coverage at actuarially fair prices.

The equilibrium in this market is straightforward. Consumers buy full indemnity insurance at actuarially fair prices, and, when sick, purchase the innovation at marginal cost from the competitive goods-producing sector. The following three conditions obtain:

1. Full insurance indemnity transfers, in the sense that $u^s_W = u^h_W$;

2. Efficient use of medical care, where, consumers face price equal to the marginal cost of production $p = MC$;

3. Zero profits for insurers and medical-providers.

B.2 The Impact of Monopoly

In the presence of a two-part pricing contract for health insurance, monopoly does not change the equilibrium price or co-payment for medical care. Its only impact is to charge an actuarially unfair premium that is used to extract consumer surplus. If consumers own the monopolist in proportion to their premia, this has no impact on the allocation or distribution of resources, because the extracted surplus does not make them poorer. If not, a simple tax-and-transfer scheme, without any pricing regulation, can redistribute surplus to achieve the competitive allocation.

The simplest way to understand this is to consider an integrated monopolist who provides both health care and insurance. In addition to being pedagogically useful, such vertical relationships are not uncommon in the health care industry. For example, Kaiser Permanente, a staff-model HMO, controlled one-third of the California HMO market in 2004 (Baumgarten, 2005). Similarly, a 1999 FTC study found that drug companies owned or had a significant affiliation with PBMs that account for majority of the PBM activity. Indeed, in 1994 independent PBMs accounted for less than 30% of prescriptions (Levy, 1999).[3]

The monopolist maximizes profits subject to the consumer's participation constraint. The amount of surplus the monopolist can extract depends on the consumer's next available outside option. Without loss of generality, suppose there are no other firms available to supply health care or insurance of any kind.[4] Therefore, the reservation utility level

$\overline{U} = \sigma u^s(W + \pi, 0) + (1 - \sigma)u^h(W + \pi, 0)$ is utility under autarky, where firm profits are π, and we

---

[3] Since 1994 some pharmaceutical companies have divested their stock holdings in PBMs, but still maintain strategic interests in them (Martinez, 2002).

[4] Changing this assumption affects only the level of rents earned by the firm, which we show below to be (largely) neutral in this problem.

assign ownership of the firm to consumers. This results in the profit-maximization problem for the monopolist:

$$\max_{\tau^s, \tau^h} (1-\sigma)\tau^h + (m - MC)\sigma q^* - \sigma\tau^s$$

$$s.t. \ \sigma u^s(W - mq^* + \tau^s + \pi, q) + (1-\sigma)u^h(W - \tau^h + \pi) \geq \overline{U} \tag{3}$$

This problem has the following first-order conditions (simplified by using the consumer's optimality condition for q):

$$[\tau^s] : \mu u_W^s = 1 + (MC - m)q_W$$

$$[\tau^h] : \mu u_W^h = 1 \tag{4}$$

$$[m] : q(1 - \mu u_W^s) = (MC - m)q_m$$

The utilization of medical care is identical to the first-best if the monopolist sets the copayment equal to marginal cost.

Intuitively, price equals marginal cost, because this strategy maximizes the consumer surplus available for extraction. The following arguments formalize this intuition. Suppose that $m > MC$. Define q^*, τ^{s*}, τ^{h*}, and m^* as the contract values in the initial (putative) equilibrium. Consider the alternative insurance contract that sets m equal to MC. It is easy to verify that the new contract increases ex post consumer surplus by more than $q^*(m^* - MC)$. Therefore, there exists some $\varepsilon > q^*(m^* - MC)$ such that the consumer strictly prefers the contract $(\tau^{s*} - \varepsilon, \tau^{h*}, MC)$ to $(\tau^{s*}, \tau^{h*}, m^*)$. Moreover, the new contract is strictly more profitable than the old one, because the reduction in the indemnity transfer (ε) exceeds the value of the revenue lost from the price reduction, $q^*(m^* - MC)$. The existence of the alternative contract contradicts the initial equilibrium.

Now suppose $m < MC$. Define q^*, τ^{s*}, τ^{h*}, and m^* as the contract values in this putative equilibrium. Consider the alternative insurance contract that sets m equal to MC. It is easy to

verify that, when the copayment rises to MC, the consumer's loss in surplus is strictly less than $q^*(MC - m^*)$, as long as the demand for medical care is downward-sloping. Therefore, there exists some $\varepsilon' < q^*(MC - m^*)$ such that the consumer strictly prefers the contract $(\tau^{s*} + \varepsilon', \tau^{h*}, MC)$ to $(\tau^{s*}, \tau^{h*}, m^*)$. The marginal cost contract is strictly more profitable for the firm than the old one, because in the initial equilibrium the firm is losing $q^*(MC - m^*)$ on sales. The existence of the alternative contract thus contradicts the initial equilibrium. Therefore, $m = MC$.

