Insurance and Innovation in Health Care Markets

Darius Lakdawalla and Neeraj Sood

RAND and NBER

July 5, 2005

Abstract

Innovation policy often involves an uncomfortable trade-off between rewarding innovators sufficiently and providing the innovation at the lowest possible price. However, in health care markets characterized by uncertainty and insurance, society may be able to ensure efficient rewards for inventors *and* the efficient dissemination of inventions. Health insurance resembles a two-part pricing contract in which a group of consumers pay an up-front fee ex ante in exchange for a fixed unit price ex post. This can allow innovators to extract sufficient profits — from the ex ante payment — but still sell the good at marginal cost ex post. As a result, we show that complete, efficient, and competitive health insurance markets lead to perfectly efficient innovation and utilization, even when moral hazard exists. Conversely, incomplete insurance markets lead to inefficiently low levels of innovation. Second, an optimally designed public health insurance system can solve the innovation problem by charging ex ante premia equal to consumer surplus, and ex post co-payments at or below marginal cost. When these quantities are unknown, society can almost always improve static and dynamic welfare by covering the uninsured with contracts that mimic observed private insurance contracts.

A. Introduction

The difficulty of encouraging innovation is well-appreciated (Nordhaus, 1969; Wright, 1983). Innovators need to reap profits in the event of a successful innovation, but profits for a producer often are often at odds with social efficiency. Often, society must make a difficult choice between ex ante efficiency or ex post efficiency. Governments and societies are grappling with the competing needs of rewarding innovators and providing health care to sick patients.

This is often viewed as a zero-sum game between today's patients, who want cheap health care, and tomorrow's patients, who want rapid innovation. However, in the particular context of health care, society may be able to achieve efficiency for both today's and tomorrow' patients. The unique and important role of insurance in these markets explains why. 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. Such two-part pricing contracts can guarantee both the efficient utilization of a product for today's patients, *and* a sufficient level of profit to induce innovation on behalf of tomorrow's patients (Oi, 1971). While this is well-understood, two-part pricing is rarely feasible on a large-scale. However, the existence of health insurance as a two-part pricing strategy changes the terms of the debate between ex post and ex ante efficiency.

The efficient solution to the innovation problem requires both that the innovation be sold at marginal cost ex post, and that the innovator receive ex post profits equal to the net consumer surplus associated with the innovation. When it is feasible, two-part pricing can accomplish both those goals simultaneously. An innovator can charge an upfront fee equal to net consumer surplus, a fee that then allows consumers to buy as much of the innovation as they like at marginal cost. The analogy to health insurance is fairly direct. A health insurance plan with a co-payment equal to marginal cost, and an ex ante premium equal to net consumer surplus would achieve efficiency.

The mechanics of the insurance contract are similar to a two-part pricing contract, but more importantly, the uncertain demand for a health-care innovation plays a fundamental role in ensuring the practicality of this approach. It is often difficult to find and contract with groups of potential consumers ex ante, but group health insurance provides a natural and practical way to do so. Moreover, when consumers differ, it is necessary but very difficult to extract ex ante payments that accurately reflect the varying levels of surplus each consumer derives. However, uncertain demand facilitates this process, because a great deal of heterogeneity emerges ex post, after the contract is written. Consumers may thus be induced to pay their expected surplus ex ante, at which point there is more similarity among them.

Relying on the idea of health insurance as a two-part pricing contract, we show that complete and competitive health insurance markets ensure efficiency in both utilization and innovation, because they deliver the efficient two-part pricing strategy. Therefore, completing insurance markets always improve the efficiency of discovery and utilization. Even when moral hazard exists, competitive insurance markets yield the second-best allocation of resources that represents the best outcome achievable by society.

This suggests that society can improve ex ante and ex post efficiency simply by promoting efficiency in the insurance market. Completing such markets always improves the efficiency of innovation; even in the presence of moral hazard, complete insurance markets never lead to over-provision of innovation, as may be argued. Indeed, patent monopolies granted to innovators only introduce inefficiency to the extent that health insurance markets are incomplete or non-competitive.

In some cases, the government may be unable to ensure efficiency in the private market for insurance. If so, there is a unique justification for public health insurance, as a means of ensuring ex ante and ex post efficiency in the market for health care. Our model also provides guidance for the optimal design of a public health insurance scheme: co-payments ought to be set to marginal cost, while premia ought to equal consumer surplus. When regulators cannot observe one or both of these quantities, a practical and often welfare-improving strategy is to mimic observed private health insurance contracts for the same goods and services.

We develop our argument by analyzing three progressively less ideal contexts, and showing how health insurance markets can lead to first-best or second-best efficiency in all these different settings. As a benchmark, we begin with first-best efficiency, where all consumers are identical ex ante, and all ex post heterogeneity is fully observable to the innovator and to insurance companies. In this classical setting, the first-best is achievable with price-discrimination. We then move to the case of moral hazard: while innovators and insurers know the distribution of demands ex ante, they cannot observe ex post which consumers are the heaviest demanders. Incomplete information bars us from the first-best outcome, but competitive health insurance contracts markets still match the second-best efficient outcome. Finally, we consider the case in which a new innovator must compete with an incumbent. This can lead to rent-seeking behavior, where a new entrant invests excessively in innovation simply to secure some of the incumbent's profits. Here, we show that unregulated and competitive health insurance markets remain optimal, and that any necessary policy intervention ought to take the form of a lump-sum tax on the ex post profits of the new innovator.

B. First-Best Insurance and Innovation

It is well-known that ex post and ex ante efficiency are often at odds in the case of innovation. On the one hand, the inventor ought to receive the full social surplus associated with his invention. Internalizing the full value leads to efficient investments in innovation ex ante. However, efficient utilization of the product requires that it be sold at marginal cost. This leaves little room for profit.

There are a few important cases where ex post and ex ante efficiency can be reconciled. The traditional case is that of perfect price-discrimination. When heterogeneity is fully observable by all parties in the economy, the first-best allocation is achievable simply by granting a patent monopoly and ensuring the existence of a competitive insurance market. The monopolist engages in perfect price discrimination ex post, and consumers insure themselves fully so that consumption is equal across all types. While this case is rarely observed, it serves as a benchmark of perfect efficiency. As a result, we begin by analyzing the joint determination of innovation and insurance when information about demand is perfect.

B.1 The Pareto Optimum

Suppose society is deciding how much to spend researching a new innovation. Spending r resources yields the probability of discovery $\rho(r)$. Suppose consumers vary in their health and in their demand for the new innovation. Formally, suppose they are indexed by i, which is distributed uniformly over the continuum [0,1]. Ex ante, individuals do not know what value of i will be realized for them. Expected utility is the uniform average of utility across $i \in [0,1]$.

The fraction of consumers σ falls ill: if $i \leq \sigma$, the consumer is sick. Sick consumers experience a loss of utility L. The health-care innovation can partially restore health, but its effectiveness varies. Define its utility value to consumer i as $v(i)$, where v falls with i. If W is the wealth of the consumer then utility in each state i,is defined as follows:

$$\begin{aligned} &u(W), \textit{if healthy} \\ &u(W-L)+v(i), \textit{ if sick and uses innovation} \\ &u(W-L), \textit{if sick but does not use innovation} \end{aligned} \tag{1}$$

Assume that: (1) Every consumer is endowed with wealth W, (2) The innovation can be produced ex post at marginal cost MC, and (3) The social marginal utility of resources is μ. The efficient allocation of resources maximizes expected social surplus according to:[1]

$$\begin{aligned} &\max_{q,c(i),r}\ \rho(r)\left[\int_0^1 (u(c(i)))di + \int_0^q (v(i))di\right] + (1-\rho(r))u(W-r) \\ &s.t. \int_0^1 c(i)di + \sigma L + MC * q \leq W - r \end{aligned} \tag{2}$$

The solution to this problem is the familiar one of: (1) Full insurance, (2) Utilization of the innovation until marginal cost equals marginal willingness to pay, and (3) The marginal cost of research equals the expected increment to social surplus associated with the innovation. The first order conditions are:

$$\begin{aligned} &\rho(r)u'(c(i)) = \lambda, \forall i \\ &\rho(r)v(q) = \lambda * MC \\ &\rho'(r)\left[\left(\int_0^1 (u(c(i)))di + \int_0^q (v(i))di\right) - u(W-r)\right] = \lambda + (1-\rho(r))u'(W-r) \end{aligned} \tag{3}$$

We can simplify these expressions as follows:

---

[1] In the absence of discovery, there is no gain to redistributing resources across states of the world, because the marginal utility of consumption is already equated. This owes itself to the additive separability of health and consumption, but it is an inessential and merely simplifying assumption.

$$\frac{v(q^*)}{u'(W-r-MC*q^*-\sigma L)}=MC$$
$$c(i)=c=W-r-MC*q^*-\sigma L,\forall i \tag{4}$$
$$\frac{\rho(r)u'(c)+(1-\rho(r))u'(W-r)}{\rho'(r^*)}=u(W-r-MC*q^*)+\int_0^{q^*}v(i)di-u(W-r)$$

These three equations embody the three conditions above — efficient utilization, full insurance, and a research decision that internalizes the full increment to social surplus.

