Journal of Health Economics 73 (2020) 102329

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/econbase

Insurance access and demand response: Pricing and welfare implications☆

David Besanko[a], David Dranove[a,*], Craig Garthwaite[b]

[a] Kellogg School of Management, Northwestern University, United States
[b] Kellogg School of Management, Northwestern University and NBER, United States

ARTICLE INFO

Article history:
Received 18 June 2019
Received in revised form 14 April 2020
Accepted 18 April 2020
Available online 23 May 2020

JEL classification:
I11
I13
L13

Keywords:
Pharmaceutical prices
Health insurance access
Liquidity constraints
Complementary monopoly

ABSTRACT

We present a model in which health insurance allows liquidity-constrained patients access to otherwise unaffordable treatments. A monopolist's profit-maximizing price for an insured treatment is greater (for any cost sharing) than it would be if the treatment was not covered. Consumer surplus may also be less. These results are based on a different mechanism than would operate in a standard moral hazard model. Our model also provides an economic rationale for the common claim that pharmaceutical firms set prices that exceed the value their products create. We show this problem is exacerbated when health insurance covers additional monopoly-provided services.

1. Introduction

Standard insurance theory suggests that the primary motive for the purchase of insurance is to smooth consumption across different states of the world. The welfare economics of insurance has traditionally focused on the trade-off between this consumption smoothing benefit and the costs of moral hazard, and this trade-off has been extensively studied in the health insurance literature (Pauly, 1968; Feldstein, 1973; Friedman, 1974; Feldman and Dowd, 1991; Manning and Marquis, 1989; Manning and Marquis, 1996; Newhouse, 1993; Einav et al., 2013).

But relatively unique among insurance products, health insurance provides a second benefit to consumers: the ability to pre-pay for a package of medical services, some of which would be unaffordable in the absence of insurance. That is, health insurance provides access to health care by helping to eliminate liquidity constraints for the purchase of expensive procedures and treatments. Nyman (1999, 2003) provided the first economic analysis of this access benefit. This access benefit of health insurance has not been lost on policymakers—and arguably served as a motivation for recent large expansions of social and private health insurance programs. For example, both the Affordable Care Act (ACA) and the creation of Medicare Part D (which expanded coverage for pharmaceutical products for Medicare beneficiaries), were supported, in part,

☆ We would like to thank Dan Barron, Amitabh Chandra, Niko Matouschek, FionaScott Morton, Matthew Notodiwidgo, and Michael Powell for their helpfulcomments and suggestions. We would also like to acknowledge Matthew Schmidtfor his helpful research assistance in the early stages of this project.

* Corresponding author.
E-mail addresses: d-besanko@kellogg.northwestern.edu (D. Besanko), d-dranove@kellogg.northwestern.edu (D. Dranove), c-garthwaite@kellogg.northwestern.edu (C. Garthwaite).

https://doi.org/10.1016/j.jhealeco.2020.102329
0167-6296/

because they would allow patients to access treatments and services that they otherwise would have been unable to afford. In fact, much of the most recent debate about the future of the ACA's ban on insurers considering the existence of pre-existing conditions or the practice of medically underwriting policies has focused on the access benefits of insurance rather than the more traditional consumption smoothing benefits.[1]

Given the increasing policy importance of the access motive of health insurance, this paper presents a theoretical analysis of the economics of a health insurance market in which this relatively unique feature of health insurance is front and center. To highlight the distinctive role of the access motive in shaping market outcomes we deliberately abstract away from both the consumption smoothing benefit of insurance and the moral hazard from an individual over-utilizing medical services. At first blush, one might imagine that the insights of a theoretical model featuring the access motive would be straightforward: the ability to obtain health insurance would eliminate liquidity constraints and increase consumer surplus for individuals who would be unable to afford expensive health care without it. This, in fact, is the conclusion reached by Nyman (1999) who estimated that in the U.S. in the late 1980s the access benefit of health insurance was about three times as large as the consumption smoothing benefit.

In this paper we show that the access motive for health insurance may not increase consumer surplus—and may even decrease it—if sellers of medical services exercise market power in setting prices. We study a setting in which liquidity constrained individuals have the opportunity to purchase health insurance in a perfectly competitive market. In the baseline version of the model, the insurance plan covers a share of the cost of a specialized treatment sold by a monopolist (e.g., a new pharmaceutical treatment for cancer). To abstract from the consumption smoothing motive, consumers are assumed to be risk neutral. To abstract from moral hazard based on an individual's overconsumption of medical services, consumers are assumed to purchase at most a single unit of the treatment, and the value of the treatment for any consumer is assumed to exceed the treatment's incremental cost.[2]

If the price of the covered was exogenous, as in the Nyman (1999, 2003) model, access to health insurance would increase aggregate consumer surplus. But when we allow for endogenous pricing of the treatment by a monopoly innovator, we show that the monopolist's profit-maximizing price for an insured treatment is greater (for any level of cost sharing in the insurance contract) than it would be if consumers did not have access to insurance. The impact of monopoly pricing could be so strong that consumer surplus might be lower when consumers are not insured than when they have full insurance. Thus, the market power of providers can potentially enable them to convert the entirety of the access benefit of insurance into producer profit.

The prediction that insurance coverage can raise health care prices would seem to be no different than the one that arises in a traditional model of insurance with moral hazard. In fact the mechanism underlying our results is very different from that in a standard moral hazard model. In such a model, insurance reduces the out-of-pocket amount paid by consumers for covered medical services, leading providers who possess market power to raise prices. Importantly, this mechanism for a price increase can be countered through increased cost sharing, i.e., in the standard model a provider's profit-maximizing price goes down as cost sharing goes up. By contrast in our model, we show that there can exist a range over which increased cost sharing can lead the monopolist to *increase* prices—a result that is important for welfare considerations and does not arise from a standard moral hazard model. The key difference is that in a standard moral hazard model, greater cost sharing induces a *substitution effect*, i.e., with greater cost sharing, consumers face a higher out-of-pocket price, inducing them, on the margin, to substitute away from the insured service. In our model, by contrast, cost sharing entirely works through an *income effect*. Cost sharing deters low-income individuals from purchasing insurance, in part because they may not be able to afford the cost share, thus decreasing the value of the insurance product. This raises the average income of those who purchase insurance, which may encourage the monopolist to raise its price, even though these individuals are paying a higher portion of the price themselves. We note that increased cost sharing leading to a decline in insurance purchases is not an abstract concern or construct of our model. During the 2017 debate about "repeal and replace" of the ACA, the Congressional Budget Office estimated that lower income consumers would not purchase high deductible insurance because cost sharing decreased the value of insurance.[3] Similarly, recent work by Geruso et al. (2017) suggests that some insurers attempt to use formulary construction and cost sharing for expensive pharmaceuticals to make their policies unattractive to particularly expensive enrollees.

We then extend our model to a situation in which insurers are required to cover not just the monopoly-provided service but also a set of core medical goods and services sold at competitive prices (e.g. hospital and physician services as well as generic medications).[4] Thus, unlike our baseline model which implicitly assumed that consumers can purchase coverage *a la carte*, in this extension consumers face the choice between an insurance plan that covers a fixed bundle of services or going without coverage. This extension more closely approximates the structure of the health insurance market in the United States. Consistent with the results in our basic model, full insurance results in a higher

[1] After all, if an individual is already diagnosed with a condition that requires a known and expensive treatment and wants to purchase health insurance, the product they are purchasing is not "insurance" in the typical consumption smoothing sense.

[2] Given our focus on high-cost treatments, this is not a very strong assumption.

[3] Congressional Budget Office Cost Estimate, "American Health Care Act," (March 13, 2017), https://www.cbo.gov/sites/default/files/115th-congress-2017-2018/costestimate/americanhealthcareact.pdf (accessed August 20, 2018).

[4] We are not suggesting that all medical services and generic drugs are priced at competitive levels, nor do our qualitative findings materially differ if we allow for supracompetitive pricing.

price for the monopoly service and lower consumer surplus than would arise if the insurance plan did not cover the monopoly service. We also show that the monopolist will charge more than its product is worth to consumers. Our model demonstrates that when insurers are compelled to cover both competitive and monopolized services in a single bundle, a monopoly seller may set a price that captures some of the value created in the competitive markets for other health services, value that was previously captured by consumers. The monopolist stops raising price only when it reaches the point that insurance becomes unaffordable to so many consumers that sales of their product start to fall. Though our model is too stylized to offer precise numerical predictions, an example that broadly fits real world data confirms some of the model's implications. The example illustrates that the profit-maximizing price for the product is far greater in the presence of health insurance than it would be for an uninsured population, and consumer surplus decreases dramatically.

We further show that this problem is exacerbated when the insurance bundle must include several monopolized products, such as pharmaceuticals treating different diseases, which is again a feature of most insurance products in the United States and a requirement of Medicare Part D. This result stems from the Cournot complementary monopolist problem where each innovator sets prices without internalizing its effect on the market. Indeed, if the number of monopolized products is sufficiently large, we show that coverage of those products can not only decrease consumer surplus but also total surplus. Importantly, this loss in total welfare is not due to the key force that would operate in a standard moral hazard model—overconsumption of covered medical services. Nor is it akin to the welfare loss in Nyman's model which is due to liquidity-constrained consumers lacking access to affordable insurance. The welfare loss in our model arises because the increase in the prices of the monopolized services eventually becomes so large that lower valuation consumers no longer purchase insurance and therefore forego access to both the monopolized services and the core services. Importantly, this means that even in a situation where insurers could eliminate moral hazard (a primary focus of the design of current insurance contracts) there are still other channels through which insurance can reduce welfare.

While our model can be used to explain any part of the medical sector where providers have some degree of market power, it seems especially applicable to the pharmaceutical sector—a market that is characterized by many products with high and rising prices. These unprecedented high prices have recently received considerable attention from both policymakers and economists including calls for policies that either directly reduce prices (e.g. explicit price controls) or indirectly do so (e.g. increased bargaining rights for government purchasers). Given the potential negative effects on innovation and access that could come from such policies, it is important to understand the determinants of the market prices and their effect on welfare. Our model may help explain the recent increase in the number of oncology products that are covered by health insurance but have been found to not be cost-effective at current prices (Managan, 2015; Loftus, 2015; Walker, 2015).

The novelty of our contribution is in its combined focused on the access motive from health insurance and the role of medical supplier market power when the purchase of health insurance is driven by the access motive. As noted earlier, Nyman (1999, 2003) importantly provides the first formal model of the welfare implications of the access motive. Nyman's model assumes that prices are exogenous and, in doing so, likely overstates the magnitude of the access benefit accruing to consumers. Other papers in the health economics literature have studied the implications of endogenous price setting by providers of health services that have market power, but they focus on settings that are economically different from ours. For example, Jena and Philipson (2013) show that when a cost-effectiveness measure is based on endogenous prices, a technology or treatment that appears to be more cost-effective than another can actually be less cost-effective when judged on the basis of its social cost. In such cases, policies aimed at raising overall cost-effectiveness could have the counterproductive effect of raising health care spending and increasing the adoption of inefficient treatments. Lakdawalla and Sood (2009) formulate a model in which a monopoly health care provider sells to an insurance firm that, in turn, determines a two-part insurance contract for consumers (premium and cost share). They demonstrate that when different types of consumers can be perfectly sorted into different insurance contracts, health insurance markets can eliminate the deadweight loss from monopoly pricing by the health care provider. When this is not possible, monopoly pricing leads to some consumers remaining uninsured, but it does not result in underconsumption of medical care by consumers. As in these papers, a key assumption in our model is that medical care is priced by a firm with market power. But in contrast to Jena and Philipson (2013), our model focuses on the interaction of medical care pricing and insurance, and unlike Lakdawalla and Sood (2009), health insurance in our model serves as a device to enable liquidity-constrained households to obtain access to medical care as opposed to providing consumption smoothing. As emphasized above, this latter distinction leads to very different implications about the roles of cost sharing and endogenous pricing of medical treatments and is increasingly policy relevant.