Since the monopolist sets the co-payment equal to marginal cost, the first-order conditions for τ^s and τ^h imply that the consumer will be fully insured in the sense that $u_W^s = u_W^h$.

Finally, profits must be positive, because the participation constraint binds. Suppose, to the contrary, that profits are zero. This implies that the consumer's utility will be equal to that of autarky, which is lower than in the first-best equilibrium. If profits are zero and utility is lower than under competition, the competitive allocation offers higher total surplus. The monopolist should thus be choosing a different allocation. The equilibrium contract under monopoly can now be summarized as:

1. Full insurance indemnity transfers, in the sense that $u_W^s = u_W^h$;

2. Consumers face the price equal to the marginal cost of production $p = MC$;

3. Positive profits for the monopolist-insurer, by means of actuarially unfair premia;

There is one remaining result to show: the monopoly allocation is Pareto-equivalent to the competitive allocation. In particular, when consumers own the firm, monopoly produces the same level of consumer utility as competition. Define π^* as the equilibrium level of monopoly profit. The problem in 3 can be equivalently rewritten as:

$$\max_{\tau^s, \tau^h} \sigma u^s(W - mq^* + \tau^s + \pi^*, q^*) + (1 - \sigma)u^h(W - \tau^h + \pi^*)$$
$$s.t. (1 - \sigma)\tau^h + (m - MC)\sigma q^* - \sigma\tau^s \geq \pi^*$$

8

Now observe that this problem is the same as maximizing the following over $(\tau^h - \pi^*)$ and $(\tau^s + \pi^*)$:

$$\max_{(\tau^s+\pi^*),(\tau^h-\pi^*)} \sigma u^s(W - mq^* + (\tau^s + \pi^*), q^*) + (1-\sigma)u^h(W - (\tau^h - \pi^*))$$
$$s.t.\ (1-\sigma)(\tau^h - \pi^*) + (m - MC)\sigma q^* - \sigma(\tau^s + \pi^*) \geq 0$$

This is just the competitive insurer's problem, of choosing transfers that maximize consumer utility subject to a zero profit constraint. The consumer's maximum utility will thus be identical to that under competition.

B.3 Separating the Insurance- and Goods-Producers

The preceding analysis demonstrated the use of health insurance contracts as a means of surplus-extraction by considering a single firm that provided both insurance and goods. Such a model is directly relevant for vertically integrated firms like staff-model HMO's, or pharmaceutical firms integrated with PBM's, but its results also apply to markets where insurance and health-care provision are separated. Analytically, we consider the case of a monopoly goods-provider interacting with a competitive insurance market. Later, we discuss how the results generalize to the case of bilateral monopoly between an insurer and goods-producer. Both these cases produce efficient outcomes. If consumers receive all the firms' rents in proportion to their utilization of the good, the monopoly distribution of resources is identically equivalent to the competitive distribution. If not, simple tax-and-transfer schemes can produce an equivalent outcome without regulating the goods market.

The representative insurer faces a monopolist selling the good. In negotiating with the insurer, the monopolist is able to specify both a price and a quantity, or equivalently, a quantity and a total fixed fee. This type of contracting is often observed in health care markets, where quantities are either explicitly named (e.g., by a pharmaceutical wholesaler), or tied to a nonlinear price schedule

(e.g., in the form of quantity discounts, rebates, and the like). For example, contracts between PBMs and pharmaceutical firms are of two types – non-capitated and capitated.[5] Non-capitated contracts usually specify a list price or "wholesale acquisition costs" and terms for determining discounts or rebates. Rebates are usually tied to the dollar or unit sales of a particular drug product. For example, 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. Similarly complex pricing arrangements are also common between hospitals and insurers (Melnick, 2004).

The ability to set both a price and a quantity is important. When the monopolist is able to specify only one of these, we revert to the analysis of monopoly articulated by Gaynor, Haas-Wilson, and Vogt (2000), where the usual societal losses are incurred.[6] Specifying both prices and quantities for heterogeneous consumers is quite impractical in the absence of two-part health insurance. The provider would need to specify a different price-quantity pair, or two-part pricing menu for each of these heterogeneous consumers. The two-part structure of health insurance provides a natural and practical way to do so.

The insurer takes as given a fixed quantity and a fixed fee associated with that quantity. Since he is competing for a contract from a monopolist, he must maximize his gross profits — gross

---

[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.