B.2 The Competitive Equilibrium

Consider the case where a patent monopoly is granted to the innovator in the second period. Assume further that there exists a perfectly competitive insurance market. The monopolist sets quantity and prices, while consumers decide how much insurance to purchase. To close the economy, suppose also that consumers own equal shares in the innovating firm, which earns ex post profits π.

Without an innovation, the consumer's problem is trivial. In the event of an innovation, she solves:

$$\begin{aligned}&\max_{c(i)}\int_0^1 u(c(i))di+\int_0^{q^*}v(i)di\\ &s.t.\int_0^1 c(i)di+\int_0^{q^*}p(i)di+\sigma L\le W-r+\pi\end{aligned} \tag{5}$$

This has the full insurance solution:

$$c(i)=W-r+\pi-\int_0^{q^*}p(x)dx-\sigma L,\forall i \tag{6}$$

The net transfer to each state i can be written as:

$$\begin{aligned}&\tau(i)=p^*(i)+(1-\sigma)L-\int_0^{q^*}p^*(x)dx,\ \text{if sick and user}\\ &\tau(i)=(1-\sigma)L-\int_0^{q^*}p^*(x)dx,\text{if sick and non user}\\ &\tau(i)=-\int_0^{q^*}p^*(x)dx-\sigma L,\text{if healthy}\end{aligned} \tag{7}$$

Taking the consumer's optimal insurance decisions as given, the risk-neutral monopolist maximizes expected profits, but subject to the constraint that consumers using the innovation cannot do better by deviating away from use:

$$\begin{aligned}&\max_{r,p(i),q}\ \rho(r)\left[\int_0^q p(i)di - MC*q\right]-r \\ &s.t. u(W-p(i)+\tau^*(i)-L)+v(i)\geq u(W+\tau^*(i)-L), \forall i\leq q\end{aligned} \tag{8}$$

This has the first order conditions:

$$\begin{aligned}&p(q^*)=MC \\ &\rho'(r)\left[\int_0^q p(i)di - MC*q\right]=1 \\ &v(i)=u(W+\tau^*(i)-L)-u(W-p(i)+\tau^*(i)-L), \forall i\leq q\end{aligned} \tag{9}$$

The competitive equilibrium defined by these conditions is equivalent to the Pareto-Optimum in equation 4. First, observe that the ability of the monopolist to extract full consumer surplus implies that: $p(i) \approx \frac{v(i)}{u'(c)}$. Therefore, in competitive equilibrium, $\frac{v(q^*)}{u'(c)} = MC$, which matches the condition for first-best efficiency.

Second, in competitive equilibrium, consumption in each state is equal to:

$$c = W + \pi - r - \int_0^q p(x)dx - \sigma L = W - r - MC*q - \sigma L \tag{10}$$

This is identical to competitive equilibrium consumption.

Finally, taking a first-order approximation to the condition for first-best research yields:

$$\frac{1}{\rho'(r)} = \int_0^q \frac{v(i)}{u'(c)} di - MC * q \tag{11}$$

Since $p(i) \approx \frac{v(i)}{u'(c)}$, this condition is met in competitive equilibrium.

B.3 The Competitive Equilibrium Without Price Discrimination

Notice that the competitive equilibrium with a monopolist patent holder and indemnity insurance only produced the first best outcome when the monopolist was allowed to price discriminate. The ability to set to prices based on the health (willingness to pay) of consumers was the key to achieving dynamic efficiency. However, it is likely that legal, political and social restrictions impede the monopolist's ability to price discriminate, especially when sicker consumers have higher willingness to pay. This naturally raises the question: Can the first-best be achieved when legal, social or political restrictions prohibit price discrimination?

In this section we show that the first best can be achieved even with restrictions on price discrimination. We argue that health insurance enables the monopolist to solve the dynamic efficiency problem even when the monopolist is not allowed to charge higher prices to consumers with a higher willingness to pay. If structured properly, an insurance contract can function as a two-part pricing scheme, where an insurer allows its insureds to pay marginal cost for drugs in the form of a co-payment, but then transfers an upfront payment to the drug manufacturer that is equal to the drug's total social value. This scheme leads to the first-best level of innovation *and* the first-best level of drug utilization[2]. The key to the success of this scheme is that consumers do not know their willingness to pay ex-ante. This makes drug purchase a risky decision and therefore creates a demand for insurance. The insurance market in turn enables the monopolist to extract consumer surplus.

---

[2] This result only holds when the consumption of the innovation has no external effects. Intervention in this market might be warranted when the innovation has consumption externalities. Philipson and Mechoulan (2005) discuss appropriate market interventions in the presence of technological change and consumption externalities.

Consider a health insurance contract where consumers are charged a premium I and pay copay m for the purchase of the innovation. Consumers also receive insurance pay-outs $K(i)$ depending upon their health i. We assume that the health insurance market is competitive and insurers make zero profits. The monopolist charges a fixed fee F to supply the innovation to insurers. We show that this market produces the first best outcome with the following insurance contract and fixed fee:

$$
\begin{aligned}
&I^* = \left[\frac{u(W - r^* - MC * q^* - \sigma L) + \int_0^{q^*} v(i)di - u(W - r^*)}{\rho(r^*)u'(W - r^* - MC * q^* - \sigma L) + (1 - \rho(r^*))u'(W - r^*)} \right] + \sigma L + MC * q^* \\
&m^* = MC \\
&K(i) \quad = L \ if \ i \ \in \left(q^*, \sigma\right) \\
&\qquad\quad = L + MC \ if \ i \ \hat{I} \left(0, q^*\right) \\
&F^* = I^* - \sigma L
\end{aligned}
\tag{12}
$$

Under this contract insurers make zero profits as they pass their entire surplus to the monopolist as a fixed fee for purchasing the innovation. Also, notice that copays equal marginal costs under this insurance contract. Therefore consumers with $i < q^*$ consume the product. This is the condition for first-best utilization, where consumers with willingness to pay below marginal costs are excluded from the market for the innovation.

The profits of the monopolist under this contract are:

$$
\begin{aligned}
\pi^* &= F^* - MC * q^* \\
&= \left[\frac{u(W - r^* - MC * q^* - \sigma L) + \int_0^{q^*} v(i)di - u(W - r^*)}{\rho(r^*)u'(W - r^* - MC * q^* - \sigma L) + (1 - \rho(r^*))u'(W - r^*)} \right]
\end{aligned}
\tag{13}
$$

Equation (13) shows that under this insurance contract and fixed fee the monopolist is able to extract the entire social surplus due to the innovation. Clearly, this contract maximizes profits from the production of the innovation, as the social surplus from the innovation is the maximum profit that can be extracted from the innovation.

The risk-neutral monopolist chooses r to maximizes profits from R&D:

$$\max_{r} \pi^{R\&D} = \rho(r)\pi^* - r \tag{14}$$

The first order condition for the monopolist is:

$$\pi^* = \frac{1}{\rho'(r)} \tag{15}$$

Substituting for π^* in equation (15) shows that the first order condition for the monopolist is exactly identical to the first order condition for maximizing dynamic efficiency.