The remainder of this paper is organized as follows. Section 2 analyzes the pricing of a pharmaceutical product in the face of consumer liquidity constraints. We contrast cases in which insurance does and does not cover the treatment, while allowing for the possibility that insurance involves cost sharing. We characterize consumers' optimal purchase decisions and the implied demand curve for the treatment; the profit-maximizing price of the treatment; and the resulting levels of consumer and total surplus. Section 3 extends our model to the case in which insurance covers a bundle of medical products and services. It also considers the possibility of multiple innovators. Section 4 discusses the policy implications of our analysis. Section 5 summarizes and concludes. Proofs of all propositions and derivations of some key expressions are in the Appendix.

2. The model

Our analysis focuses on the pricing of an innovative treatment (hereafter the treatment) for a serious but rare illness that is covered by insurance. To build intuition, we consider a simple case in which insurance covers only this treatment. We relax this assumption below.

Speaking about his theory of the demand for health insurance, Nyman (2003, p. 3) points out that "risk preferences are largely extraneous to the new theory." To focus attention on the access motive for health insurance in our model, we eliminate the effect of risk preferences by assuming that consumers are risk neutral. This implies, of course, that the consumption smoothing motive for insurance is absent.

Potential consumers of the treatment differ only in the amount of their liquid wealth $W \in [\underline{W}, \overline{W}]$, where $\underline{W} \geq 0$. The cumulative distribution function for W is $F(W)$ and its corresponding density function is $f(W) = F'(W)$, which we assume is bounded on $\left[\underline{W}, \overline{W}\right]$ and strictly positive on $(\underline{W}, \overline{W})$

A consumer can be in one of two mutually exclusive states, $s = \{0, 1\}$. In state $s = 0$, the consumer is healthy and does not need the treatment. In state $s = 1$, the consumer experiences the illness for which the treatment is required. The probability of state 1 is $\rho_I \in (0, 1)$. Consumer utility is

$$U(x, z, s) = \begin{cases} z & s = 0 \\ V(W)x + z - L & s = 1 \end{cases},$$

where $z \in [0, \infty)$ is the quantity of a numeraire consumption good; $x \in \{0, 1\}$ denotes the decision to purchase the treatment (where $x = 1$ constitutes purchasing the treatment); L is a fixed utility loss if the individual needs medical services, and $V(W) > 0$ is the incremental value of the treatment. The utility losses plays no essential role in our analysis, so we set $L = 0$.

The price P of the treatment is determined monopolistically by the innovator of the treatment whose marginal cost, for simplicity, is assumed to be 0.[5] This innovator's price is endogenous, and its determination is a key focus of our analysis.

Consumers have the opportunity to purchase a health insurance contract with a coinsurance rate $\sigma \in [0, 1]$ and a premium $\phi > 0$. When $\sigma = 0$, the treatment is fully covered. As σ increases, the consumer's out-of-pocket expense σP goes up. No insurance coverage corresponds to $\sigma = 1$. Throughout, we take σ to be exogenous, but below we discuss welfare implications of varying σ. Health insurance is assumed to be sold in a competitive market.

A consumer's choice problem proceeds in two stages. Before learning the state s, a consumer decides whether to purchase insurance, and the insurance market attains equilibrium. Then, depending on the individual's first-stage choice and the realized state, the individual chooses x and z. Fig. 1 summarizes the timing of both consumer choice and market activity.

Throughout the analysis, we make the following assumptions:

Assumption 1. $\varepsilon_{WV}(W) \equiv \frac{V'(W)W}{V(W)} \in (0, 1)$ for all W.

Assumption 2. $V(W) > W$ for all W.

Assumption 3. $\rho_I < \frac{W_V^*}{V(W_V^*)}$, where W_V^* is the solution to $\max\limits_{W \in [\underline{W}, \overline{W}]} V(W)[1 - F(W)]$.

Assumption 4. $\frac{f(W)}{1-F(W)}$ is non-decreasing in W.

Assumption 5. $V(W)[1 - F(W)]$ is strictly log-concave in W.

Consistent with the evidence in Viscusi and Aldy (2003), Assumption 1 says the wealth elasticity of the value of life, $\varepsilon_{WV}(W)$, is between 0 and 1.[6] Assumption 2 may seem less natural: after all, why would an individual value something at more than their own ability to pay for it inasmuch as standard economic theory equates value with the willingness to pay. Following standard theory, we would conclude that someone who has, say, only $1,000 in liquid assets and therefore does not purchase a $2,000 life-saving treatment must value that treatment at less than $2,000. However, this revealed preference logic ignores the reality of liquidity constraints. Assumption 2 is tantamount to saying that if people could borrow against future earnings to save their lives, everyone would do so. We can put this another way: suppose that you encountered a gunman who gave the choice of "your liquid assets or your life." If the value of your life was capped by your liquid assets, then you would necessarily be indifferent between these two choices. Yet you, and surely most everyone, would strongly prefer keeping your life over keeping your liquid assets. Your life must be worth more.

In light of Assumption 2, Assumption 3 puts an upper bound on the likelihood that consumers will need the treatment.[7] This is consistent with the notion that the illness necessitating the treatment is serious but relatively rare.[8] The remaining two assumptions are regularity con-

[5] An implication of this assumption is that the benefit of the treatment exceeds the marginal cost for all consumers experiencing the illness. Thus, there is no moral hazard utilization in our model.

[6] The implication that $V'(W) > 0$ is also in line with Becker's (2007) model of health as human capital. Equation (1) in that article expresses the statistical value of life as a function of wealth and expected survivorship, and it is assumed that the valuation of a given improvement in survivorship—$V(W)$ in our model—is "rising in initial wealth."

[7] We explain the significance of W_V^* below.

[8] This assumption is empirically plausible. To illustrate, the value of a statistical life in the U.S. is roughly $10 million 2017 dollars (Kniesner and Viscusi, 2019). If liquid wealth is taken to equal annual income, the median U.S. income in 2017 was about $61,000. The ratio of latter to the former is about 0.06, or 6,000 per 100,000. It is not difficult to find examples of serious diseases whose incidence falls well below this rate. For example, the incidence of hepatitis B in the mid-2000s was about 2.1 per 100,000 (Kim, 2009).

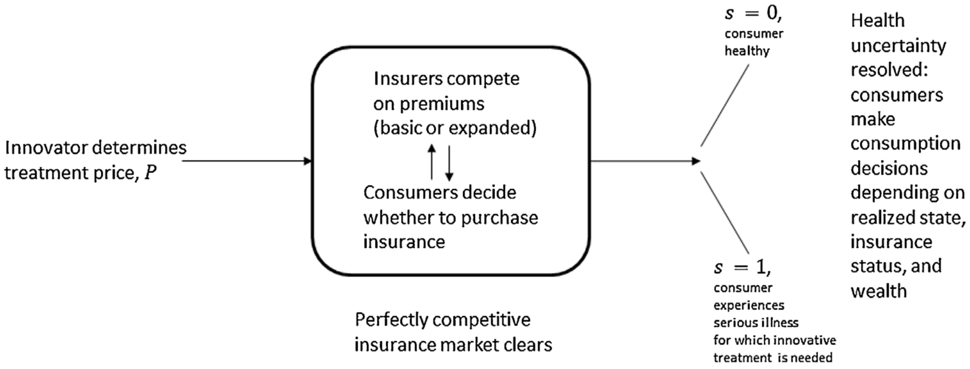

Innovator determines treatment price, *P*
Insurers compete on premiums (basic or expanded)
Consumers decide whether to purchase insurance
Perfectly competitive insurance market clears
s = 0, consumer healthy
s = 1, consumer experiences serious illness for which innovative treatment is needed
Health uncertainty resolved: consumers make consumption decisions depending on realized state, insurance status, and wealth

Fig. 1. The timing of consumer choice and market activity.

ditions: they ensure that second-order conditions for the innovator's profit-maximization hold.[9],[10]

2.1. *The access benefit of health insurance when the treatment price is endogenous*

We begin by deriving the demand curve for the treatment, which we then use to characterize the innovator's profit-maximizing price. We show how this price varies with the coinsurance rate. By comparing the profit-maximizing price at $\sigma = 1$ (no insurance) to profit-maximizing prices when $\sigma < 1$, we identify how providing access to insurance with different degrees of cost sharing affects the innovator's pricing decision. We then, in turn, identify the benefit that consumers receive from having access to health insurance, as well as the total benefit realized by society.

2.1.1. *Market demand curve for the treatment*

In the Appendix, we derive the expected utility of a consumer of type W from not purchasing insurance and the expected utility from purchasing an insurance contract (ϕ, σ):[11]

$$EU^{NI}(W) = \begin{cases} W & W < P \\ W + \rho_I\,[V(W) - P] & W \geq P \end{cases} . \tag{1}$$

$$EU^{I}(W, \phi, \sigma) = \begin{cases} W - \phi & \phi \leq W < \phi + \sigma P \\ W - \phi + \rho_I\,[V(W) - \sigma P] & W \geq \phi + \sigma P \end{cases} . \tag{2}$$

With these expressions we can establish:

Lemma 1. *(a) If a person can afford insurance but would then be left with insufficient wealth to pay the out-of-pocket price σP—i.e.,$W \in [\phi, \phi + \sigma P)$—then that person is strictly better off not purchasing insurance in the first place; (b) If a person purchases insurance and can afford to purchase the treatment at the out-of-pocket price σP—i.e.,$W \geq \phi + \sigma P > P$—then if they become ill, they will purchase the treatment.*

Thus, those who cannot afford the out-of-pocket costs associated with insurance will not purchase insurance in the first place. Those who can afford the out-of-pocket costs will, in fact, use their insurance if they become ill. Because every consumer paying a premium ϕ receives an insurance reimbursement $(1 - \sigma)P$ if they become ill, the perfectly competitive equilibrium premium, $\phi^*(\sigma)$, must equal $\rho_I(1 - \sigma)P$.