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

of the fee paid to the monopolist — subject to the participation of the consumer. Given the pre-specified quantity q^*, we can write the insurer's new problem as:

$$
\begin{aligned}
G(q^*) &= \max\nolimits_{\tau^h,\tau^s,m} (1-\sigma)\tau^h + m\sigma q(m,W+\tau^s) - \sigma\tau^s \\
&\text{s.t. } \sigma u^s(W - mq + \tau^s, q) + (1-\sigma)u^h(W - \tau^h) \geq \overline{U} \\
&\text{and } \sigma q(m, W + \tau^s) \leq \sigma q^*
\end{aligned}
\tag{5}
$$

Associating the multipliers μ and η with the two constraints, respectively, this problem has the following first-order conditions:

$$
\begin{aligned}
[\tau^s] &: \mu u_W^s = 1 + q_W(\eta - m) \\
[\tau^h] &: \mu u_W^h = 1 \\
[m] &: q(1 - \mu u_W^s) = (\eta - m)q_m
\end{aligned}
\tag{6}
$$

Note that these first-order conditions are identical to the case of the integrated insurer, except that MC is replaced by η. The envelope theorem implies that $\dfrac{dG}{dq^*} = \sigma\eta$, the change in expected profits associated with an increase in the monopolist's offer of quantity. Since the monopolist can extract all gross profits, he will equate this marginal change to the expected marginal cost of output, σMC, making these first-order conditions identical to those of the integrated case.

Formally, the monopolist selling quantity q^* is able to charge a fee equal to $G(q^*)$. Therefore, the profit-maximizing monopolist solves:

$$
\max\nolimits_{q^*} G(q^*) - MC * \sigma q
\tag{7}
$$

The first-order condition for this problem implies that $G'(q^*) = \sigma MC$, or that $\eta = MC$. As a result, this equilibrium is identical to that produced by the integrated insurer.

C. Innovation Incentives

A major reason for monopolies in health care is the use of patents to encourage innovation. While patents improve dynamic efficiency, two well-known sources of dynamic and static

11

inefficiency remain (Shavell and van Ypersele, 1998). First, incentives to invest in research remain inadequate, because monopoly profits are less than the social surplus created by the innovation. Second, patents encourage innovation at the expense of static inefficiency from monopoly loss. Two-part health insurance can solve both these problems in health care markets – it limits static inefficiency by subsidizing medical care, and at the same time delivers social surplus to a monopolist in the form of the extracted premium. Thus, it can produce better dynamic incentives for innovation, even while it decreases the static costs associated with encouraging innovation. The only danger arises not from patent protection, but from failure in the insurance market: if health insurance is inefficiently cheap or over-provided (due, for example, to government subsidies), the result will be excessive amounts of innovation (Garber, Jones, and Romer, 2006).

C.1 The Efficient Allocation

It is well-known that competition does not produce first-best outcomes with innovation. Therefore, to calculate the efficient allocation we must solve the Pareto problem. In addition to the structure developed earlier, suppose that the good in question must be developed through research. Society can spend resources r on the research process, and the probability of discovering the new good is $\rho(r)$. $\overline{U}^N$ is maximum utility without the invention. The first-best efficient allocation solves the following (equal weights) Pareto problem:

$$
\begin{aligned}
&\max_{r,c^s,c^h,q} \rho(r)\big(\sigma u^s(c^s,q) + (1-\sigma)u^h(c^h,0)\big) + (1-\rho(r))\overline{U}^N \\
&s.t.\ \rho(r)\big(\sigma c^s + (1-\sigma)c^h + MC*\sigma q\big) \leq \rho(r)(W-r)
\end{aligned}
\tag{8}
$$

Conditional on the innovation being discovered, the efficient allocation shares all the features of the first-best competitive equilibrium without innovation: full insurance and utilization up to the point

where marginal benefit equals marginal cost.[7] Formally, we can characterize it using the following simplified first-order conditions:

$$u_W^s = u_W^h$$

$$\frac{u_q^s}{u_W^s} = MC \tag{9}$$

$$\rho'(r)\left[\sigma u^s(c^s,q) + (1-\sigma)u^h(c^h) - \overline{U}^N\right] = u_W^h$$

The third condition, unique to the innovation problem, stipulates that the marginal value of investing in innovation is equal to its marginal opportunity cost.

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

Above, we showed that the vertical integration of insurer with goods-producer had little impact on the allocation, provided that monopolists can engage in nonlinear pricing. Therefore, we analyze this problem in the expositionally simpler context of the integrated insurer-producer-innovator. Defining the innovator's ex post profits (in the event of discovery) as π^d, and assuming consumers own the firm, the integrated innovator solves the problem:

$$\max\nolimits_{\tau^s,\tau^h,m,r} \rho(r)\left((1-\sigma)\tau^h + (m-MC)\sigma q^* - \sigma\tau^s\right) - r$$

$$s.t. \ \sigma u^s(W - mq^* + \tau^s + \pi^d, q) + (1-\sigma)u^h(W - \tau^h + \pi^d) \geq \overline{U} \tag{10}$$

$\overline{U}$ is maximum utility for the consumer who chooses not to contract with the innovator. This formulation assumes that in the absence of discovery, the firm is simply a competitive insurer earning zero profit. Conditional on discovery, this firm faces the same problem as the integrated insurer in Section B. It shares all its first-order conditions, but adds an equilibrium condition for innovation, as follows:

---

[7] Since we are considering the case of a single innovation, we rule out the possibility of insuring against the failure to innovate, which would require the possibility of transferring resources across the "innovation" and "no innovation" states.

$$[\tau^s]:\mu u_W^s = 1 + (MC - m)q_W$$
$$[\tau^h]:\mu u_W^h = 1$$
$$[m]:q(1 - \mu u_W^s) = (MC - m)q_m \tag{11}$$
$$[r]:\rho'(r)\big((1 - \sigma)\tau^h + (m - MC)\sigma q^* - \sigma\tau^s\big) = 1$$

By the same arguments made in Section B, we can show that $m = MC$. This will then imply full insurance, according to the first-order conditions for τ^s and τ^h. This implies that, conditional on discovery, the provision of insurance and the invented good are Pareto-optimal. It remains to show that investment in research is also efficient. We will do so by showing that the private return to innovation equals the social return.

The private return to innovation is the ex post return earned by the innovator, or $\pi^d + r$. On the other hand, the social return to invention is the total (monetized) gain enjoyed by consumers as a result of the innovation's discovery:

$$\frac{\sigma\big(u^s(W - MC * q + \tau^s + \pi^d, q) - u^s(W - r,0)\big)}{u_W^s} +$$
$$\frac{(1 - \sigma)\big(u^h(W - \tau^h + \pi^d,0) - u^h(W - r,0)\big)}{u_W^h}$$

Since the consumer's reservation utility constraint holds at equality, we know that:

$$\frac{\sigma\big(u^s(W - MC * q + \tau^s + \pi^d, q) - u^s(W + \pi^d,0)\big)}{u_W^s} +$$
$$\frac{(1 - \sigma)\big(u^h(W - \tau^h + \pi^d,0) - u^h(W + \pi^d,0)\big)}{u_W^h} = 0 \tag{12}$$

Taking first-order approximations to $u^s(W + \pi^d,0)$ and $u^h(W + \pi^d)$, we obtain:

$$\frac{\sigma\big(u^s(W - MC * q + \tau^s + \pi^d, q) - u^s(W - r,0)\big)}{u_W^s} +$$
$$\frac{(1 - \sigma)\big(u^h(W - \tau^h + \pi^d,0) - u^h(W + \pi^d,0)\big)}{u_W^s} = \pi^d + r$$

This demonstrates equality between the private and social returns to innovation.

14

C.3 Subsidies for Employer Provided Insurance

The analysis above considered an unregulated, unsubsidized, and competitive insurance market. In practice, however, employer-based health insurance premia are implicitly subsidized, because they are tax-exempt. This affects the optimal level of the insurance premium generally, along with the incentive to innovate, but it does not affect the optimal copayment, or static efficiency in the goods market.

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 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. Notice that we continue to have the result that two-part pricing erases static losses from monopoly, even in the context of innovation.

D. Moral Hazard

The presence of moral hazard is largely responsible for inducing the two-part structure of health insurance. This makes it important to show that moral hazard does not change the implications we have developed. Absent any other market failures, moral hazard leads to a second-best equilibrium. Monopoly with two-part health insurance also achieves this competitive outcome.

Studying the moral hazard problem requires incorporating some additional consumer heterogeneity. We continue to assume that consumers are indexed by $h \in [0,1]$, and distributed uniformly over this interval. We also keep the assumption that the fraction σ fall sick, or all

consumers for whom $h \leq \sigma$. Sick consumers place value on the medical care good, while healthy consumers do not. Therefore, insurers can easily distinguish healthy from sick patients. However, information on the severity of illness is incomplete. Patients with lower values of h are sicker, but the insurer cannot observe this. Therefore, even though they may benefit from more insurance than the less ill patients, there is no way for the insurer to make payments contingent on actual underlying health state. Payments can only be contingent on the consumer's observed decision to purchase the medical good or not. It is impossible to insure all consumers fully. The result is a second-best solution, where the insurer charges co-payments below marginal cost. This results in "over-utilization" relative to the first-best, but this is a welfare-enhancing means of delivering some additional insurance in the face of informational incompleteness.

D.1   The Typical Competitive Problem

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

In this case of heterogeneous consumers, we adopt the simplifying convention that consumers either purchase one unit of the medical good, or none at all. That is, consumers vary along the extensive margin only. The firm's optimization problem can be written as:

$$\max\nolimits_{I, m \leq MC} \int_0^1 u(W - I - mq^*, q^*, h)dh$$
$$s.t. \ I + (m - MC)E(q^*) \geq 0 \tag{13}$$
$$and \ q^* = q(W - I, m, h)$$

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

$$[I]: Eu_W = \mu\left(1 - (m - MC)E(q_W)\right)$$

$$[m]: \int_0^1 \frac{q}{E(q)} u_W(W - I - mq, q, h)dh = \mu\left(1 - (m - MC)\frac{-E(q_m)}{E(q)}\right) \qquad (14)$$

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.[8] This fact, coupled with the two first-order conditions, implies that

$$(m - MC)\left(\frac{E(q_m)}{E(q)} - E(q_W)\right) > 0 \qquad (15)$$

and that $m < MC$. 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, the only way to provide insurance in this limited information case 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. Competitive markets deliver the second-best efficient allocation that maximizes consumer well-being, subject to the economy's information constraints.