Finally, consumption in each health state given this insurance contract is:

$$\begin{aligned} &W - r + \pi^* - I^*, \text{ if i } \in (\sigma,1) \\ &W - r + \pi^* - I^* + K(i), \text{ if i } \in (\sigma,\mathrm{q}^*) \\ &W - r + \pi^* - I^* + K(i) - m^*, \text{ if i } \in (0,\mathrm{q}^*) \end{aligned} \tag{16}$$

Substituting the insurance contract from equation (12) and the monopolist profits from equation (13) in the above equation yields that consumption in each health state i is:

$$c(i) = W - r - \sigma L - MC * q^* \tag{17}$$

Thus, this insurance contract and fixed fee also yields full insurance for consumers, consequently maximizing consumer surplus. Therefore the insurance contract and fixed fee schedule characterized in equation (12) yields the first best outcome as:

- Utilization of the innovation is optimal
- Investment in R&D is optimal
- The monopolist maximizes profits, consumers maximize expected utility, and a competitive insurance industry earns zero profits

C. Second-Best With Hidden Information

Often, and particularly in the case of health, it is very difficult to verify the extent of illness or the true demand for a health-care innovation (Arrow, 1963). However, it is extremely easy to verify whether a consumer chooses to use an innovation. One often observes health insurance contracts that reimburse consumers when they use an innovation, but it is very rare to find a "true indemnity" contract where consumers are reimbursed based on their underlying health state. As a result, we now consider the case where contracts can be made contingent on a consumer's decision to purchase the innovation, but not on the true state i. The incompleteness of information means that we will no longer attain the first-best Pareto Optimum, but we can analyze the second-best efficient allocation and a competitive equilibrium.

C.1 The Pareto Optimum

Define c_D as the consumption of a demander (gross of health losses), and c_N as consumption for a non-demander (also gross of health losses). Insurance contracts can be written on the basis of observed demand, but not on the basis of type i. The second-best efficient allocation of resources maximizes expected utility for consumers, subject to resource constraints, and the incentive compatibility of the chosen allocation. Incentive compatibility requires that sick demanders are exactly as well off as the marginal non-sick demander; otherwise, there are incentives for consumers around the margin to "cheat" by picking the other group's allocation. The second-best allocation thus solves:

$$\begin{aligned} &\max_{q,c_D,c_N,r}\ \rho(r)\left[qu(c_D - L) + (\sigma - q)u(c_N - L) + (1-\sigma)u(c_N) + \int_0^q (v(i))di\right] + \\ &(1-\rho(r))[\sigma u(W - r - L) + (1-\sigma)u(w - r)] \\ &s.t.\ qc_D + (1-q)c_N + MC * q \le W - r \\ &\quad u(c_D - L) + v(q) = u(c_N - L) \end{aligned} \tag{18}$$

In addition to the two constraints (at equality), the second-best efficient allocation is characterized by four first-order conditions, where $\rho(r)\lambda$ and $\rho(r)\mu$ are the (scaled) Lagrange multipliers associated with the resource and incentive compatibility constraints, respectively:

$$
\begin{aligned}
&[q] : \left(u(c_D - L) - u(c_N - L) + v(q)\right) = \lambda\left(MC + c_D - c_N\right) - \mu v'(q) \\
&[c_D] : \left(qu'(c_D - L)\right) = \lambda q - \mu u'(c_D - L) \\
&[c_N] : \left[(\sigma - q)u'(c_N - L) + (1 - \sigma)u'(c_N)\right] = \lambda(1 - q) + \mu u'(c_N - L) \\
&[r] : \rho'(r)\left[\begin{array}{l}\left(qu(c_D - L) + (\sigma - q)u(c_N - L) + (1 - \sigma)u(c_N) + \int_0^q (v(i))di\right) - \\ \left[\sigma u(W - r - L) + (1 - \sigma)u(W - r)\right]\end{array}\right] = \\
&(1 - \rho(r))\left[\sigma u'(W - r - L) + (1 - \sigma)u'(W - r)\right] + \rho(r)\lambda
\end{aligned} \tag{19}
$$

While it is not possible to solve for an explicit equilibrium without imposing functional form restrictions, several qualitative results can be proven from the equilibrium conditions. The second-best equilibrium involves: (1) Partial but incomplete insurance, (2) More than first-best utilization of the innovation, but (3) Less than first-best investment in research.

The incentive compatibility constraint proves there cannot be complete insurance, since the constraint requires that $c_D < c_N$. However, there is some insurance provided to the demanders of the innovation.

First note that the marginal utility of wealth is less than the marginal utility of consumption in the poorest state, or $\lambda < u'(c_D - L)$.[3] As a result, the multiplier μ must be negative,

---

[3] Using the conditions for c_D and c_N to solve for λ yields: $\lambda = u'(c_D - L)\dfrac{\sigma + (1 - \sigma)\dfrac{u'(c_N)}{u'(c_N - L)}}{q + (1 - q)\dfrac{u'(c_D - L)}{u'(c_N - L)}}$. Since $c_D < c_N$, $\dfrac{u'(c_D - L)}{u'(c_N - L)} > 1$. In addition, $\dfrac{u'(c_N)}{u'(c_N - L)} < 1$. Therefore, the numerator is strictly less than unity, while the denominator is strictly greater than unity. The result then follows.

according to the first order condition for c_D. Some algebraic manipulation of the six equilibrium conditions allows us to express the resource constraint multiplier in terms of equilibrium quantities:

$$\lambda(MC + c_D - c_N) = \mu v'(q) \tag{20}$$

Since $\lambda > 0$ and $\mu < 0$, it must be true that $c_D + MC > c_N$, so that demanders do not bear the full cost of the innovation.

The latter result also implies that utilization exceeds the first-best level. Incentive compatibility requires that private marginal cost equals private marginal benefit, in the sense that $v(q) = u(c_N - L) - u(c_D - L)$. Insurance implies that private marginal cost is less than social marginal cost. As a result, we end up with over-utilization. This is the classic moral hazard that results when underlying demand is unobservable.

Finally, innovation must be less than first-best. Due to hidden information, second-best maximum social surplus will be strictly less than the first best, and so will the returns to innovation. The result is less innovation, even though moral hazard induces over-utilization compared to the first-best.

C.2 The Competitive Equilibrium

Since information is hidden, the monopolist cannot practice perfect price-discrimination. However, since consumers are ex ante identical, it can engage in two-part pricing, which can also lead to the extraction of consumer surplus. Suppose there is a perfectly competitive insurance industry. The innovator charges each insurer an upfront license fee F, but then sells each unit of output for a constant price p. Insurers sell insurance policies to consumers for the ex ante insurance premium I, which entitles the consumer to purchase the innovation from the insurer at the co-payment m. Markets arranged in this way will produce the second-best efficient outcome with hidden information.

Note that this arrangement is equivalent to one in which the innovator charges an ex ante license fee and an ex post unit price to *consumers*, who can then purchase insurance contracts that pay out contingent on purchase of the drug. We choose to model the insurer as an intermediary simply because it is closer to the way health care markets actually function.

C.2.1 ***Second-Best Efficiency of Utilization***

In competitive equilibrium, the consumer chooses the states of the world in which to purchase the innovation, taking as given the insurance contract offered by the insurance industry.

$$\max_q \; qU(W+\pi-L-m-I)+(\sigma-q)U(W+\pi-L-I)+(1-\sigma)U(W+\pi-I)+\int_0^q v(i)di \tag{21}$$

The consumer's optimal utilization decision sets the marginal benefit of the innovation equal to the private marginal cost:

$$v(q)=U(W+\pi-L-I)-U(W+\pi-L-m-I) \tag{22}$$

Note that this is equivalent to the second-best utilization condition: $v(q)$ is equal to the difference in utility across the consuming and non-consuming states of illness.