Now consider the decision to purchase insurance by those consumers who are the only candidates to do so. Substituting $\phi^*(\sigma) = (1 - \sigma)\rho_I P$ into (2) we get

$$EU^I(W, \phi^*(\sigma), \sigma) = W + \rho_I\,[V(W) - P]\,,$$
$$W \geq [(1 - \sigma)\rho_I + \sigma]\,P. \tag{3}$$

For people who could afford to pay the out-of-pocket cost for the treatment if they are insured and become ill but could not afford it without insurance, i.e., $W \in [(1 - \sigma)\rho_I + \sigma]\,P, P)$,

$$EU^I(W, \phi^*(\sigma), \sigma) - EU^{NI}(W) = \rho_I\,[V(W) - P]\,.$$

These consumers purchase insurance if and only if $V(W) \geq P$. For people who could afford to purchase the treatment if they are uninsured, i.e., $W \geq P$,

$$EU^I(W, \phi^*(\sigma), \sigma) - EU^{NI}(W) = 0.$$

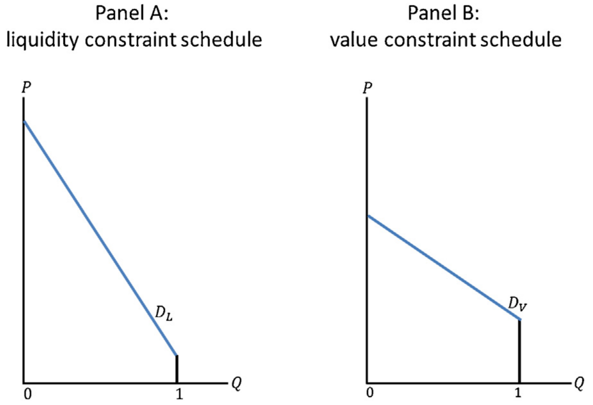
These consumers are indifferent between purchasing and not purchasing insurance, and we assume that they opt to purchase.[12] The foregoing logic implies that the set of peo-

[9] Assumption 4 is a standard regularity condition in models of pricing and auctions that is satififed by many common distributions. Assumption 5 is satisfied if and only the (absolute value) of the price elasticity of demand for the treatment *assuming everyone is fully insured* $(\sigma = 0)$ is locally increasing in P in a neighborhood of the profit-maximizing price. This holds for any demand curve that is strictly concave in price, for example, a linear demand curve, and for any demand curve that is less convex than a constant elasticity demand curve.

[10] In our formulation, the value of the treatment V is a deterministic function of wealth W. A more general model would posit a distribution $F(\xi)$ of consumer types ξ that jointly determine wealth and valuation of life, $W(\xi)$, $V(\xi)$. One can interpret $V(W)$ as the expectation of $V(\xi)$ conditional on $W(\xi) = W$, i.e., $V(W) = \int_{\{\xi | W(\xi)=W\}} V(\xi) dF(\xi)$.

[11] Note that EU^I_i is defined only for wealth levels such that $W \geq \phi$. When $W < \phi$ consumers cannot afford to purchase insurance.

[12] This assumption can be justified on a number of grounds including empirical plausibility (there are few if any examples of wealthy families that go without health insurance even though they could afford all conceivable treatments) and the possibility that insurance provides additional benefits not included in our model (e.g., discounted provider prices). In addition, it is unlikely to be optimal for the producer of a patented pharmaceutical to set a price to serve only a population of high income individuals who have foregone insurance.


Panel A: liquidity constraint schedule

Panel B: value constraint schedule

Fig. 2. The demand curve for the treatment has two parts. The first part, D_L^* in panel A, is the liquidity constraint. The second part, D_V^* in panel B, is the value constraint.

ple that purchase insurance when the treatment price is P is such that

$$W \ge [(1-\sigma)\rho_I+\sigma]P \quad \text{and} \quad W \ge V^{-1}(P), \tag{4}$$

where $V^{-1}(\cdot)$ is the inverse of $V(\cdot)$. Finally, note that individuals who are uninsured—namely $W < [(1-\sigma)\rho_I+\sigma]P < P$—could not afford the treatment if they became ill. This implies that the set of consumers who purchase insurance perfectly overlaps with the set of consumers who purchase the treatment if they become ill. These insights imply that the demand curve for insurance and the demand curve for the treatment coincide.

Proposition 1. *The demand curve for insurance—and thus the demand curve for the treatment—is the measure of W satisfying (4), i.e.,*[13]

$$Q = D(P,\sigma) = 1 - F(\max\{[(1-\sigma)\rho_I+\sigma]P, V^{-1}(P)\}).$$

This demand curve is made up of two parts as depicted in Fig. 2.[14] The first part, labeled D_V in panel A of Fig. 2, represents the demand for insurance if the treatment had infinite value, so that the only factor limiting insurance (and thus treatment demand) is liquidity. We call this the *liquidity constraint*. Along this schedule, consumers purchase insurance (and thus the treatment) only if $W \ge (1-\sigma)\rho_I+\sigma$. The equation of the liquidity constraint is thus $D_L(P,\sigma) = 1 - F([(1-\sigma)\rho_I+\sigma)]P)$. The second part, labeled D_V in panel B of Fig. 2, represents the demand for insurance if there was no liquidity constraint, so that the only limiting factor is the value of the treatment. We call this the *value constraint*. Along this schedule, consumers purchase insurance (and thus the treatment) only if $V(W) \ge P$. The equation of the value constraint is thus $D_V(P) = 1 - F(V^{-1}(P))$.

Fig. 3 depicts the three potential demand curves for the treatment. In panel A, D_L lies everywhere below D_V on the quantity axis between 0 and 1, and the therefore the demand curve coincides with D_L. In this case, for any treatment price P there are fewer people who can afford to purchase insurance and make the copayment if they become ill than there are people who get a non-negative net benefit from the treatment, and thus demand for the treatment is determined by affordability, not value. In panel B, D_L intersects D_V at a quantity $Q_K(\sigma)$ strictly between 0 and 1. Because consumers will purchase the insurance they use to obtain the treatment only if they can both afford insurance and think it is worth it, the demand curve corresponding to the situation in panel B is the lower envelope of D_L and D_V. In panel C, D_V lies everywhere below D_L and therefore the demand curve coincides with D_V. In this case, for any price P, any consumer that derives non-negative net benefit from the treatment can afford to purchase and use insurance, and thus demand for the treatment is determined by value and not affordability.

Where the value and liquidity constraints cross in panel B of Fig. 3, Assumption 1 implies that the liquidity constraint is less price elastic than the value constraint.[15] Because the value of the treatment rises less than proportionately as wealth increases (which, as we noted, is consistent with empirical evidence), for any given increase in price, there will be a larger number of consumers who find the treatment no longer valuable than who find it no longer affordable. (Think of the extreme case in which value of the treatment is a constant independent of wealth; an increase in the price of the treatment above this value reduces the proportion of consumers who find the treatment valuable to zero, while a decrease in price increases that proportion to one. In that case, the value constraint would be perfectly elastic.) We discuss below the economic significance of the differences in the elasticities at the kink in the demand curve.

The role of the coinsurance rate

In general, we cannot rule out any of the cases depicted in Fig. 3, and determination of the profit-maximizing price must take into account each possibility. Fortunately, these cases vary systematically with the coinsurance rate σ. With no insurance ($\sigma = 1$), the liquidity constraint lies below the value constraint for all $Q \in [0,1]$—or equivalently, $D_L(P,1) < D_V(P)$ for all P.[16] With no insurance, the liquidity constraint thus represents the demand curve for the treatment as in panel A of Fig. 3. In other words, with no insurance—or by continuity, with a coinsurance rate sufficiently close to 1—liquidity restricts access to the treatment.

Because $\frac{\partial D_L(P,\sigma)}{\partial \sigma} = -f([(1-\sigma)\rho_I+\sigma)]P)(1-\rho_I) < 0$, while $D_V(P)$ is independent of σ, as the coinsurance rate decreases the liquidity constraint rotates rightward, while the value constraint remains fixed. Eventually, the liquidity constraint intersects the value constraint at $Q \in (0,1)$, giving rise to the situation in panel B of Fig. 3. As the coinsurance rate decreases further, the kink in the demand curve in panel B, point $(Q_K(\sigma), P_K(\sigma))$, slides to the

[13] Strictly speaking, the expected quantity of the treatment purchased is $\rho_I D(P,\sigma)$. To simplify verbiage, throughout we refer to $D(P)$ as both the demand curve for insurance and the demand curve for the treatment.

[14] These figures depict the case in which W has a uniform distribution, and $V(W)$ is linear in W.

[15] See the Appendix for the proof.

[16] See the Appendix for the proof.

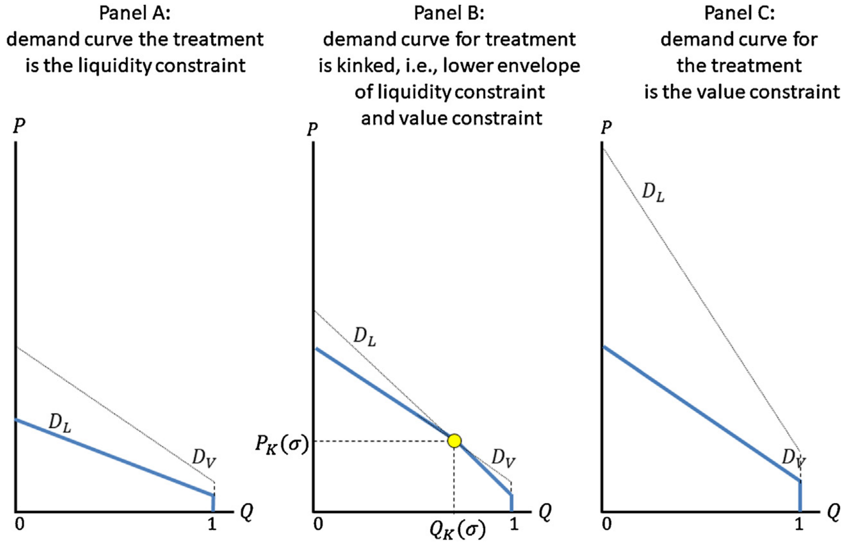


Fig. 3. The demand curve for the treatment could be D_L^*, D_V^*, or a kinked demand curve given by the lower envelope of D_L^* and D_V^*.

southeast—i.e., $\frac{dP_K(\sigma)}{d\sigma} > 0$ and $\frac{dQ_K(\sigma)}{d\sigma} < 0$.[17] For sufficiently small σ, the liquidity constraint *may* be entirely above the value constraint as in panel C of Fig. 3, though it need not.[18] Still, as we show below, when $\sigma = 0$, the kinked demand curve may include a large enough portion of the value constraint that the innovator's profit-maximizing price ends up along the value constraint.