---

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

D.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 that is second-best efficient, just like the competitive allocation, but he cannot solve the underlying informational problem.

Consider an insurer who is also a monopoly provider of the good with uncertain demand. We consider an existing good, rather than an innovative one, although the extension to innovation is straightforward. The insurer 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 insurer thus solves:

$$
\begin{aligned}
&\max_{I,m} I + (m - MC)E(q(W - I, m, h)) \\
&s.t. \int_0^1 u(W - I - mq + \pi, q, h)dh \geq \overline{U}
\end{aligned}
\tag{16}
$$

It is straightforward to prove the mechanical equivalence between this problem and the competitive one, so long as consumers own the firm. Formally, if we define $\overline{\pi}$ as the equilibrium monopoly profit level, the above problem is equivalent to:

$$
\begin{aligned}
&\max_{I,m} \int u(W - I - mq + \overline{\pi}, q, h)dh \\
&s.t. I + (m - MC)E(q) \geq \overline{\pi}
\end{aligned}
\tag{17}
$$

Substituting in the reservation profit constraint allows us to rewrite this as:

$$
\max_m \int_0^1 u(W - m(q - E(q)) - MC * E(q), q, h)dh,
\tag{18}
$$

which is exactly equivalent to the displaced version of the problem in 13.

A more informal but also more illuminating proof demonstrates exactly why the monopolist chooses to solve the second-best Pareto problem. The reservation utility condition can be first-order approximated by:

$$I + mq \approx \frac{1}{u_W} \int_0^1 u_q(W - I, 0, h)q(W - I, m, h)dh \equiv CS(W - I, m), \qquad (19)$$

where CS is monetized consumer surplus. In words, the monopolist can extract in total revenues no more than gross consumer surplus from use of the good. Therefore, the monopolist's problem under two-part health insurance pricing is equivalent to:

$$\max_{I,m} CS(W - I, m) - MC * E(q(W - I, m, h)) \qquad (20)$$

This is the second-best Pareto problem, which maximizes social surplus given the economy's contracting constraints. The monopolist can maximize profits by first maximizing gross consumer surplus, and then extracting this.

E. Monopolistic Competition

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, one strictly preferring A and 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.

E.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:

$$\max\nolimits_{I^j,m^j \leq MC} \int_0^1 u^j(W - I^j - m^j q^{j*}, q^{j*}, h)dh$$
$$s.t.\ I^j + (m^j - MC)E(q^{j*}) \geq 0 \qquad\qquad (21)$$
$$and\ q^{j*} = q^j(W - I^j, m^j, h)$$

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

$$[I]: Eu_W^j = \mu\left(1 - (m^j - MC)E(q_W^j)\right)$$

characterized by:
$$[m]: \int_0^1 \frac{q^j}{E(q^j)} u_W^j (W - I^j - m^j q^j, q^j, h) dh = \mu\left(1 - (m^j - MC)\frac{-E(q_m^j)}{E(q^j)}\right)$$

(22)

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

E.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:

$$\max_{I^A, m^A} I^A + (m^A - MC)E(q^A)$$

$$s.t. \int_0^1 u^A(W - m^A q^{A*} - I^A + \pi^A, q^A, h)dh \geq$$

$$\int_0^1 u^A(W - m^B q^{B0} - I^B + \pi^A, q^B, h)dh \qquad (23)$$

$$q^{A*} = q^A(W - I^A, m^A, h)$$

$$q^{B0} = q^{BA}(W - I^B, m^B, h)$$

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 have several times earlier, define $\bar{\pi}^A$ as the optimal level of profit that solves

the firm's problem. The problem in 23 can be equivalently written as:

$$\max_{I^A, m^A} \int_0^1 u^A(W - m^A q^{A*} - I^A + \bar{\pi}^A, q^A, h)dh$$

$$s.t. I^A + (m^A - MC)E(q^A) \geq \bar{\pi}^A \qquad (24)$$

$$q^{A*} = q^A(W - I^A, m^A, h)$$

The displaced version of this problem is identical to the displaced version of the competitive problem

in 21. This demonstrates that monopolistic competition produces the same allocation as pure

competition.

F. Adverse Selection

The basic logic of health insurance as two-part pricing also holds up under another common

failing of insurance markets — adverse selection. 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, in the face of asymmetric

information.