C.2.2 ***Second-Best Efficiency of State-Specific Consumption***

Taking as given the offer of the innovator, the representative insurer maximizes profits by choosing its contract parameters and its purchases of the good from the innovator, subject to consumers' participation in the insurance market, ex post incentive compatibility,[4] and demand function $d(m)$.

$$\max_{I,m} I + md(m)^* - F - pd(m)^*$$
$$s.t.\ d(m)u(W + \pi - L - m - I) + (\sigma - d(m))u(W + \pi - L - I) + (1 - \sigma)u(W + \pi - I) + \int_0^{d(m)} v(i)di$$
$$\geq \sigma u(W + \pi - L) + (1 - \sigma)u(W + \pi)$$
$$u(W + \pi - L - I) - u(W + \pi - L - m - I) \leq v(q)$$
(23)

Note that maximizing profits subject to a reservation utility condition is equivalent to maximizing utility subject to a nonnegativity constraint on profits (i.e., a "reservation profits level").[5] Moreover, choosing a co-payment subject to a demand function is equivalent to choosing a level of quantity subject to remaining on the demand function. Defining willingness to pay as $T(q)$ (a scalar multiple of $v(q)$), staying on the demand function requires that $m = T(q)$. Therefore, we can rewrite the problem as one whose notation conforms more closely to the second-best Pareto problem.

$$\max_{q,I} qu(W + \pi - L - T(q) - I) + (\sigma - q)u(W + \pi - L - I) + (1 - \sigma)u(W + \pi - I) +$$
$$\int_0^q v(i)di$$
$$s.t.\ F + pq \leq I + qT(q)$$
$$u(W + \pi - L - I) - u(W + \pi - L - T(q) - I) \leq v(q)$$
(24)

In addition to the two constraints at equality, the first order conditions for this problem are:

---

[4] Consumers purchasing the invention must be better off purchasing than not purchasing, from their ex post point of view.

[5] This problem is associated with an equilibrium condition that sets the consumer's expected utility equal to her reservation utility level. Note that the profit-maximizing version of the problem is analogously associated with a zero profit condition in equilibrium.

$$[I]: \lambda - \mu[u'(W + \pi - L - T(q) - I) - u'(W + \pi - L - I)] =$$
$$qu'(W + \pi - L - T(q) - I) + (\sigma - q)u'(W + \pi - L - I) + (1 - \sigma)u'(W + \pi - I)$$
$$[q]: T'(q)[\mu u'(W + \pi - L - v(q) - I) + qu'(W + \pi - L - v(q) - I) - \lambda q] + \lambda(p - T(q))$$
$$= \mu v'(q) \tag{25}$$

The competitive condition for I matches the sum of the second-best conditions for c_N and c_D. The competitive condition for q matches the sum of the second-best conditions for q and c_D (which is equivalent to the term in square brackets being zero), provided that the innovator prices at marginal cost.

Consumption allocations are consistent with the second-best, because the innovator will choose to price at marginal cost in competitive equilibrium. The innovator's ex post decision involves maximizing her profits subject to the participation of insurance companies. Define $\pi^I(F, p)$ as the maximum profits earned by the representative insurer when faced with the fixed fee F and supply price p. The innovator's problem can be expressed as:

$$\begin{aligned} &\max_{p,F} F + pq(p) - MCq(p) \\ &s.t.\ \pi^I(F, p) \geq 0 \end{aligned} \tag{26}$$

This has the first order conditions:

$$\begin{aligned} &[F]: 1 + \lambda\pi_F^I = 0 \\ &[p]: q(p) + pq'(p) - MCq'(p) + \lambda\pi_p^I = 0 \end{aligned} \tag{27}$$

Applying the envelope theorem to the insurer's profit function yields the results that $\pi_p^I = -q$ and $\pi_F^I = -1$. This then implies that $p = MC$.

C.2.3 *Second-Best Efficiency of the Innovation Decision*

The innovator invests in research to maximize ex ante profits, according to:

$$\begin{aligned} &\max_{p,F,r} \ \rho(r)[F + pq(p) - MCq(p)] - r \\ &s.t. \ \pi^I(F, p) \geq 0 \end{aligned} \tag{28}$$

Competitive innovation is characterized by:

$$\rho'(r)[F + q(p - MC)] = \rho'(r)[I + q(m - MC)] = 1 \tag{29}$$

Since the consumer's receives only her reservation utility level, the insurer extracts all her surplus. This implies that I is equal to net consumer surplus at the co-payment, or $I = \frac{\int_0^q v(i)di}{\lambda} - mq$. Therefore, competitive ex post profits satisfy:

$$\pi = \frac{\int_0^q v(i)di}{\lambda} - MC * q \tag{30}$$

The private return to innovation shown above matches the social return, given in the first order condition for r, in equations 19. In competitive equilibrium, the social return to innovation can be written as:

$$\frac{1}{\lambda}\left[\begin{aligned} &\left(-qv(q) + \sigma u(W - (m - MC)q - r - L) + (1 - \sigma)u(W - (m - MC)q - r) + \int_0^q (v(i))di\right) \\ &- [\sigma u(W - r - L) + (1 - \sigma)u(W - r)] \end{aligned}\right] \tag{31}$$

This simplification also embeds three results: (1) The consumer indifference condition; (2) Zero profits in the insurance industry; and (3) Equal marginal utility across the "discovery" and "non-discovery" states. In a world with many simultaneous innovation projects underway, there is minimal social risk posed by the success or failure of any single innovation project. Therefore, we can safely regard marginal utility as invariant to the success or failure of any one innovation.

We can monetize the above expression by taking first-order approximations of the differences in utility. This procedure, along with the fact that $m \approx \frac{v(q)}{\lambda}$, yields the final simplification:

$$Social\,return = \left[\frac{\int_0^q v(i)di}{\lambda} - MC * q \right] \tag{32}$$

This is the same as the private return to innovation and guarantees the second-best level of research.

C.3 The Competitive Equilibrium Without Ex Ante Pricing

In addition to its usual static inefficiencies, incompleteness in the insurance market has adverse effects on *dynamic* efficiency in innovation. If some people are uninsured, or if insurers have market power, monopolists may not be able to write and enforce efficient pricing contracts with insurers. It is infeasible for innovators to write contracts with every potential consumer ex ante. Therefore, if some consumers do not participate in an insurance pool, they are not open to such two-stage contracts. Public insurance systems also pose problems for this type of contracting: Medicare or Medicaid may be quite reluctant to hand over all surplus to an innovator. More generally, innovators may worry about the threat of price regulation if they attempt to extract the full value of consumer surplus from insurers.

For these and related reasons, two-stage contracting may be infeasible.

The alternative is for the monopolist to sell directly to consumers at a fixed price, while consumers can purchase insurance payable in the event of purchase[6]. The result is the under-

[6] Gaynor, Haas-Wilson and Vogt (2000) consider this case and show that despite moral hazard induced by competitive health insurance markets, a reduction in the price of medical care is

provision of innovation and a decline in social surplus. Utilization may be greater or less than second-best utilization, depending on the relative size of the monopolist's incentive to restrict quantity versus consumers' willingness to subsidize ex post consumption of the innovation through an insurance premium.

Consider an environment where an insurer can reimburse a consumer if he purchases a product, but not otherwise. However, the product is sold directly by the innovator to the consumer. The innovator sells at the single price p. Insurers sell contracts that reduce the ex post price in exchange for an actuarially fair ex ante payment; reducing the ex post price by τ costs $I = \tau q$ ex ante. The consumer chooses a level of insurance and ex post consumption maximize utility. She chooses from an array of actuarially fair insurance contracts, and she can choose only time-consistent insurance contracts, where she has no incentive to deviate ex post. The latter requirement implies that ex ante consumption decisions must maximize utility ex post.

$$\max_{q,\tau,I} qu(W + \pi - L - (p - \tau) - I) + (\sigma - q)u(W + \pi - L - I) + (1 - \sigma)u(W + \pi - I) + \int_0^q v(i)di$$
$$s.t.\ I \geq \tau q$$
$$u(W + \pi - L - (p - \tau) - I) + v(q) \geq u(W + \pi - L - I)$$
(33)

The consumer's optimality conditions are given by:

$$[q] : \lambda\tau = \mu v'(q)$$
$$[\tau] : \lambda q = (q + \mu)u'(W + \pi - L - (p - \tau) - I)$$
$$[I] : \lambda + \mu(u'(W + \pi - L - I) - u'(W + \pi - L - (p - \tau) - I)) = qu'(W + \pi - L - (p - \tau) - I) +$$
$$(\sigma - q)u'(W + \pi - L - I) + (1 - \sigma)u'(W + \pi - I)$$
(34)

---

always welfare enhancing. However they do not consider the role of higher prices in encouraging innovation.