This discussion has two notable implications. First, if the treatment price is *exogenous*, then whenever consumer demand is constrained by liquidity (panels A or B), providing access to insurance increases consumer surplus. This confirms the key insight of Nyman (1999, 2003) about the consumer surplus generated by the access benefit when prices are fixed. We can see the Nyman access benefit in our model because, as just noted, a decrease in cost sharing from the no-insurance level $\sigma = 1$ shifts the liquidity constraint rightward. Thus, for any exogenous treatment price P, more generous insurance expands the set of consumers who can obtain the treatment. The second implication pertains to the innovator's profit-maximization problem formally analyzed in the next section: we show that a *small* increase in the coinsurance rate above 0 has *no impact* on the profit-maximizing price of the treatment. This seems counterintuitive. After all, an increase in σ increases the out-of-pocket price paid by consumers. Why would that not affect the demand curve for the treatment and thus the profit-maximizing price? The formal answer is that, as just discussed, for a small positive coinsurance rate, the profit-maximizing price of the treatment is determined by the value constraint, which does not depend on σ. Thus, the profit-maximizing price is the same as it would be with full insurance. But that begs a question: why would demand for the treatment *ever* be *independent* of the coinsurance rate, even when that rate is small? The answer is that in our model insurance is a pre-paid plan that enables a consumer to purchase the treatment at the coinsurance rate. This makes the treatment affordable to consumers who could not afford the full price. As we showed, consumers *only* obtain the treatment through this pre-paid plan, i.e., uninsured consumers do not obtain the treatment. The plan has an up-front premium of $(1-\sigma)\rho_I P$, an expected copayment of $\rho_I \sigma P$, and an expected benefit of $\rho_I V(W)$. *If a consumer is not liquidity constrained*—i.e., the consumer can afford the insurance premium and the copayment—the consumer perceives the net benefit of the pre-paid plan, and thus the treatment, to be $\rho_I V(W) - \rho_I \sigma P - (1-\sigma)\rho_I P = \rho_I (V(W) - P)$. This means that demand for the treatment depends on the full price P, not the out-of-pocket price σP.[19] Put simply, insurance enables consumers to purchase the treatment that was unaffordable without insurance. If consumers can afford the insurance premium, they will purchase the treatment, regardless of the coinsurance rate.

If a consumer *is liquidity constrained*, the coinsurance rate does matter because it affects the affordability of insurance and therefore affects the innovator's price.

2.1.2. *Profit-maximizing price of the treatment*

The innovator's profit-maximization problem is

$$\max_P \rho_I D(P, \sigma) P.$$

[17] See the Appendix for the proof.

[18] In the Appendix, we provide a sufficient condition for the value constraint to lie entirely below the liquiditiy constraint when there is full insurance.

[19] We can make the point about the "irrelevance" of the coinsurance rate in a different way. Suppose that demand was not liquidity constrained. Now suppose that there is an unexpected increase in the coinsurance rate from σ_1 to $\sigma_2 < 1$. Wouldn't some insured consumers refrain from purchasing the treatment? The answer is no. Since the liquidity constraint was not binding, any consumer who purchases insurance is such that $V(W) - P \geq 0$. For any such consumer, $V(W) - \sigma_2 P > 0$, so they will all purchase the treatment at the higher coinsurance rate, just as they would at the lower rate. That is, once non-liquidity constrained consumers purchase the pre-paid plan, they will use it whatever the coinsurance rate is.

In the Appendix we show that the profit-maximizing price $P^*(\sigma)$ of the treatment, the corresponding quantity $Q^*(\sigma)$ of those who become insured (and thus who purchase the treatment if they become ill) depend on the coinsurance rate σ and satisfy:

$$P^*(\sigma) \;=\; \begin{cases} V(W_V^*) & \sigma \in [0, \sigma_a] \\ P_K(\sigma) & \sigma \in [\sigma_a, \sigma_b] \\ \dfrac{W_L^*}{(1-\sigma)\rho_I + \sigma} & \sigma \in [\sigma_b, 1] \end{cases} .$$

$$Q^*(\sigma) \;=\; \begin{cases} 1 - F(W_V^*) & \sigma \in [0, \sigma_a] \\ Q_K(\sigma) & \sigma \in [\sigma_a, \sigma_b] \\ 1 - F(W_L^*) & \sigma \in [\sigma_b, 1] \end{cases} .$$

where:

1. W_V^* solves $\max\limits_{W \in [\underline{W}, \overline{W}]} V(W)(1-F(W))$;[20]
2. W_L^* solves $\max\limits_{W \in [\underline{W}, \overline{W}]} \frac{W}{(1-\sigma)\rho_I + \sigma}(1-F(W))$;
3. If $\underline{W} f(\underline{W}) < 1$, then $W_L^* > W_V^*$.[21] If $\underline{W} f(\underline{W}) \geq 1$, then $W_L^* = W_V^* = \underline{W}$.
4. σ_a and σ_b are given by

$$\sigma_a = \frac{\dfrac{W_V^*}{V(W_V^*)} - \rho_I}{1 - \rho_I} > 0.$$

$$\sigma_b = \frac{\dfrac{W_L^*}{V(W_L^*)} - \rho_I}{1 - \rho_I} < 1.$$

If $\underline{W} f(\underline{W}) < 1$, then $0 < \sigma_a < \sigma_b < 1$; if $\underline{W} f(\underline{W}) \geq 1$, $) < \sigma_a = \frac{\frac{\underline{W}}{V(\underline{W})} - \rho_I}{1-\rho_I} = \sigma_b < 1$.

The term W_V^* is the wealth of the marginal consumer when the innovator's optimal solution occurs along the value constraint (but not the liquidity constraint), and W_L^* is the wealth of the marginal consumer when the innovator's optimal solution occurs along the liquidity constraint (but not the value constraint). Let us first consider the case in which $\underline{W} f(\underline{W}) < 1$. In this case, $W_L^* > W_V^*$, and Fig. 4 illustrates the solution to the innovator's profit-maximization problem. When the coinsurance rate is sufficiently high—$\sigma \in [\sigma_b, 1]$—consumers are liquidity constrained, and the profit-maximizing quantity and price are determined by the liquidity constraint. The number of people purchasing insurance and obtaining the treatment remains unchanged at $1 - F(W_L^*)$ as σ increases above σ_b. This may seem surprising: why wouldn't less wealthy consumers continue to drop out as the coinsurance rate increases? This would, in fact, be the case if the innovator kept its price the same, but as Fig. 4 shows, the innovator profit maximizing price falls to compensate for a higher σ. The reason the profit-maximizing quantity does not change is because the price elasticity of demand along the liquidity constraint at any fixed *quantity* Q remains unchanged as σ varies.[22]

By contrast, when the coinsurance rate is sufficiently low—$\sigma \in [0, \sigma_a]$—insurance coverage is generous, and the profit-maximizing quantity and price are determined by the value constraint. The wealth of the marginal consumer in this case is W_V^*. Because $W_L^* > W_V^*$, more generous insurance expands the set of people who have access to the treatment. This tells us that the Nyman access motive is at work even with an endogenous treatment price.

For intermediate coinsurance rates, $\sigma \in (\sigma_a, \sigma_b)$, the profit-maximizing quantity and price occur at the kink in the demand curve. Because $\frac{dP_K(\sigma)}{d\sigma} > 0$ and $\frac{dQ_K(\sigma)}{d\sigma} < 0$, as coverage becomes more generous (σ decreases), more consumers purchase insurance and have access to the treatment on increasingly generous terms. Interestingly, as Fig. 4 shows, the price of the treatment falls as insurance coverage becomes more generous over this range. We offer an intuition for this surprising result in Section 2.1.4.

Finally, as the graph in the upper-right portion of Fig. 4 illustrates, if $V(W_V^*) > W_L^*$ the profit-maximizing price of the treatment when consumers have access to insurance at any coinsurance rate $\sigma \in [0, 1)$ exceeds the profit-maximizing price when consumers have no insurance ($\sigma = 1$). Thus, although decreasing σ gives rise to the Nyman access motive, with consumers having more access to more generous insurance, they could still end up paying a higher price for the treatment. However, this need not happen. As shown in the lower-right portion of Fig. 4, if $V(W_V^*) < W_L^*$, the price of the treatment when consumers have access to sufficiently generous insurance will be less than the price with no insurance. We discuss the significance of these results on the treatment price below.

Now consider the case in which $\underline{W} f(\underline{W}) \geq 1$. Here the innovator's problem entails a corner solution in which $W_L^* = W_V^* = \underline{W}$, and thus $Q^*(\sigma) = 1$ for all $\sigma = [0, 1]$. In this case there is no access benefit from insurance because the innovator serves the entire market even if consumers are uninsured. In contrast to Fig. 4, there is no intermediate range of σ at which both the liquidity and value constraints bind. Instead, for σ equal to or close to 1 (i.e., no

[20] In what follows, we focus on the case in which $W_V^* > \underline{W}$, so that at the profit-maximizing price of the treatment, some consumers do not purchase insurance (and thus the treatment). It is straightforward to show that a sufficient condition for this is $\frac{V'(\underline{W})}{V(\underline{W})} \geq f(\underline{W})$. This would hold, for example, if the density function $f(\cdot)$ goes to zero as $W \to \underline{W}$. It is not essential that we rule out a corner solution in which the entire market is served. All the insights in this section hold if $Q^*(\sigma) = 1$ for σ in a neighborhold of 0. Indeed, in the unform distribution example in the next section, a corner solution of this sort arises.

[21] We provide an economic interpretation of the term $\underline{W} f(\underline{W})$ below.

[22] We can see this by considering the inverse of the liquidity constraint:

$$P_L(Q) = \frac{F^{-1}(1-Q)}{(1-\sigma)\rho_I + \sigma}.$$

It is straightforward to verify that the inverse elasticity $\frac{dP_L(Q)}{dQ} \frac{Q}{P_L(Q)}$ is independent of σ.

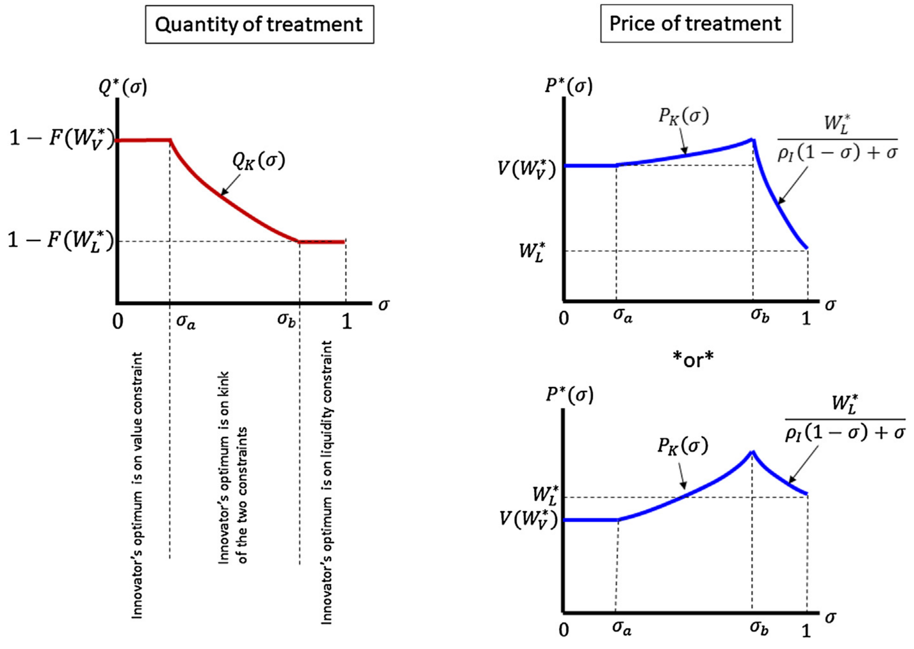


Fig. 4. Innovator's profit-maximizing quantity and price as a function of the coinsurance rate: $\underline{\hat{W}} \leq 1$.

insurance or insurance has a high coinsurance rate), the solution to the innovator's profit-maximization problem occurs along the liquidity constraint, in which case the price of the treatment is $\frac{\underline{W}}{(1-\sigma)\rho_I+\sigma}$. As σ decreases this price goes up. Eventually, the point is reached ($\sigma = \frac{\frac{\underline{W}}{V(\underline{W})}-\rho_I}{1-\rho_I}$) at which the solution to the innovator's problem occurs along the value constraint—though still where $Q^* = 1$. At this point the price of the treatment becomes $V(\underline{W})$ and remains at this level as σ falls to 0. Because $P^*(0) = V(\underline{W}) > \underline{W} = P^*(1)$, the price of the treatment is higher under full insurance than it is no insurance. Overall, then, when $\underline{W}f(\underline{W}) \geq 1$, there is no access benefit from insurance, but there is an adverse price effect from the exercise of monopoly power by the innovator.