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 distributions of chronically ill and not chronically ill people. The health 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).

F.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:

$$\max_{m^c, I^c} \int_0^1 u(W - I^c - m^c q^c, q^c, h)\mu^c(h)dh$$
$$s.t.$$
$$[\gamma]: \int_0^1 u(W - I^n - m^n q^n, q^n, h)\mu^n(h)dh$$
$$\geq \int_0^1 u(W - I^c - m^c q^c, q^c, h)\mu^n(h)dh$$
$$[\beta]: I^c + \int_0^1 q^c \mu^c(h)dh(m^c - p) \geq 0$$
$$q^n \equiv q(W - I^n, m^n, h), q^c \equiv q(W - I^c, m^c, h)$$

(25)

This problem has the following first-order conditions:

$$[I]: E_c(u_W) - \gamma E_n(u_W) = \beta(1 - E_c(q_W)(m^c - p))$$
$$[m]: E_c(u_W \frac{q^c}{E_c(q^c)}) - \gamma E_n(u_W \frac{q^c}{E_c(q^c)}) = \beta(1 - \frac{E_c(-q_m)}{E_c(q_c)}(m^c - p))$$

(26)

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 this case, however, the second-best moral hazard contracts may sometimes 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.

$$v(I,m) \equiv \max_q \int_0^1 u(W - I - mq(W - I, m, h), q(W - I, m, h), h)\mu(h)dh \qquad (27)$$

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:

$$-\frac{dI}{dm}\Big|_{v^c} > -\frac{dI}{dm}\Big|_{v^n} \qquad (28)$$

This is the typical "single-crossing" property from Rothschild and Stiglitz's (1976) analysis of adverse selection.[9] Second, a given change in the co-payment rate has a bigger impact on a firm's profits, so that:

$$-\frac{dI}{dm}\Big|_{\pi=0} = \frac{E(q) + (MC - m)E(q_m)}{E(q_w)(MC - m)}$$

$$-\frac{dI}{dm}\Big|_{\pi^c=0} > -\frac{dI}{dm}\Big|_{\pi^n=0}$$

(29)

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 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.

---

[9] $-\dfrac{dI}{dm} = \dfrac{E(u_w q)}{E(u_w)}$. 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.
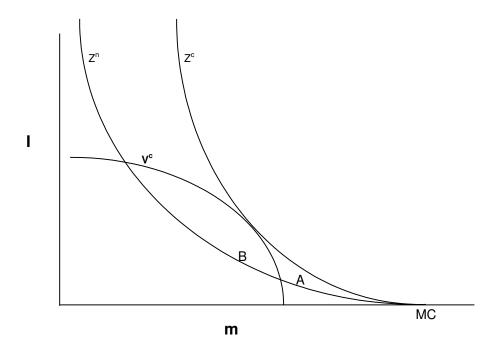
Figure 1: Equilibrium with adverse selection and moral hazard.

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.

F.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.

$$\max_{m^c, I^c, m^n, I^n} I^c + (m^c - MC)\int_0^1 q^c \mu^c(h)dh + I^n + (m^n - MC)\int_0^1 q^n \mu^n(h)dh$$

s.t.

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

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

$$\int_0^1 u(W - I^c - m^c q^c, q^c, h)\mu^c(h)dh \qquad (30)$$

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

$$\int_0^1 u(W - I^c - m^c q^c, q^c, h)\mu^c(h)dh \geq \bar{u}^c$$

$$\int_0^1 u(W - I^n - m^n q^n, q^n, h)\mu^n(h)dh \geq \bar{u}^n$$

$$q^n \equiv q(W - I^n, m^n, h), q^c \equiv q(W - I^c, m^c, h)$$

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 30 is equivalent to the pair of maximization

problems below:

$$\max_{m^c, I^c} I^c + (m^c - MC)\int_0^1 q^c \mu^c(h)dh$$

s.t.

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

$$(31)$$

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

$$\int_0^1 u(W - I^c - m^c q^c, q^c, h)\mu^c(h)dh \geq \bar{u}^c$$

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

$$\max_{m^n, I^n} I^n + (m^n - MC) \int_0^1 q^n \mu^n(h) dh$$

$$s.t.$$

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

$$\geq \int_0^1 u(W - I^n - m^n q^n, q^n, h) \mu^c(h) dh \tag{32}$$

$$\int_0^1 u(W - I^n - m^n q^n, q^n, h) \mu^n(h) dh \geq \overline{u}^n$$

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

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 $\overline{\pi}^n$ as the equilibrium profit associated with the solution to 32. If so, then 32 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 $\overline{\pi}$, incentive-compatibility, and utility at least equal to $\overline{u}^n$:

$$\max_{m^n, I^n} \int_0^1 u(W - I^n - m^n q^n + \overline{\pi}^n, q^n, h) \mu^n(h) dh$$

$$s.t.$$

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

$$\geq \int_0^1 u(W - I^n - m^n q^n, q^n, h) \mu^c(h) dh \tag{33}$$

$$I^n + (m^n - MC) \int_0^1 q^n \mu^n(h) dh \geq \overline{\pi}^n$$

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

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

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

$s.t.$

$$\int_0^1 u(W - I^{c*} - m^{c*}q^c, q^c, h)\mu^c(h)dh \tag{34}$$

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

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

This problem is identical to the displaced version of the competitive problem in 25.[10] Therefore, the monopoly allocation is identical to the competitive one.