In this environment, the consumer's decisionmaking is efficient, even though the innovator's might not be. Formally, the consumer's first-order conditions match the conditions for social efficiency, provided that they face efficient pricing (i.e., at marginal cost): the condition for q matches the second-best efficiency condition for q; the condition for τ matches the second-best condition for c_D; the condition for I matches the sum of the second-best conditions for c_D and c_N.

Departures from the second-best originate in the innovator's problem, when she is unable to extract consumer surplus through a two-stage pricing strategy. The monopolist maximizes profits, taking as given the consumer's insurance and demand decisions. Approximating the marginal utility of consumption for sick demanders as γ, the monopolist solves:

$$\begin{aligned} &\max_{q,p}\ pq - MC * q \\ &s.t.\ \ p = \left(\frac{v(q)}{\gamma} + \tau\right) \end{aligned} \tag{35}$$

This has the first order condition:

$$\frac{v'(q)q}{\gamma} + \left(\frac{v(q)}{\gamma} + \tau\right) = MC \tag{36}$$

The innovator has the standard incentives of a monopolist. Price exceeds marginal cost, because of the incentive to raise price by restricting quantity. However, unlike the standard monopoly problem, utilization may be above or below the first- or second-best level, because consumers face the price $\frac{v(q)}{\gamma} = p - \tau < p$. Utilization can be either above or below the second-best level, depending on the shape of the willingness to pay function and the consumer's desire for insurance. It is not possible to determine the impact on utilization without specific assumptions on functional form (Garber, Jones, and Romer, 2005).

However, it is clear that consumer welfare in every state of the world is lower than in the second-best, because the monopolist charges a unit price that is higher than marginal cost.[7] It is also clear that the monopolist's profits are lower than when she has access to two-stage pricing. The absence of two-stage contracting thus leads to inefficiently low levels of innovation, consumer welfare, and social surplus.

Moreover, it is also clear that adding an insurance market improves consumer welfare, increases the profits of the innovator, and raises the level of innovation, compared to the equilibrium that would exist without such a market. This leads to the following welfare-ordering:

$$\begin{aligned}&No\ Insurance <_W \\ &Insurance + Monopoly\ pricing <_W \\ &Insurance + Two - Stage\ pricing =_W \\ &Pareto\ Optimum\end{aligned} \tag{37}$$

D. Incremental Innovation

Above, we considered the case of a brand-new innovation. In practice, new innovations must compete with existing innovations for customers, and monopolistic competition is the norm. Our basic results are unchanged if the incumbent's profits remain the same after the entry of the new firm. However, in cases where the new entrant is able to cannibalize some of the incumbent's profits, the result is too much innovation, because the entrant spends on innovation simply to initiate a transfer of resources, rather than a creation of wealth. Efficiency can be maintained if the entrant is made to pay an ex post tax on profits that is equal to the decline in profits of the incumbent.

[7] The first-order effect of this is to lower utility in every state of the world: even non-consuming states face higher insurance premia.

D.1 The Pareto Optimum

An incremental innovation can be thought of in the context of the spatial model of consumers developed earlier. The case where a new innovation supplants an older innovation entirely is very straightforward and behaves exactly as the earlier single-innovator case. Consider the more complex case with a new innovation that is an improvement for some consumers, but not for other consumers. The new innovation thus splits up the market with the original innovation. Recall that the utility value of the original innovation was $v(i)$. Define $v_N(i)$ as the value of the new innovation, where there exists i_N such that $v_N(i) > v(i)$ for $i < i_N$ and $v_N(i) \le v(i)$ for $i \ge i_N$. To ensure that this point is unique, we also assume that $v_N' < v_O' < 0$. Without loss of generality, suppose that $i_N < q$, so that some consumers will use the original innovation even after the new one enters the market.[8]

Sick consumers now have three choices: purchase the original innovation O, purchase the new innovation N, or purchase nothing at all. The utility function can now be written as:

$$\begin{array}{l} u(c_D - L), i \ge \sigma \\ u(c_D^O - p_O) - L + v(i), i \le \sigma, \text{purchase } O \\ u(c_D^N - p_N) - L + v_N(i), i \le \sigma, \text{purchase } N \\ u(c_D - L), i \le \sigma, \text{purchase nothing} \end{array} \tag{38}$$

Suppose that the marginal cost of producing the new innovation is the same as the old, so that the value of the new innovation lies purely on the demand side. Define r_N as research into the new innovation, q_N as the quantity of the new innovation sold, and q_O as the quantity of the old innovation sold if the new one is discovered. Finally, define $\overline{U}(r)$ as the level of expected utility enjoyed if research fails to yield a new innovation, but the old innovation is

---

[8] Failure of this assumption necessitates analysis of another case whose results are largely similar.

available. This is the level of utility yielded by the earlier, single-innovation equilibrium; it involves two-part pricing by the innovator and partial insurance for consumers. Maximum social surplus is obtained as the solution to:[9]

$$
\begin{aligned}
&\max_{r,q_N,q_O,c_D^N,c_D^O,c_N} \rho(r)\Big\{q_N u(c_D^N - L) + q_O u(c_D^O - L) + (\sigma - (q_N + q_O))u(c_N - L) + (1-\sigma)u(c_N) + \\
&\int_0^{q_N} v_N(i)di + \int_{q_N}^{q_N+q_O} v_O(i)di\Big\} + (1-\rho(r))\overline{U}(r) \\
&s.t.\ q_N c_D^N + q_O c_D^O + (1-(q_N+q_O))c_N + MC*(q_O+q_N) \le W - r\ [\rho(r)\lambda] \\
&u(c_D^N - L) + v_N(q_N) = u(c_D^O - L) + v_O(q_N)\ [\rho(r)\mu_N] \\
&u(c_D^O - L) + v_O(q_N + q_O) = u(c_N - L)\ [\rho(r)\mu_O]
\end{aligned}
\tag{39}
$$

Equilibrium is characterized by the following first order conditions:

$$
\begin{aligned}
&[q_N] : -\lambda\left(c_D^N - c_N + MC\right) + \mu_N\left(v_N'(q_N) - v_O'(q_N)\right) + \mu_O v_O'(q_N + q_D) = 0 \\
&[q_O] : -\lambda\left(c_D^O - c_N + MC\right) + \mu_o\left(v_O'(q_N + q_O)\right) = 0 \\
&[c_D^N] : q_N u'(c_D^N - L) - \lambda q_N + \mu_N u'(c_D^N - L) = 0 \\
&[c_D^O] : q_O u'(c_D^O - L) - \lambda q_O - \mu_N u'(c_D^O - L) + \mu_O u'(c_D^O - L) = 0 \\
&[c_N] : (\sigma - (q_N + q_O))u'(c_N - L) + (1-\sigma)u'(c_N) - \lambda(1-(q_N+q_O)) - \mu_O u'(c_N - L) = 0 \\
&[r] : \rho'(r)\Big\{q_N u(c_D^N - L) + q_O u(c_D^O - L) + (\sigma - (q_N + q_O))u(c_N - L) + (1-\sigma)u(c_N) + \\
&\qquad \int_0^{q_N} v_N(i)di + \int_{q_N}^{q_N+q_O} v_O(i)di - \overline{U}\Big\} = \rho(r)\lambda - (1-\rho(r))\overline{U}'(r)
\end{aligned}
\tag{40}
$$

The key difference between this case and the earlier case lies in the level of optimal research spending. When other inventions are present, the efficient return to research is equal to the *increment* in social surplus induced by the new innovation. As a result, optimal research

[9] Note that there is also a third, non-binding, incentive compatibility constraint that guarantees that the marginal user of the new innovation is better off than the marginal non-user of either innovation.

spending is lower when other inventions are present, because the incremental gain in social surplus from the innovation is less.

At the second-best allocation: (1) More insurance is provided for the newer innovation, because it confers more value on its users than the original innovation; (2) Partial insurance is provided to users of both innovations; (3) Both inventions are over-utilized relative to the first-best; and (4) Innovation is less than first-best.

<u>*More insurance is provided for the newer innovation.*</u> We prove this by contradiction. Suppose that $c_D^N \leq c_D^O$. Since the marginal utility of ex post wealth must be less than the marginal utility in the most impoverished state, it must be true that $\lambda < u'(c_D^N - L)$.[10] The first order condition for c_D^N then implies that $\mu_N < 0$. Adding up the first order conditions for q_N and q_O yields $\lambda(c_D^N - c_D^O) = \mu_N(v_N'(q_N) - v_O'(q_N))$. Therefore, since $v_N'(q_N) - v_O'(q_N) < 0$, it must be true that $c_D^N > c_D^O$, but this contradicts our original assumption and proves the claim.