The term $\underline{W}f(\underline{W})$ corresponds to the (absolute value) of the price elasticity along the liquidity constraint at the point at which the entire market is served. To the extent that the lower bound of the wealth distribution $\underline{W} \approx 0$, in order for $\underline{W}f(\underline{W}) \geq 1$ to hold, the wealth distribution would have to have an extraordinarily "fat" lower tail, i.e., virtually the entire population would need to be clustered at the very bottom of the wealth distribution. Because this does not describe conditions in the U.S., we think it best to regard $\underline{W}f(\underline{W}) < 1$ rather than $\underline{W}f(\underline{W}) \geq 1$ as the normal case. However, wealth distributions in very poor countries may have very fat lower tails, so the case in which $\underline{W}f(\underline{W}) \geq 1$ is not completely far fetched.

2.1.3. Implications for total and consumer surplus

Recalling that consumers who do not purchase insurance also do not purchase the treatment if they become ill, while all those who do purchase insurance also obtain the treatment if they become ill, consumer surplus can be expressed as

$$CS(P,\sigma) = \int_{\underline{W}}^{W(P,\sigma)} tf(t)dt + \int_{W(P,\sigma)}^{\overline{W}} \left\{t + \rho_I\left[V(t) - P\right]\right\} f(t)dt$$
$$= W_M + \rho_I \int_{W(P,\sigma)}^{\overline{W}} V(t)dt - \rho_I PD(P,\sigma) \quad .$$

where W_M is the mean of the wealth distribution, and $W(P,\sigma) = \max\{[(1-\sigma)\rho_I + \sigma]P, V^{-1}(P)\}$ is the wealth of the marginal consumer when the price is P and the coinsurance rate is σ. Because the innovator's expected profit is $\rho_I D(P,\sigma)P$, total surplus is

$$TS(P,\sigma) = W_M + \rho_I \int_{W(P,\sigma)}^{\overline{W}} V(t)dt.$$

Finally, let $CS^*(\sigma) = CS(P^*(\sigma),\sigma)$ and $TS^*(\sigma) = CS(P^*(\sigma),\sigma)$ equal the levels of consumer and total surplus induced by the innovator's profit-maximizing price when the coinsurance rate is σ.

We first show that full insurance maximizes total surplus.

Proposition 2. *Total surplus attains its maximum value when consumers have access to full insurance, i.e., $TS^*(0) = \max_{\sigma \in [0,1]} TS^*(\sigma)$. If $\underline{W}f(\underline{W}) < 1$, so that $\underline{W} \leq W_V^* < W_L^*$, total surplus is strictly higher when consumers have full insurance than when they have no insurance or insurance with sufficiently high cost sharing, $TS^*(0) > TS^*(\sigma)$ for $\sigma \in (\sigma_a, 1]$.*

Proposition 2 implies that taking into account the impact of insurance on price setting, in those circumstances in which insurance expands access to the treatment ($\underline{W}f(\underline{W}) < 1$), this access is socially beneficial. However, this access benefit does not necessarily flow to consumers: consumer surplus may not go up. To illustrate this, we establish:

Lemma 2. *(a) If $\underline{W}f(\underline{W}) \geq 1$, so that $\underline{W} = W_V^* = W_L^*$, $CS^*(\sigma)$ attains its maximum value when there is no insurance, i.e., $\sigma = 1$; (b) If $\underline{W}f(\underline{W}) < 1$, so that $\underline{W} \leq W_V^* < W_L^*$, $CS^*(\sigma)$ attains its maximum value at either full insurance ($\sigma = 0$) or no insurance ($\sigma = 1$).*

The result in (a) flows directly from our discussion above: when $\underline{W}f(\underline{W}) \geq 1$ and the innovator prices the treatment to serve the entire market even with no insurance, there is no access benefit from insurance, but an adverse price effect. Consumer surplus is thus unambiguously higher when there is no insurance. Part (b) of Lemma 2 arises because while the access benefit from insurance becomes larger as we move from no insurance to full insurance (reflected in Proposition 2), the effect of price on consumer welfare is ambiguous because the innovator's profit-maximizing price is non-monotonic in σ, as can be seen in Fig. 4.

However, when $\underline{W}f(\underline{W}) < 1$ the sign of $CS^*(0) - CS^*(1)$ is ambiguous. This is because it embodies the offsetting effects of better access under full insurance but a potentially higher price of the treatment. We can see this from the expression for $CS^*(0) - CS^*(1)$:

$$CS^*(0) - CS^*(1) = \int_{W_V^*}^{W_L^*} \left[V(t) - V(W_V^*)\right] dt - [V(W_V^*) - W_L^*][1 - F(W_L^*)]. \tag{5}$$

The first term in (5), which is unambiguously positive, is the benefit to consumers from increased access under full insurance. The second term, which could be positive or negative, is the effect on consumer welfare from the difference in the innovator optimal prices under full insurance and no insurance. If the price of the treatment when there is full insurance is at least as low as it is when there is no insurance, as in the lower right panel of Fig. 4, moving from no insurance to full insurance unambiguously *increases* consumer surplus: the lower price of the treatment reinforces the access benefit. If, on the other hand, the price of the treatment when there is full insurance is higher than it is with no insurance, then consumer surplus might still be higher with full insurance, as illustrated in Fig. 5, but it also could be negative, as illustrated in Fig. 6. (In each figure, area C corresponds to the first term in (5), while are E corresponds to the second term.) In Fig. 5 the price of the treatment is only slightly higher when there is full insurance, but in this case the welfare gain from the increase in access due to insurance is large enough to more than offset the greater expenditures required of consumers. By contrast in Fig. 6, the increase in access from insurance is more modest, and the price of the treatment under insurance is elevated more than it is in Fig. 5, so the consumer welfare gain from increased access is not large enough to offset the increase in the price of the treatment.

Whether the access benefit dominates the price effect (in those cases in which the price effect is disadvantageous to the consumer) depends on the price elasticity along the value constraint. The situation in Fig. 6 involves less price elasticity along the value constraint (at any price) than that in Fig. 5. It is straightforward to show that

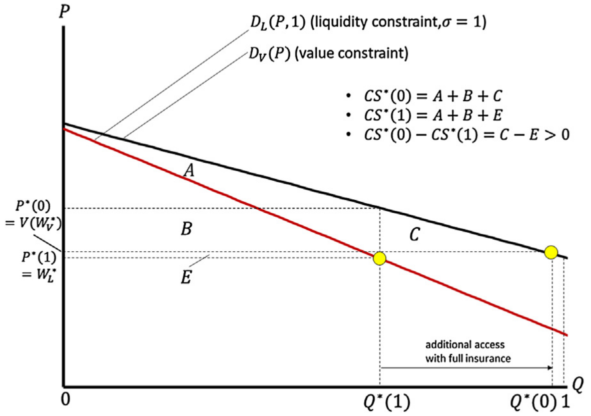


Fig. 5. $CS^*(0) - CS^*(1)$ = area C - area E, which in this case is positive. Thus, consumer welfare is higher with full insurance than with no insurance.

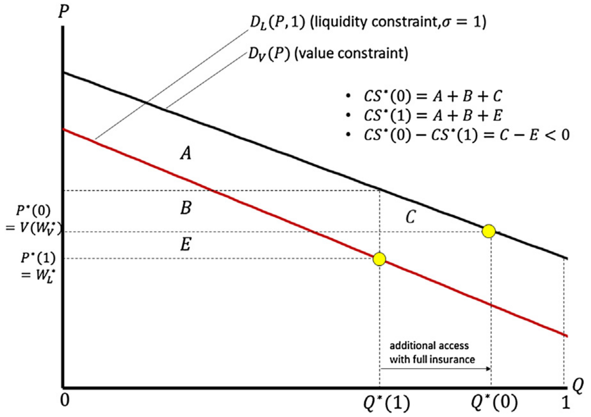


Fig. 6. $CS^*(0) - CS^*(1)$ = area C - area E, which in this case is negative. Thus, consumer welfare is lower with full insurance than with no insurance.

$$\left|\frac{dD_V(P)}{dP}\frac{P}{D_V(P)}\right| = \frac{\frac{f(V^{-1}(P))V^{-1}(P)}{1-F(V^{-1}(P))}}{\varepsilon_{VW}(V^{-1}(P))},$$

so a less price elastic value constraint is associated with a valuation function $V(W)$ that is more elastic in wealth. We can formalize this in the following way. Consider the case in which $\underline{W}f(\underline{W}) < 1$, so an access benefit from insurance can arise. Let $\theta \in (\underline{\theta}, \overline{\theta})$ be a parameter that for any W increases the wealth elasticity of the value of the treatment, i.e.,

$$\frac{\partial \varepsilon_{VW}(W, \theta)}{\partial \theta} > 0 \tag{6}$$

Suppose, further that for all $W \in [\underline{W}, \underline{W}]$, $\lim_{\theta \to \overline{\theta}} \varepsilon_{VW}(W, \theta) = 1$.[23] Then we can establish

Proposition 3. *There exists a critical value of the elasticity parameter $\hat{\theta} \in (\underline{\theta}, \overline{\theta})$ such that if $\theta \geq \hat{\theta}$, the treatment price is*

[23] An example of a specification that satisfies (6) and $\lim_{\theta \to \overline{\theta}} \varepsilon_{VW}(W, \theta) = 1$,

higher under full insurance than under no insurance and consumer surplus is lower, i.e., $P^*(0,\theta) > P^*(1)$ and $CS^*(0,\theta) < CS^*(1)$.

Table 1 summarizes our results on consumer surplus in terms of properties of the liquidity and value constraints.

2.1.4. Isn't this just standard moral hazard?

In a traditional model of health insurance with moral hazard, insurance reduces the out-of-pocket amount paid by consumers for covered medical services, leading providers of those services who possess market power to raise price. Because of this, providing access to medical services may actually decrease consumer surplus. One might ask whether the possibility that insurance could lead to an increase in the price of the treatment simply reflects this well-known intuition.

The answer is "no." The mechanism in our model differs from that of the standard moral hazard model, and the two models give different predictions about cost sharing and price.