G. Deadweight Loss Reduction due to Prescription Drug Insurance

In this section, we calculate—in a "back-of-the-envelope" fashion—how much health insurance lowers deadweight loss in the US market for pharmaceuticals, where patents create a considerable amount of market power. Specifically, we estimate the reduction in deadweight loss that would obtain if we provided the average uninsured consumer with the average prescription drug insurance policy. This calculation illustrates the empirical significance of our key idea – that health insurance can significantly reduce the deadweight loss from monopoly pricing by lowering marginal prices for consumers. We abstract from moral hazard and adverse selection, in order to focus on efficiency losses from monopoly alone.

To calculate the reduction in deadweight monopoly loss, we need to estimate prices and quantity under three scenarios: no insurance, insurance, and competitive marginal cost pricing. We begin with prescription drug use and expenditure data from the 1998 Medical Expenditure Panel Survey (MEPS), which allows us to estimate the quantity of drugs consumed and prices faced by uninsured consumers. Next, we use data on the average rate of cost-sharing for the insured

---

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

population, to calculate the unit price that consumers actually face. Estimates of the price elasticity

of demand for pharmaceuticals then imply the quantity that would be demanded if the uninsured

population acquired the "average" prescription drug insurance policy. Finally, we use estimates of

the mark-up on prescription drugs to compute the competitive marginal cost price of drugs.

Estimates of the elasticity of demand with respect to co-payments are then used to infer the quantity

that would prevail under marginal cost pricing.

Data from the 1998 MEPS show that the uninsured (defined as those without prescription

drug insurance) consumed an average of 8 prescriptions per year and faced an average price per

prescription of $29.[11] In the same year, the average insured consumer paid 40% of the costs of

prescription drugs out-of-pocket. In other words, if the uninsured acquired the "average" prescription

drug insurance policy they would have paid $11.50 ($29*0.40) per prescription. Typically, long-run

generic prices (assumed to be equal to marginal cost) are approximately 10% of the prices charged

for the corresponding on-patent drug (Lakdawalla, Philipson, and Wang, 2006). The standard theory

of monopoly would then imply, based on a 90% mark-up by monopolists, a price elasticity of

uninsured demand around 1.1, or the inverse of the markup. Based on this price elasticity, and an

insured price of $11.50, the uninsured would demand 13.3 prescriptions per year if they received the

average insurance policy. The mark-up of 90% also implies that the marginal cost per prescription is

$2.90. The elasticity of demand at $11.50 may differ from the optimal monopoly elasticity.

However, Goldman et al (2004) have empirically estimated this elasticity to be 0.6. Reducing price

from $11.50 to $2.90 would cause consumers to demand 19.3 prescriptions per year, at the point of

marginal cost pricing.

---

[11] The MEPS is a representative sample of the US civilian noninstitutionalized population; all our
numbers reflect the behavior of this population.

Using a linear approximation to the demand curve, deadweight loss associated with a

particular change in price and quantity is simply the area of the "triangle," or $\frac{1}{2}(-\Delta p)(\Delta q)$—one

half times the reduction in price, times the increase in quantity. Therefore, the per capita deadweight

loss without insurance is approximately \$148, obtained as $0.5 * \overbrace{[19.3 - 8]}^{dQ} * \overbrace{[29 - 2.9]}^{dP}$. Similarly, the

per capita deadweight loss with insurance would be approximately \$26, or

$0.5 * \overbrace{[19.3 - 13.3]}^{dQ} * \overbrace{[11.5 - 2.9]}^{dP}$. Thus, this simple calculation shows that prescription drug insurance

can reduce the deadweight loss from monopoly pricing by more than 82%.

H. Conclusions and Implications for Policy

Two-part pricing is well-known as a solution to the deadweight loss from monopoly, but it is

frequently impractical. In health care markets, the observed structure of insurance contracts provides

a means for achieving the efficient outcomes associated with two-part pricing. While it is not a

panacea for informational problems in the insurance market, it can be an ideal solution to static

deadweight losses from monopoly, as we have shown. By partially decoupling monopoly profits

from consumer prices, two-part health insurance can play an important role in the efficient delivery

of health care, even in the presence of market power.