In this particular case, the equilibrium users of the new innovation have a higher average willingness to pay for the innovation than the users of the old innovation. Therefore, they are more willing to transfer resources to the states in which the new innovation is used than those

---

[10] Using the conditions for c_N, c_D^O, and c_D^N to solve for λ yields:

$$\lambda = u'(c_D^N - L)\frac{\sigma u'(c_N - L) + (1-\sigma)u'(c_N)}{(1-(q_N + q_O))u'(c_D^N - L) + q_N u'(c_N - L) + q_O u'(c_N - L)\dfrac{u'(c_D^N - L)}{u'(c_D^O - L)}}$$

By assumption, $u'(c_D^N - L) \geq u'(c_D^O - L)$. Moreover, the incentive compatibility constraints require that $c_N > c_D^N$. Therefore, the denominator is strictly greater than $u'(c_N - L)$, but the numerator is strictly smaller than this quantity. This implies the result.

in which the old innovation is used. This forms the basis of our result. It is not generically true that more insurance is provided to the newer innovation, but it is always true that more insurance flows to the invention with the higher-value users.

<u>*Partial insurance is provided for both innovations.*</u> Since $c_D^N < c_D^O$ and $\lambda(c_D^N - c_D^O) = \mu_N(v_N'(q_N) - v_O'(q_N))$, it must be true that $\mu_N < 0$. Moreover, since $\lambda < u'(c_D^O - L)$, the condition for c_D^O implies that $\mu_N > \mu_O$, which ensures that $\mu_O < 0$ as well. Since $\mu_O < 0$, the condition for q_O implies that $MC > c_N - c_D^O$, which implies partial insurance for the old innovation. Since $\mu_O, \mu_N < 0$, the condition for q_N implies that $MC > c_N - c_D^N$, which implies partial insurance for the new innovation as well.

This finding implies the last two results. Since partial insurance is provided for both innovations, *both inventions are over-utilized* in the sense that some individuals use them whose benefit is less than marginal cost. Finally, since incomplete information lowers the total ex post consumer surplus, *innovation will be less than the first-best* in this context.

D.2 The Competitive Equilibrium

With more than one firm present, the efficiency of the competitive equilibrium depends on how the incumbent innovator responds to the new innovator's entry. The two-stage pricing contract yields the second-best allocation of resources if the incumbent's profits are unaffected. If, however, the new innovator captures some of the incumbent's profits, the result is excessive innovation, as the new innovator seeks to capture some of the incumbent's rents. Efficiency can be restored if the new innovator is charged a lump-sum tax on ex post profits, equal to the change in the incumbent's profits.

Insurers now sell two insurance contracts. The first transfers resources to the insured when she buys innovation O, and the second transfers resources if innovation N is purchased.

Since the contracts can be offered separately, insurers must make zero profits on each of them. In turn, insurers contract with the innovators to purchase the right to buy the innovation. Innovators employ a two-part pricing strategy, where they charge an ex ante fee coupled with an ex post unit price.

D.2.1 *Second-Best Efficiency of Utilization*

In competitive equilibrium, the consumer chooses the states of the world in which she purchases an innovation, and which innovation she purchases. She takes as given the insurance contracts offered by the insurance industry. The contract associated with the new innovation N is defined by the premium, copayment pair (I_N , m_N), while (I_O , m_O) define the contract for the old innovation. The consumer solves:

$$\begin{aligned}&\max_{q_N, q_O} q_N u(W + \pi - L - m_N - I_N - I_O) + q_O u(W + \pi - L - m_O - I_N - I_O) + \\&(\sigma - (q_N + q_O))u(W + \pi - L - I_N - I_O) + (1 - \sigma)u(W + \pi - I_N - I_O) + \\&\int_0^{q_N} v_N(i)di + \int_{q_N}^{q_N + q_O} v_O(i)di\end{aligned} \tag{41}$$

The consumer's optimal utilization decision sets the marginal benefit of the innovation equal to the private marginal cost:

$$\begin{aligned}&[q_O] : v_O(q_N + q_O) = u(W + \pi - L - I_N - I_O) - u(W + \pi - L - m_O - I_N - I_O)\\&[q_N] : v_N(q_N) = u(W + \pi - L - I_N - I_O) - u(W + \pi - L - m_N - I_N - I_O) + v_O(q_N) - v_O(q_N + q_O)\end{aligned} \tag{42}$$

The first condition is identical to the second-best incentive compatibility condition for q_O. The second is identical to a linear combination of the two second-best incentive-compatibility conditions for q_O and q_N.

D.2.2 *Second-Best Efficiency of State-Specific Consumption*

Insurers now sell two insurance contracts, one for each of the innovations available. Taking as given the offers of the innovators, the representative insurer decides how to price its contracts by maximizing its profits, subject to a reservation utility level U_R for consumers, ex post incentive-compatibility, and the consumer's demand function. This yields the problem:

$$
\begin{aligned}
&\max_{I_N, I_O, m_N, m_O} I_N + I_O + m_N d(m_N) + m_O d(m_O) - F_N - F_O - p_N F_N - p_O F_O \\
&s.t.\ EU(I_N, I_O, m_N, m_O) \geq U_R \\
&u(W + \pi - L - T_N(q_N) - I_N - I_O) + v_N(q_N) \geq u(W + \pi - L - T_O(q_N + q_O) - I_N - I_O) + v_O(q_N) \\
&u(W + \pi - L - T_O(q_N + q_O) - I_N - I_O) + v_O(q_N + q_O) \geq u(W + \pi - L - I_N - I_O)
\end{aligned} \tag{43}
$$

In equilibrium, profits on each of the types of insurance contracts will be zero. The reservation utility level U_R is the level of utility the consumer obtains from contracting only with the incumbent monopolist.[11] Since any insurer has the option of providing a single insurance contract, this constraint must be satisfied.

As discussed in Section C.2.2, the insurer's profit-maximization problem is equivalent to maximizing the consumer's utility subject to a nonnegativity constraint on profits, ex post incentive-compatibility, and consumer demand. In this case, there are two nonnegativity constraints on profits, corresponding to the zero-profit equilibrium conditions. Satisfying the consumer's demand conditions requires that the copayment rate equal the marginal person's willingness to pay. In equilibrium, the consumer's utility must be equal to U_R. Defining $T_N(q_N)$ and $T_O(q_N + q_O)$ as the willingness to pay functions, we can write the representative insurer's problem as:

[11] Without loss of generality, we assume that this level of utility is higher than contracting exclusively with the new entrant, and higher than utility from no insurance at all.

$$
\begin{aligned}
&\max_{q_N, I_N, q_O, I_O} q_N u(W + \pi - L - T_N(q_N) - I_N - I_O) + q_O u(W + \pi - L - T_O(q_N + q_O) - I_N - I_O) + \\
&(\sigma - q_N - q_O) u(W + \pi - L - I_N - I_O) + (1 - \sigma) u(W + \pi - I_N - I_O) + \int_0^{q_N} v_N(i)di + \int_{q_N}^{q_N + q_O} v_O(i)di \\
&s.t. \; I_N + q_N T_N(q_N) - F_N - p_N q_N \geq 0 \\
&\quad I_O + q_O T_O(q_N + q_O) - F_O - p_O q_O \geq 0 \\
&u(W + \pi - L - T_N(q_N) - I_N - I_O) + v_N(q_N) \geq u(W + \pi - L - T_O(q_N + q_O) - I_N - I_O) + v_O(q_N) \\
&u(W + \pi - L - T_O(q_N + q_O) - I_N - I_O) + v_O(q_N + q_O) \geq u(W + \pi - L - I_N - I_O)
\end{aligned}
\tag{44}
$$