The standard intuition in a traditional moral hazard model suggests that reduction in access from well-designed cost sharing would increase consumer surplus by decreasing the overutilization of medical services and inducing providers of insured services to lower their price. But in our model, that is not necessarily the case. Indeed, given the Nyman access motive that operates in our model, increased cost sharing could simultaneously limit access and *increase* the price of medical services.

To see this, recall the intuition about how cost sharing affects pricing when cost sharing is low. When $\sigma = 0$, the demand curve for the treatment is the value constraint. For sufficiently small increases in cost sharing, all consumers who think insurance is worth it can afford the premium and cost share. The demand curve thus continues to be the value constraint, and the optimal price is unchanged. But as cost sharing increases further, the liquidity constraint begins to come into play. As the coinsurance rate rises, it eventually reaches the level σ_a at which the profit-maximizing quantity and price along the value constraint $(1 - F(W_V^*), V(W_V^*))$ coincide with the point $(Q_K(\sigma_a), P_K(\sigma_a))$ at which at which the value constraint intersects the liquidity constraint for coinsurance rate σ_a in Fig. 7. For a range of coinsurance rates above σ_a, the profit-maximizing price and quantity occur at the kink in the demand curve. In such cases, the treatment is priced so that the consumer who finds the treatment to be just worth the price is just able to afford insurance. Fig. 7 illustrates why it makes sense for the innovator to price the treatment this way. When the coinsurance rate is just slightly greater than σ_a, say $\hat{\sigma}$, the liquidity constraint pivots slightly to the southwest, and the demand curve is now *EKF*. Because the higher coinsurance rate reduces affordability, it places a more onerous constraint on the innovator's pricing decision. If the innovator wanted to continue to sell the same amount, it would have to lower its price to $P' < V(W_V^*)$. But as discussed above, the liquidity constraint is less price elastic than the value constraint for σ close to σ_a, so this is not an optimal response to the slightly higher coinsurance rate. A more profitable response is to move up to the kink $(Q_K(\hat{\sigma}), P_K(\hat{\sigma}))$ corresponding to the slightly higher coinsurance rate, resulting in $H - G$ more profit than it gets by lowering price to P'.

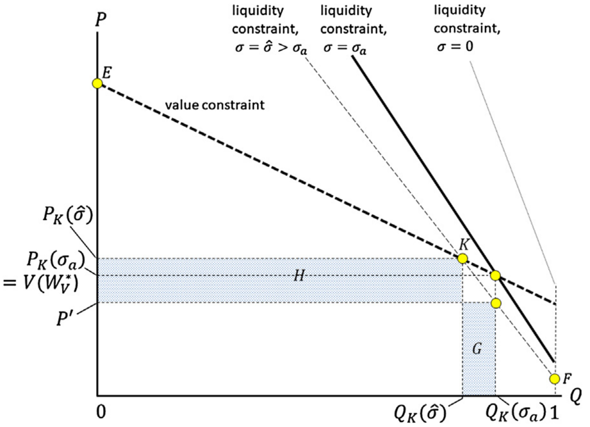

P
liquidity constraint, $\sigma = \hat{\sigma} > \sigma_a$
liquidity constraint, $\sigma = \sigma_a$
liquidity constraint, $\sigma = 0$
E
value constraint
$P_K(\hat{\sigma})$
$P_K(\sigma_a)$
$= V(W_V^*)$
P'
K
H
G
F
0
$Q_K(\hat{\sigma})$
$Q_K(\sigma_a)$
1
Q

Fig. 7. How increased cost sharing could simultaneously limit access and increase the price of medical services.

This is very different from the traditional moral hazard intuition where increases in the cost sharing rate induce a more price elastic demand curve and a lower price of an insured medical service. Here, increases in the cost sharing rate (over a certain range) can induce a *less price elastic* demand curve and a *higher price* of an insured service. In a nutshell, a greater cost sharing rate brings into play liquidity as a factor in pricing as some individuals find they cannot afford the out-of-pocket expenses of the treatment. The innovator exploits this by setting a higher price than it would at a lower coinsurance rate. We note that this is not an abstract concern. Recent insurance products, particularly those on the ACA marketplaces, have meaningful cost sharing in the form of large deductibles and high coinsurance rates. Even with existing limits on out-of-pocket spending for families, many patients likely find themselves unable to afford expensive treatments, and some may find that the insurance product is no longer attractive.

Now, eventually, as the coinsurance rate increases even further, something analogous to the traditional intuition takes over. As Fig. 4 shows, for $\sigma > \sigma_b$ the innovator's optimum moves off the kink, and its profit maximizing price equals the most profitable price along the liquidity constraint alone, $\frac{W_L^*}{(1-\sigma)\rho_I+\sigma}$. As σ increases over this range, the liquidity constraint pivots inward, acting as a tighter constraint on what the innovator can charge for the treatment.

Even still, as shown in the lower right-hand panel of Fig. 4, it is possible that the price with no cost sharing could actually be lower than it is when the coinsurance rate is very high. This outcome would not be an implication of a conventional moral hazard model.

as well as our maintained assumptions on $V(W,\theta)$ is

$$V(W,\theta) = AW^{\theta},$$

where $\theta \in (0,1)$, and $A > \overline{W}^{1-\theta}$.

Table 1
Properties of liquidity and value constraints.

Properties of liquidity and value constraints	Access benefit?	Price effect?	Impact on consumers?
· Liquidity constraint has inelastic and elastic regions and value constraint is relatively more price elastic	Yes	Ambiguous	Ambiguous
· Liquidity constraint has inelastic and elastic regions and value constraint is relatively less price elastic	Yes	Adverse	Negative
· Liquidity constraint only has an elastic region	No	Adverse	Negative

To sum up, in a standard model of insurance, greater cost sharing increases the price elasticity of demand of an insured medical service. In our model, greater cost sharing could decrease the price elasticity of demand, though it must eventually increase it as the coinsurance rate approaches 100 percent. This difference arises because in the standard model, greater cost sharing induces a *substitution effect*, i.e., with greater cost sharing, consumers face a higher out-of-pocket price, inducing them, on the margin, to substitute away from the insured service. In our model, by contrast, cost sharing works entirely through an *income effect*. Higher cost sharing deters low-income individuals from purchasing insurance, in part because they may not be able to afford the cost share. This raises the average income of those who purchase insurance, which may encourage the monopolist to raise its price.

3. Health insurance as a bundle

In the analysis above, insurance covered only the treatment. This would be appropriate in settings in which consumers can purchase insurance à la carte. But in practice, health insurance usually covers a fixed bundle of services, and consumers cannot choose which to purchase insurance for and which not to. In some settings such as Medicare Part D and the ACA marketplaces, minimum coverage standards create mandatory bundles of services that define a legally acceptable insurance contract. In this section, we adapt the model to deal with the case in which consumer choice is limited to purchasing insurance that covers a range of services or no insurance at all.

We do so by extending the model to include a third state of the world, in which the consumer would also benefit from a package of well-established "core" medical services. Thus a consumer can be in one of three mutually exclusive states, $s = \{0, 1, 2\}$, where states 0 and 1 are as above, and in $s = 2$ the consumer experiences a medical condition for which the core medical services are required. The probability of state 1 continues to be ρ_I, and the probability of state 2 is $\rho_C \in (0, 1)$, where $0 < \rho_C + \rho_I < 1$. Consumer utility is

$$U(x, y, z, s) = \begin{cases} z & s = 0 \\ V(W)x + z & s = 1 \\ By + z & s = 2 \end{cases},$$

where x and z are as they were above; $y \in \{0, 1\}$ denotes the decision to purchase the package of core medical services (where $y_i = 1$ constitutes buying the package), and B is the value of core medical services (for simplicity, assumed to be the same for all consumers).[24] In contrast to the treatment, the price of core medical services is assumed to be determined competitively and equals $C \in [0, B)$.[25]

3.1. Pricing the treatment under expanded insurance

We now assume that consumers can choose to purchase an expanded insurance policy that covers both the treatment and core medical services or to go without insurance. Having made the point that the mechanism underlying the innovator's pricing behavior in our model differs from standard moral hazard intuition, we simplify the analysis by dropping cost sharing from our model and consider the case in which the expanded health insurance policy provides full coverage. We simplify things further by considering a tractable example in which W has a uniform distribution on $[\underline{W}, \underline{W} + 1]$, and thus

$$F(W) = \max\left\{\min\left\{W - \underline{W}, 0\right\}, 1\right\}.$$

Further, we analyze a linear value-of-life function

$$V(W) = \underline{V} + \alpha(W - \underline{W}),$$

where $\alpha \in (0, 1)$, $\underline{V} > \underline{W}$, $\underline{V} + \alpha > \underline{W} + 1$, and $\rho_I < \frac{\underline{W}}{\underline{V}}$. This specification satisfies Assumptions 1-5.[26]

To derive the demand curve for the treatment, we make four additional parameter assumptions:

Assumption 6. $\underline{W} + 1 < C$.

Assumption 7. $\underline{W} \geq \frac{\rho_C C}{1 - \rho_I}$.

Assumption 8. $\rho_I(\underline{V} + \alpha) < \rho_C(B - C)$.

Assumption 9. $\underline{V} > \alpha$.

Assumption 6 implies that no consumer could afford to pay out of pocket for core medical services. It is plausible in a setting characterized by significant medical cost inflation, as in the U.S.[27] Assumption 7 places a lower

[24] It would not be difficult to extend our model so that B depended on wealth. However, our major insights and intuitions would not change if we did so, and making B independent of wealth simplifies the exposition of the model.

[25] This assumption is made both for analytical tractability and to reflect that in practice basic medical services tend to be priced more competitively than patented pharmaceutical products. If we had assumed that the core services were priced non-competitively, we would have an analysis that would be analogous to that which we present in Section 3.2.

[26] Given Assumption 9 below that $\underline{V} > \alpha$, the problem $\max_{W \in [\underline{W}, \underline{W}+1]} V(W)[1 - F(W)]$ has a corner solution at $\underline{W}$, so $\frac{W_V^*}{V(W_V^*)} = \frac{\underline{W}}{\underline{V}}$ and thus Assumption 3 holds.

[27] Here and elsewhere, we implicitly ingnore the wealthiest tail of the wealth distribution. This would be problematic for us if the innovator

bound on consumer wealth and in particular implies that $\underline{W} > \rho_C C$, i.e., even a consumer with the lowest wealth could afford an actuarially fair “basic” health insurance policy covering only core services. Together, Assumptions 6 and 7 can be shown to give rise to a Nyman access benefit for a basic health insurance plan: without insurance consumers would be unable to afford core medical services, but all would purchase a basic plan covering only those services if expanded insurance was unavailable. Assumption 8 implies that the expected social value of the treatment is less than the expected social value of core medical services. It is plausible in light of the notion that the illness covered by the treatment is rare, while the core medical services cover an array of established treatments that have sufficiently high value when needed. Assumption 9 will help bring into sharp focus an important consequence of having multiple innovators in the section below. Dropping it would not materially change our results or the economic intuition underlying them. Below we present numerical calculations that illustrate that Assumptions 6-9 can be satisfied by empirically plausible parameter values.