A review of trends in 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 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 might not significantly reduce social welfare. The optimal design of

insurance contracts can limit or eliminate deadweight losses from monopoly in the goods market. To be sure, monopoly can change the distribution of resources, if patients are not proportionate shareholders. However, society can achieve any distribution it likes — along the Pareto-frontier — simply by taxing profits and transferring them to the appropriate consumers. Breaking up the monopoly is not necessary, and neither is direct price regulation. Indeed, the equivalence between the monopoly and competitive outcomes means that resources spent breaking up a monopoly or regulating prices leave society strictly worse off.

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. In the presence of health insurance, deadweight loss from monopoly arises only if: health care providers are separated from insurers; and providers use simple linear pricing contracts with insurers. If these conditions obtain, breaking up a monopoly or oligopoly is socially desirable (as in Gaynor, Haas-Wilson, and Vogt, 2000). However, the same outcomes can be achieved by encouraging or requiring vertical integration between the monopolist and the health insurance market. In effect, giving *more* vertical market power to a health care monopolist can actually reduce deadweight loss in this case. 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 also improves dynamic incentives, because it allows patent monopolists to extract the maximum amount of consumer surplus associated with their inventions. The result is improved static and dynamic efficiency. In this context, longer patents may have rather limited social costs in terms of deadweight loss from monopoly but considerable social benefits. Taken together, these arguments suggest that competition may do little to improve static efficiency, and that

competition—even monopolistic competition—may do harm to dynamic efficiency. An important caveat here, however, is that government subsidies to the insurance market can lead to over-innovation, although optimal copayments prevent these policies from harming static efficiency.

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.

References

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

Baumgarten, A. (2005). "California Health Care Market Report 2005." California HealthCare Foundation. California HealthCare Foundation.

Garber, A. M., C. I. Jones and P. M. Romer (2006). "Insurance and Incentives for Medical Innovation." National Bureau of Economic Research Working Paper 12080. Cambridge, MA: National Bureau of Economic Research.

Gaynor, M. and D. Haas-Wilson (1999). "Change, consolidation, and competition in health care markets." <u>Journal of Economic Perspectives</u> 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: The Heinz School, Carnegie-Mellon.

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." <u>Journal of Political Economy</u> 108(5): 992-1005.

Goldman, D. P., G. F. Joyce, J. J. Escarce, J. E. Pace, M. D. Solomon, M. Laouri, P. B. Landsman and S. M. Teutsch (2004). "Pharmacy benefits and the use of drugs by the chronically ill." <u>Jama</u> 291(19): 2344-50.

Lakdawalla, D., T. J. Philipson and R. Wang (2006). "Intellectual Property and Marketing." National Bureau of Economic Research Working Paper 12577. Cambridge, MA: National Bureau of Economic Research.

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. Bureau of Economics/Federal Trade Commission.

Martinez, B. (2002). Medco finds that old ties bind in contract with parent Merck after spinnoff, it agrees to maintain market share for Merck medicines. <u>The Wall Street Journal</u>. New York.

Subcommittee on Health of the House Committee on Ways and Means (2004). <u>Hospital Pricing and the Uninsured</u>. House Committee on Ways and Means. March 9, 2004.

Nordhaus, W. (1969). "Invention, Growth and Welfare." <u>Cambridge, MA: MIT Press</u>.

Oi, W. Y. (1971). "A Disneyland Dilemma: Two-part Tariffs for a Mickey Mouse Monopoly." <u>Quarterly Journal of Economics</u> 85: 77-96.

Pauly, M. (1968). "The Economics of Moral Hazard: Comment." <u>American Economic Review</u> 58(3): 531-537.

Pauly, M. V. (2005). "What Measures of Costs and Benefits for Drugs in Cost Effectiveness Analysis." Wharton School Working Paper. Philadelphia, PA: Wharton School.

Philipson, T. J. and A. Jena (2006). "Surplus Appropriation from R&D and Health Care Technology Assessment Procedures." National Bureau of Economic Research Working Paper 12016. Cambridge, MA: National Bureau of Economic Research.

Rothschild, M. and J. E. Stiglitz (1976). "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information." <u>Quarterly Journal of Economics</u> 90(4): 630-49.

Shavell, S. and T. van Ypersele. (1998). "Rewards versus Intellectual Property Rights." Retrieved December 1998, from SSRN: <<u>http://ssrn.com/abstract=145292</u>> or DOI: 10.2139/ssrn.145292 <<u>http://dx.doi.org/10.2139/ssrn.145292</u>>

Wright, B. D. (1983). "The Economics of Invention Incentives: Patents Prizes and Research Contracts." <u>American Economic Review</u> 73: 691-707.

Zeckhauser, R. (1970). "Medical Insurance: A Case Study of the Tradeoff Between Risk Spreading and Appropriate Incentives." <u>Journal of Economic Theory</u> 2(1): 10-26.