In addition to the four constraints, and an equilibrium condition that guarantees consumer utility U_R, the first order conditions for this problem are:

$$
\begin{aligned}
&[I_N] : \lambda_N - \mu_N[u'(W + \pi - L - T_N(q_N) - I_O - I_N) - u'(W + \pi - L - T_O(q_N + q_O) - I_O - I_N)] \\
&- \mu_O[u'(W + \pi - L - T_O(q_N + q_O) - I_N - I_O) - u'(W + \pi - L - I_N - I_O)] = \\
&q_N u'(W + \pi - L - T_N(q_N) - I_N - I_O) + q_O u'(W + \pi - L - T_O(q_N + q_O) - I_N - I_O) + \\
&(\sigma - q_N - q_O) u'(W + \pi - L - I_N - I_O) + (1 - \sigma) u'(W + \pi - I_N - I_O) \\
&[I_O] : \lambda_O - \mu_N[u'(W + \pi - L - T_N(q_N) - I_O - I_N) - u'(W + \pi - L - T_O(q_N + q_O) - I_O - I_N)] \\
&- \mu_O[u'(W + \pi - L - T_O(q_N + q_O) - I_N - I_O) - u'(W + \pi - L - I_N - I_O)] = \\
&q_N u'(W + \pi - L - T_N(q_N) - I_N - I_O) + q_O u'(W + \pi - L - T_O(q_N + q_O) - I_N - I_O) + \\
&(\sigma - q_N - q_O) u'(W + \pi - L - I_N - I_O) + (1 - \sigma) u'(W + \pi - I_N - I_O) \\
&[q_N] : T_N'(q_N)[\mu_N u'(W + \pi - L - T_N(q) - I) + q_N u'(W + \pi - L - T_N(q_N) - I_N - I_O) - \lambda_N q_N] + \\
&T_O'(q_N + q_O)\begin{bmatrix} -\mu_N u'(W + \pi - L - T_O(q_N + q_O) - I_N - I_O) + \mu_O u'(W + \pi - L - T_O(q_N + q_O) - I_N - I_O \\ q_O u'(W + \pi - L - T_O(q_N + q_O) - I_N - I_O) - \lambda_N q_O \end{bmatrix} \\
&\lambda_N(p_N - T_N(q_N)) \\
&= \mu_N(v_N'(q_N) - v_O'(q_N)) + \mu_O(v_O'(q_N + q_O)) \\
&[q_O] : T_O'(q_N + q_O)\begin{bmatrix} \mu_O u'(W + \pi - L - T_O(q_N + q_O) - I_N - I_O) + \\ q_O u'(W + \pi - L - T_O(q_N + q_O) - I_N - I_O) - \lambda_O q_O \\ -\mu_N u'(W + \pi - L - T_O(q_N + q_O) - I_N - I_O) \end{bmatrix} + \lambda_O(p_O - T_O(q_N + q_O)) \\
&= \mu_O v_O'(q_N + q_O)
\end{aligned}
\tag{45}
$$

These competitive first order conditions are simply linear combinations of the efficiency conditions. Note first that the conditions for I_N and I_O imply that $\lambda_N = \lambda_O$, so that the first two competitive first order conditions are identical. Moreover, it is easily confirmed that the competitive conditions for the insurance premium are both equivalent to the sum of the efficiency conditions for c_D^N, c_D^O, and c_N. Moreover, provided that the competitive price p equals marginal cost, the other two competitive conditions are also equivalent to their efficiency counterparts. Examining the condition for q_N, the two terms in square brackets are zero according to the efficiency conditions for c_D^N and c_D^O, respectively. The remaining terms match the efficiency condition for q_N, provided that price equals marginal cost. Similarly, in the condition for q_O, the term in square brackets is zero according to the condition for c_D^O, and the remaining terms match the efficiency condition for q_O, provided price equals marginal cost.

It remains to verify that both innovators will choose unit prices that are equal to marginal cost. Each innovator's ex post decision involves maximizing her profits subject to the participation of insurance companies. Define $\pi^I(F, p)$ as the maximum profits earned (on a particular contract) by the representative insurer when faced with the fixed fee F and supply price p. Each innovator's problem can be expressed as:

$$\begin{aligned} &\max_{p,F}\ F + pq(p) - MCq(p) \\ &s.t.\ \pi^I(F, p) \geq 0 \end{aligned} \tag{46}$$

This has the first order conditions:

$$\begin{aligned} &[F] : 1 + \lambda\pi_F^I = 0 \\ &[p] : q(p) + pq'(p) - MCq'(p) + \lambda\pi_p^I = 0 \end{aligned} \tag{47}$$

Applying the envelope theorem to the insurer's profit function yields the results that $\pi_p^I = -q$ and $\pi_F^I = -1$. This then implies that $p = MC$.

D.2.3 *Second-Best Efficiency of the Innovation Decision*

The new innovator invests in research to maximize ex ante profits, according to:

$$\max_{p_N, F_N, r} \; \rho(r)\left[F_N + p_N q_N(p_N) - MCq_N(p_N)\right] - r$$
$$s.t. \;\; \pi^I(F_N, p_N) \geq 0 \tag{48}$$

Competitive innovation is characterized by:

$$\rho'(r)F_N = \rho'(r)\left[I_N + q\left(m_N - MC_N\right)\right] = 1 \tag{49}$$

By comparison, the second-best social return can be written as:[12]

$$\frac{1}{\lambda}\left(q_N u(W + \pi - L - m_N - I_N - I_O) + q_O u(W + \pi - L - m_O - I_N - I_O) + \right.$$
$$(\sigma - q_N - q_O)u(W + \pi - L - I_N - I_O) + (1 - \sigma)u(W + \pi - I_N - I_O) + \int_0^{q_N} v_N(i)di + \int_{q_N}^{q_N + q_O} v_O(i)di$$
$$\left. - \bar{U}\right) \tag{50}$$

Using the consumer's incentive-compatibility constraints and the insurer's zero-profit conditions, we can simplify this as:

$$\frac{1}{\lambda}\left(q_N\left(v_O(q_N) - v_N(q_N)\right) - (q_O + q_N)\left(v_O(q_N + q_O)\right) + \right.$$
$$\sigma u(W - L - q_N(MC - m_N) - q_O(MC - m_O)) + (1 - \sigma)u(W - q_N(MC - m_N) - q_O(MC - m_O)) +$$
$$\left.\int_0^{q_N} v_N(i)di + \int_{q_N}^{q_N + q_O} v_O(i)di - \bar{U}\right) \tag{51}$$

Using what we learned about $\bar{U}$ in the earlier analysis, this can be further simplified:

[12] We once again employ the assumption that the marginal utility of wealth λ is the same regardless of whether this particular innovation is discovered.

$$\frac{1}{\lambda}\Big(q_N\big(v_O(q_N)-v_N(q_N)\big)-(q_O+q_N)\big(v_O(q_N+q_O)\big)+qv(q)$$
$$\sigma u\big(W-L-q_N(MC-m_N)-q_O(MC-m_O)-r\big)+(1-\sigma)u\big(W-q_N(MC-m_N)-q_O(MC-m_O)-r\big)-$$
$$\sigma u\big(W-L-q(MC-m)-r\big)-(1-\sigma)u\big(W-q(MC-m)-r\big)$$
$$+\int_0^{q_N} v_N(i)di+\int_{q_N}^{q_N+q_O} v_O(i)di-\int_0^{q} v(i)di\Big) \tag{52}$$

Using a first-order approximation to utility, and relying on the fact that λ is the marginal utility of wealth ex post, we can write this as:

$$\frac{1}{\lambda}\Big\{q_N\big(v_O(q_N)-v_N(q_N)\big)-(q_O+q_N)\big(v_O(q_N+q_O)\big)+qv(q)+\int_0^{q_N} v_N(i)di+\int_{q_N}^{q_N+q_O} v_O(i)di-\int_0^{q} v(i)di\Big\}$$
$$+\big(q(MC-m)-q_N(MC-m_N)-q_O(MC-m_O)\big) \tag{53}$$

Under the assumption that $v_O \equiv v$, or that the new innovation does not change the value of the old innovation, the marginal user of the old innovation will be the same person, or $q_O + q_N = q$.[13] This allows the further simplification:

$$\frac{1}{\lambda}\Big\{q_N\big(v_O(q_N)-v_N(q_N)\big)+\int_0^{q_N}\big(v_N(i)-v_O(i)\big)di\Big\}+\big(q_N m_N+q_O m_O-qm\big) \tag{54}$$

Using the consumer’s incentive compatibility constraint for q_N to derive the approximation $v_O(q_N)-v_N(q_N)\approx\lambda(m_O-m_N)$ allows us to write:

$$\frac{1}{\lambda}\Big\{\int_0^{q_N}\big(v_N(i)-v_O(i)\big)di\Big\} \tag{55}$$

---

[13] Departures from the second-best will occur if there are complementarities between the two innovations. In this case, the new innovator may not be able to capture the enhancement in value of the old innovation from the other monopolist, unless of course there is a single monopolist inventing both products.