With no cost sharing, everyone who purchases insurance will subsequently obtain the treatment and core medical services in those states in which they are needed. Thus, the equilibrium price of the expanded insurance policy when the price of the treatment is P is $\rho_C C + \rho_I P$. The set of people who purchase the treatment when the treatment price is P can be shown to be[28]

$$W \geq \rho_C C + \rho_I P \quad \text{and} \quad W \geq V^{-1}\left(P - \frac{\rho_C}{\rho_I}(B - C)\right) \tag{7}$$

$$\text{and} \quad W < \rho_C C + \rho_I P \quad \text{and} \quad W \geq P \quad . \tag{8}$$

The consumers represented by the inequalities in (8) are those who cannot afford to purchase expanded insurance but who could purchase the treatment if need. In the Appendix we show that Assumption 7 rules out the existence of such consumers. Thus, the only consumers who purchase the treatment are those represented by the inequalities in (7). These are the consumers who purchase expanded insurance because it is affordable ($W \geq \rho_C C + \rho_I P$) and the insurance bundle it provides is valuable ($\rho_I\left[V(W) - P\right] + \rho_C(B - C) \geq 0$). The demand curve for expanded insurance is thus[29]

$$Q = D(P) = 1 - F(\max\{\rho_C C + \rho_I P, V^{-1}(P - \frac{\rho_C}{\rho_I}(B - C))) \quad .$$

As before the demand curve is the lower envelope of a value constraint and a liquidity constraint. Using the expressions for $F(W)$ and $V(W)$, we have

$$D_L(P) = \quad \min\{\max\left\{\underline{W} + 1 - \rho_C C - \rho_I P, 0\right\}, 1\} \quad . \tag{9}$$

$$D_V(P) = \quad \min\left\{\max\left\{\frac{1}{\alpha}\left[\hat{\underline{V}} + \alpha - P\right], 0\right\}, 1\right\}, \tag{10}$$

where

$$\hat{\underline{V}} \equiv \underline{V} + \frac{\rho_C}{\rho_I}(B - C).$$

The liquidity constraint in (9) is similar to the one in the previous section but with a coinsurance rate σ of zero and shifted leftward by the portion of the cost of insurance $\rho_C C$ due to core medical services. The value constraint in (10) reflects the inclusive valuation $\underline{V} + \frac{\rho_C}{\rho_I}(B - C)$ of the bundle of services covered by insurance.

The innovator’s pricing problem is

$$\max_{Q \in [0,1]} \rho_I D(P) P.$$

As in our analysis of basic insurance, the solution to this problem may lie along the value constraint, the liquidity constraint, or the kink between the two. To illustrate the economics of the treatment pricing, however, it is useful to focus on the case in which consumer wealth does not limit the affordability of expanded insurance, and thus the demand curve for the treatment is the value constraint. A sufficient condition for this is[30]

$$\rho_I < \frac{\underline{W} - \rho_C B}{\underline{V}}. \tag{11}$$

Condition (11) is stronger than Assumption 3, but it is in the same spirit, namely that the illness for which the treatment is needed is sufficiently rare.

When the demand curve is the value constraint, then because $\hat{\underline{V}} > \underline{V} > \alpha$ (the latter inequality from Assumption 9) it is straightforward to verify that the solution to the innovator’s profit-maximization problem is $P^* = \hat{\underline{V}}, Q^* = 1$. Under this outcome, all consumers obtain insurance. They thus obtain core medical service when they need them, and they receive the treatment in the event they become ill.[31]

Because $\hat{\underline{V}}$ is the value of the treatment for the lowest wealth individual *plus* $\frac{\rho_C(B-C)}{\rho_I}$, the innovator captures the (appropriately reweighted) expected social value of the core medical services. The comparative statics implication of this is that the treatment’s price under expanded

priced the treatment so that only the wealthy could afford insurance to cover it. However, this does not appear to be case in practice since it would entail prices of innovative treatments well above even the current very high levels.

[28] See Besanko et al. (2016) for the derivation.

[29] Strictly speaking, the expected quantity of the treatment purchased is $\rho_I D(P, \sigma)$. To simplify verbiage, throughout we refer to $D(P)$ as the demand curve for insurance and the demand curve for the treatment.

[30] To see why, invert each of the expressions in (9) and (10), and note that the vertical distance between the liquidity constraint and the value constraint is

$$\frac{\underline{W} + 1 - \rho_C C}{\rho_I} - \left(\underline{V} + \frac{\rho_C(B - C)}{\rho_I} + \alpha\right) + \left(\alpha - \frac{1}{\rho_I}\right) Q, \quad Q \in [0, 1].$$

Because $\alpha < 1 < \frac{1}{\rho_I}$, this difference is strictly decreasing in Q, and thus it is strictly positive for all $Q \in [0, 1]$ if and only if it is positive at $Q = 1$, or equivalently, $0 < \rho_I < \frac{\underline{W} - \rho_C B}{\underline{V}}$. Since $B > C$, condition (11) implies $\frac{\underline{W} - \rho_C C}{\underline{V} + \rho_C(B - C)}$, so the value constraint lies strictly below the liquidity constraint and thus constitutes the market demand curve for the treatment.

[31] The derivative of total profit with respect to price is $\hat{\underline{V}} + \alpha - 2P$. This is negative for all prices greater than or equal to the highest price at which $D_V^*(P) = 1$, which is $P = \hat{\underline{V}}$. Any lower price does not change the quantity demanded, so the optimal price must be $P = \hat{\underline{V}}$.

Table 2
Comparison of outcomes when the treatment is covered and when it is not covered by health insurance.

	Treatment not covered ($000,000)	Treatment covered ($000,000)	% difference
Price of treatment	0.66	3.28	396%
Quantity of treatment	0.66	1.00	52%
Insurance premium	0.09	0.16	73%
Consumer surplus	2.88	0.44	−85%
Producer profit	0.44	3.26	652%
Total surplus	3.32	3.72	12%

insurance is higher the rarer the illness for which the treatment is needed and the more valuable are the core medical services covered by the insurance plan. But an even more fundamental implication is that the innovator's profit-maximizing price reflects the inclusive value of the insurance bundle, not just the stand-alone value of the treatment. Indeed, we can show that, despite pricing to serve the entire market, the innovator's optimal price is strictly greater than the treatment's value for *all* consumers.

Proposition 4. *When the demand curve for the treatment coincides with the value constraint, the profit-maximizing price of the treatment P^* is strictly greater than the valuation of the treatment for each consumer in the population, i.e., $P^* > V(W)$ for all $W \in [\underline{W}, \underline{W}+1]$.*[32]

Proposition 4 provides an economic rationale for the common claim that firms in the pharmaceutical sector set prices that exceed the value created by their products.

To illustrate the innovator's profit-maximizing price—and to demonstrate that our assumptions do not restrict the model's parameters in unrealistic ways—we present a "back-of-the-envelope" numerical illustration of our model. All monetary units are in hundreds of thousand of dollars:

34 $C = 1.50$ (This is roughly the midpoint of estimates of the cost of coronary artery bypass surgery in the U.S. in the early 2010s.)[33]
35 $B = 2.25$ (This implies a benefit-cost ratio for core medical services of 1.5 to 1.)
36 $\underline{W} = 0.41$ (This equals $33,000—138 percent of the poverty line income for a family four in the U.S. in 2014, which is the point at which Medicaid eligibility phases out and ACA subsidy eligibility begins—plus $8,000, the approximate value of ACA insurance subsidies for families at this income level.)
37 $\underline{W}+1 = 1.41$ (This is roughly the 90th percentile for household income in the U.S. in the 2010s.[34])
38 $\underline{V} = 1.025$ (This assumes that the value of the treatment for the lowest-income household is 2.5 times as large as that household's wealth.)
39 $\rho_C = 0.06$ and $\rho_I = 0.02$. (The ratio of these probabilities is roughly equal to the relative incidence of heart attacks per year in the U.S.—735,000—and cases of lung cancer in the U.S. each year—210,828.)[35]
40 $\alpha = 0.89$ (In line with the Viscusi and Aldy (2003) evidence on the income elasticity of the value of life referenced above, α is chosen to make $\varepsilon_{WV}(W) = 0.55$ when evaluated at the mid-point of the wealth distribution, i.e., about $92,000).

These parameter values satisfy Assumptions 6-8 and condition (11) (so demand for health insurance is value constrained).

Table 2 compares the outcome when the treatment is covered by health insurance to one in which it is not covered by health insurance. In this latter case, we assume consumers have a choice between no insurance and a basic health insurance plan that covers only core medical services. In this case, Assumptions 6-8 imply that all consumers would purchase the basic health insurance plan, and the set of people who purchase the treatment at price P when they become ill is such that $W - \rho_C C \geq P$. The resulting demand curve for the treatment is determined by the liquidity constraint and is given by, $D^0(P) = \underline{W} + 1 - \rho_C C - P$.[36] The corresponding profit-maximizing price of the treatment when it is not covered by insurance is given by

$$P^0 = \begin{cases} \underline{W} - \rho_C C & \text{if } \underline{W} - \rho_C C \geq 1 \\ \dfrac{\underline{W} + 1 - \rho_C C}{2} & \text{if } \underline{W} - \rho_C C \leq 1 \end{cases}.$$

Coverage of the treatment induces a large increase in its price (from $\$66,000$ to about $\$328,000$) and a correspondingly large decrease in consumer surplus. The reduction in consumer surplus is more than offset by an increase in the profits of the innovator, and as a result total surplus goes

[32] This result would continue to hold without Assumption 9. In that case, the solution to the innovator's profit-maximization problem could be an interior solution in which some consumers are not served. In that case, the optimal price would be $P^* = \frac{\hat{V}+\alpha}{2} \geq \hat{V}$. (The latter inequality is a neccesary condition for an interior optimum.) Following the logic of the proof of this proposition in the Appendix, it can be verified that $P^* > V(W)$ all $W \in [\underline{W}, \underline{W}+1]$.

[33] Coronary Artery Bypass Patient Education, Frequently Asked Questions, http://cabggroupproject.web.unc.edu/frequently-asked-questions/ (accessed August 25, 2015).

[34] Capping the income data in this way implicity rules out a situation where an innovator sets an implausibly high price that would put insurance beyond the means of all but a few consumers in the top several percentiles of the income distribution.

[35] In our model, the levels of these probabilities is more difficult to pin down through back-of-the-envelope calculations since the states in our model could encompass several different medical conditions.

[36] The formal derivation of the set of consumers who purchase the treatment and the demand curve $D^0(P)$ is in Besanko et al. (2016).

up. We note that the highest valuation of the treatment (i.e., $\underline{V}+\alpha$) is $191,500, so the profit-maximizing price is more than 70 percent higher than the stand-alone value of the highest willingness-to-pay consumers and more than three times as large as the stand-alone value of the lowest-valuation consumers.

As the next proposition shows, the property in this example that coverage of the treatment causes its price to go up holds more generally.