To a first-order approximation, the social return to the innovation is equal to the direct increment in consumer surplus enjoyed by the consumers of the new innovation.

We now consider the conditions under which this equals the new innovator's competitive return. In competitive equilibrium, the innovator earns ex post profits:

$$I_N + q_N(m_N - MC) \tag{56}$$

The consumer's reservation utility condition implies that $EU(I_N, I_O, m_N, m_O) - U_R = 0$. Taking first-order expansions around the points $I_N + I_O = I = 0$ yields:

$$I_N + I_O - I \approx \frac{1}{\lambda}\left\{\int_0^{q_N} (v_N(i) - v_O(i))di\right\} - q_N[m_N - m_O] \tag{57}$$

Defining $\Delta\pi_O$ as the change in the profits of the original innovator, we can write:

$$I_O - I = \Delta\pi_O + q_N(m_O - MC) \tag{58}$$

Therefore, the profits of the new innovator can be written as:

$$\pi_N = \frac{1}{\lambda}\left\{\int_0^{q_N} (v_N(i) - v_O(i))di\right\} - \Delta\pi_O \tag{59}$$

The new innovator has efficient incentives, as long as the profits of the incumbent remain unchanged. If, however, his profits fall, the new innovator has incentives to over-innovate. In this case, the entrant is inheriting some of the incumbent's profits. Since this is just a transfer rather than a real creation of resources, it gives rise to inefficient rent-seeking behavior. The result is too much innovation. An appropriate policy response is to tax the ex post profits of the entrant by exactly the amount of profit lost by the incumbent. This corrects the tendency of the new innovator to over-invest in innovation.

There is a corresponding danger that a firm entering a brand-new market will under-invest because part of its profits will be poached by a potential entrant. To address this problem as well, the tax on the new innovator can be refunded to the incumbent. Transfers among innovators can solve the efficiency problems that arise with multiple firms producing similar inventions.

E. Implications for Innovation and Insurance Policy

Departures from the first- or second-best outcomes occur if: (1) A new innovator is able to "poach" some of the profits of the incumbent; (2) The market for health insurance fails to be competitive; or (3) The market for health insurance is incomplete. The analysis above suggests the policy remedies most appropriate to these various failures. In general, policy ought to focus on intervening in the innovation market to ensure that the private return to innovation does not exceed the social return, and on completing the insurance market or making it competitive.

E.1 Innovation Market Intervention

Intervention in the market for innovation can be called for when the entry of a new innovator reduces the profits of an incumbent. In this case, the new innovator can extract both the additional consumer surplus he creates, *plus* some of the surplus that previously accrued to the original inventor. The latter portion of the private return is socially excessive, since it involves nothing more than a transfer, rather than a real creation of social wealth. As such, Pigovian taxes on ex post profits, equal to the incumbent's decline in profits, can restore first- or second-best efficiency. Note that the Pigovian tax is superior to price regulation, because it does not distort outcomes in the goods market.

E.2 Insurance Market Intervention

Market power or other incompleteness in the insurance market can compromise the efficiency of innovation investments. In the presence of market power, insurers will share some of the rents from innovation and thus depress the returns to research. Moreover, market power or other imperfections can lead to less than full insurance. Uninsured consumers do not have access to the two-part pricing contract afforded by insurance. The innovator will charge them the standard monopoly price for her invention. This leads to under-utilization in that segment of the market, and under-innovation for the entire market.

Correcting either of these static failures in the insurance market can raise the level of innovation and improve the efficiency of utilization. Aggressively promoting anti-trust policy against insurers when they cover patented innovations seems important, as does promoting access to health insurance.

E.3 Methods For Redistribution

The equilibria constructed above involve the first-best levels of innovation and health-care utilization. However, since consumers receive no surplus, it is possible that consumers do not prefer these equilibria to other, less efficient outcomes. Theoretically, it is always possible to solve distributional problems through appropriate transfers. Any such transfer would have to be made from innovators to consumers, but the method of funding such a transfer cannot affect the margins of either the innovator or the consumer.

A feasible way to achieve redistribution along the Pareto frontier is to tax the innovators expected profits. Efficient incentives are achieved when the innovator faces expected profits $d(r)CS - r$. If, instead, the innovator receives $(1 - \tau)[d(r)CS - r]$, the allocation of resources is unchanged. This requires a τ-percent tax on eventual profits, coupled with a τ-percent *subsidy* for research and development expenditure. The proceeds of the tax could then be disbursed to consumers. This could achieve any desired distribution of resources across innovators and consumers, without affecting research effort or the utilization of the new innovation.

E.4 Public Health Insurance

Our analysis suggests that the optimal design for public health insurance involves a co-payment at marginal cost minus the degree of insurance, coupled with a premium payment that equals consumer surplus plus the actuarial cost of insurance. The key implementation problem is determining these quantities accurately.

A workable alternative is to mimic observed private insurance contracts. While this may not guarantee first- or even second-best efficiency, increasing the availability of insurance at competitive prices improves ex ante and ex post welfare. Ex post welfare increases as long as people receive some insurance and utilization rises; this will always be satisfied. Ex ante welfare will increase as long as costs paid to innovators do not exceed consumer surplus. Following observed competitively determined contracts ensures that this condition will be satisfied: in a competitive market, consumers would opt out of insurance contracts that paid innovators an amount greater than consumer surplus, since the associated premium would also exceed the value of the insurance to the consumer.

The Medicare Drug Benefit is an example of a public health insurance scheme that must at least be welfare-improving according to this logic. Medicare solicits competitive bids from insurers for drug insurance. This may not yield the first- or even second-best outcome, because competitively determined prices will reflect market power or other imperfections in the insurance market. Nonetheless, it is certain that the payment made to innovators will increase their profits, but will not lead to excessive returns on their investments. Moreover, the increase in insurance and utilization will also increase static efficiency from the point of view of today's patients.

F. Conclusions

Uncertainty in the demand for health care innovation provides leverage with which to solve the nagging problem of encouraging efficient innovation ex ante while still permitting efficient utilization ex post. An insurance contract can function as a two-part pricing scheme that yields efficient outcomes, where an innovator sells his product at marginal cost, but receives an ex ante payment equal to the full value of expected consumer surplus. This analysis reveals how static failures in insurance markets can lead to dynamic inefficiencies in the market for innovation, and it also points to the importance of maintaining efficient insurance markets. Indeed, patent monopolies in and of themselves are not socially harmful if they coexist with an efficient market for health insurance.

The optimal design of a public health insurance scheme ought to couple co-payments at or below marginal cost (depending on the level of insurance desired), along with premium payments equal to actuarial cost plus the consumer surplus associated with the goods being purchased. Since these quantities are difficult to observe, a practical, welfare-improving alternative is for a public insurance system to follow the reimbursements and premium schedules determined by competitive insurers.

The link between innovation and insurance is crucial to the ex ante and ex post efficiency of health care markets. Many have argued that the presence of moral hazard in health insurance contracts can help offset the incentives for quantity-restriction faced by monopolists. In fact, the relationship between insurance and innovation is much more fundamental and less coincidental. Competitive insurance markets can structurally remove inefficiencies associated with patent monopolies, regardless of the extent of moral hazard present in preferences and information.

References

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

Garber, A. M., C. I. Jones, and P. M. Romer (2005). "Pricing and the Value of Medical Innovations." Stanford University Working Paper. Palo Alto, CA: Stanford University.

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.

Nordhaus, W. (1969). "Invention, Growth and Welfare." *Cambridge, MA: MIT Press*.

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

Philipson, T., and S. Mechoulan (2005). "Intellectual Property & External Consumption Effects:

Generalizations from Pharmaceutical Markets." University of Chicago Department of Economics. University of Chicago Department of Economics.

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