Proposition 5. *When the demand curve for the treatment coincides with the value constraint, $P^* > P^0$ i.e., the price of the treatment when it is covered by insurance is greater than the price when it is not covered. When $\underline{W} - \rho_C C \geq 1$, $Q^0 = Q^* = 1$, and when $\underline{W} - \rho_C C < 1$, $Q^0 < Q^* = 1$, i.e., the quantity of the treatment when it is covered by insurance is at least as large as it is when the treatment is not covered and it is strictly larger when consumer wealth is sufficiently low.*[37]

In Besanko et al. (2016) we presented a case study whose results were consistent with Propositions 4 and 5. In particular, we examined the pricing of oncology drugs following the 2003 passage of Medicare Part D and found that prices were generally stable until 2003 and then increased rapidly in the years following. Prices often seemed to exceed the clinical value created by the products. Moreover, the increases in price were far greater for products likely covered by Part D compared to those that were covered under the long-running Medicare Part B program. These findings are consistent with both the greater share of the population with insurance and the bundling of patented products with those that are sold in a more competitive market.

Because all consumers receive the treatment and core medical services when the treatment is covered, total surplus in this case is at least as large as it is when the treatment is not covered, and it is strictly larger when $\underline{W} < 1$, because in that case, not all consumers purchase the treatment when it is not covered. However, analogous to Section 2, the impact on consumer surplus from expanding insurance to cover the treatment is potentially ambiguous: while the quantity of the treatment is at least as large under expanded insurance as it is under basic insurance—which works to boost consumer surplus—the price of the treatment is strictly higher, which works to reduce consumer surplus. Given Assumptions 6-9, this trade-off is, on balance, unfavorable to consumers.

Proposition 6. *When the demand curve for the treatment coincides with the value constraint, consumer surplus is lower when insurance is expanded to cover the treatment.*

The impact on consumer welfare from expanding insurance can be quite large: in Table 2 consumer surplus falls by 85 percent, driven by a 396 percent increase in the price of the treatment.

Just because aggregate consumer surplus is lower when the treatment is insured does not mean that *all* consumers are worse off. When $\underline{W} - \rho_C C < 1$, lower wealth consumers who would not have had access to the treatment if it was not covered are able to obtain it when it is covered. However, the gain in surplus for these consumers is more than offset by the loss of surplus by higher wealth (and thus higher valuation) consumers who would have purchased the treatment out-of-pocket if it was not insured. This discussion brings into sharp relief the distributional consequences of covering the treatment: it transfers surplus from higher wealth consumers to the innovator, and it may also provide an access benefit to lower wealth consumers.

3.2. Multiple innovators

Health insurance typically covers a host of services, including many innovative pharmaceutical products and medical treatments. We now consider the implications of multiple innovators, and we show that in our uniform distribution-linear value function example, the single innovator case represents a *lower bound* on the increase in price and decrease in consumer surplus due to expansion of insurance coverage. We also show, in contrast to the single-innovator case, expanding insurance to cover multiple innovative treatments can result in lower total surplus than basic insurance that does not cover the innovative treatments.

Specifically, suppose that there are N distinct states in which a consumer can experience a different serious illness, with each state having a probability $\frac{\rho_I}{N}$. For each of these possible illnesses, there is a treatment that provides an (identical) value $V(W)$ to consumer with wealth W should that state occur (and no value should any other state occur). With this specification the case of $N = 1$ reduces to the model in Section 3.1. The assumption that the treatment provides a consumer the same value in each state is not terribly restrictive if we imagine that each of the N possible illnesses would, if not treated, result in death. In this case, $V(W)$ is the value to the consumer of prolonging his or her life.

Suppose the basic insurance plan covers none of the N treatments. In that case, the liquidity constraint determines the demand for each treatment in the state in which a treatment is needed. The demand curve for each of the N treatments is thus identical to the demand curve $D^0(P) = \underline{W} + 1 - \rho_C C - P$ for the single-innovator case in Section 3.1. Under the basic insurance plan, each treatment's innovator thus sets the same price P^0 that a single innovator would set.

When expanded insurance covers each of the N treatments, the demand for each treatment depends on both the affordability of the insurance bundle as well as the value each treatment provides relative to its price. Because our goal in this section is to build intuition about what could happen, as opposed to comprehensively characterizing all possible equilibrium outcomes under expanded insurance, we focus on the same tractable special case considered in Section 3.1: the value constraint lies everywhere below the liquidity constraint and there are no uninsured buyers of any of the treatments. In the Appendix we show that the

[37] This result would continue to hold without Assumption 9: the logic used in the proof of this proposition exends to the case in which there is an interior solution in which $P^* = \frac{\underline{\hat{V}}_{+\alpha}}{2}$.

demand curve for a treatment depends on the average price $P_A = \frac{\sum_{i=1}^{N} P_i}{N}$ of all treatments:

$$D(P_A) = \min\left\{\max\left\{\frac{1}{\alpha}\left[\underline{\hat{V}} + \alpha - P_A\right], 0\right\}, 1\right\}. \tag{12}$$

Consider now the (symmetric) Nash equilibrium treatment price P_N^* under expanded insurance when all innovators set their prices simultaneously, and hereafter, let P_1^* denote the equilibrium price of the treatment in the single-innovator case. (Recall $P_1^* = \underline{\hat{V}}$.) We can show that this equilibrium price is no less than—and for sufficiently large N, strictly greater than—P_1^*, and thus with N innovators, coverage of the treatment results in higher treatment prices than would be the case if none of the treatments were covered.[38]

Proposition 7. *When the demand curve for the treatment coincides with the value constraint, there exists a critical number of innovators $N^* > 1$ such that for $N \leq N^*$, $P_N^* = P_1^* > P^0$, and for $N > N^*$, $P_N^* = \frac{N}{N+1}\left(\underline{\hat{V}} + \alpha\right) > P_1^* > P^0$, i.e., when insurance covers multiple treatments, each treatment's price is at least as high as it is in the single-innovator case and with enough innovators, strictly higher. Moreover, the price of each treatment when covered by insurance is greater than the price when the treatments are not covered.*

When $N \leq N^*$, the quantity of each treatment demanded is no different than it is in the single-innovator case, i.e., $Q_N^* = 1$. But as N goes up, when each innovator considers raising the price of its own treatment, it recognizes that the (actuarially fair) price of insurance increases by an ever smaller fraction of the increment to its own treatment price. The demand perceived by each innovator, then, becomes less price elastic, which makes a higher price potentially more attractive. Eventually, N becomes large enough that it makes sense for each innovator to raise its price above the point at which some consumers drop out of the market. In equilibrium, that number is N^*. When $N > N^*$, the corresponding quantity is

$$Q_N^* = \frac{1}{\alpha}\frac{\underline{\hat{V}} + \alpha}{N+1} < 1,$$

i.e., not all consumers purchase expanded insurance and receive the treatments when needed. (Recall that in the scenario under consideration—no uninsured purchases of treatments—the demand for insurance and the demand for the treatments are one in the same.) As N increases above N^*, the price of each treatment rises and the quantity of each treatment goes down. In the limit, as $N \to \infty$, each innovator sets a price that approaches the inclusive value of the entire insurance bundle, but since the quantity of the treatments goes to zero in this case, innovators end up realizing just a tiny fraction of this value.

The insurance bundle in this case gives rise to Cournot's complementary monopoly problem (Cournot, 1927, Chapter IX). Though the model just sketched is special, the intuition is broader. By increasing its own price, each innovator imposes a negative pecuniary externality on all other innovators. As a result, each ends up setting a higher price than a single innovator would have.

Because the multiple innovator case entails at least as high a price as in the single-innovator case, it is straightforward to show that consumer surplus in the multiple-innovator case is no greater than (and sometimes lower) than it is in the single-innovator case.[39] Consequently, in light of Proposition 6, insurance covering N treatments results in a lower consumer surplus than basic insurance.

Proposition 8. *When the demand curve for the treatment coincides with the value constraint, $CS_N^* < CS^0$, i.e., consumer surplus under insurance that covers multiple treatments is less than consumer surplus when the treatments are not covered.*

Propositions 7 and 8 imply that our results in the single-innovator case—higher price and lower consumer surplus under expanded insurance than under basic insurance—hold with even more force when we adapt our model to the case of multiple innovators. It seems fair to say that our analysis with a single innovator provides a best-case scenario for the impact of expanding insurance. With more products from the innovative portion of the pharmaceutical market included in the insurance bundle, the prices of these products would be expected to go up, potentially leading to lower consumer welfare.

There is, however, one important difference between the multiple-innovator case and the single-innovator case. In the single-innovator case, the innovator priced the product so that the entire market was served, and as a result, it generated at least as much social surplus as was generated under basic insurance. This does not necessarily hold when there are multiple treatments.

Proposition 9. *When the demand curve for the treatment coincides with the value constraint, there exists $\hat{N} > 1$, such that if $N > \hat{N}$, $TS_N^* < TS^0$, i.e., when the number of treatments is sufficiently large, total surplus under insurance that covers multiple treatments is less than total surplus when the treatments are not covered.*

In the single-innovator case, when $Q^0 < 1$ under basic insurance, expanded insurance increases total surplus by enlarging the set of consumers who have access to the treatment.[40] It is true that consumers pay a high price, but this is just a transfer between consumers and producers. However, the complementary monopoly problem that arises with multiple innovators works against increasing access: each innovator raises the price of its treatment by

[38] In the Online Appendix for Besanko et al. (2016), we show that this result extends more generally: it holds when the demand curve for the treatment is the liquidity constraint, and it holds when the equilibrium with a single innovator occurs at the kink in the demand curve.

[39] This is formally demonstrated in the proof of Proposition 8.

[40] This, of course, depends on Assumption 9. If it does not hold, even with a single innovator, the innovator's profit-maximizing price may result in only part of the market becoming insured. However, having more innovators will only make that problem worse due to the complementary monopoly problem.

so much that (for *N* large enough) some individuals decide to forego expanded insurance coverage altogether. This results in fewer consumers receiving the treatments than would be the case when the treatments are not covered, which in turn decreases total surplus.

4. Policy options

Our results demonstrate that bundling together insurance for medical products sold by monopoly manufacturers with those sold in a competitive market can decrease consumer surplus and potentially total surplus. One channel for this inefficiency is that bundling can allow firms to sell products at prices that exceed the value created. An obvious question is whether there are policy solutions that can address the resulting decrease in welfare. At a high level, policy solutions could take two forms: (1) regulating the price of products sold by monopoly manufacturers; and (2) allowing greater contractual freedom in the construction of insurance plans. We will discuss the benefits and costs of each of these solutions in turn.

To the extent that we are concerned that firms can exploit the forced bundling of insurance to charge prices that exceed value, one policy solution would be to regulate the ultimate price of goods based on the value created. While this is theoretically a means of addressing this concern, it relies heavily on the willingness and ability of policymakers to accurately determine the economic value created by the product. In practice, this is difficult to accomplish. At a minimum, it would require establishing an appropriate value-based price for the medication—a process that is fraught with difficulty and assumptions. Importantly, this economic value would need to account for more than simply the direct clinical value to the patient. For example, Lakdawalla et al. (2017) demonstrate that new innovations provide insurance value by decreasing the variance of health shocks faced by everyone in the insurance pool. Such value should be accounted for in an optimal price but would be difficult for regulators to calculate.

This becomes even more difficult in a world of heterogeneous treatment effects—which is increasingly a feature of the pharmaceutical market. Determining the appropriate value of a treatment is quite important for economic welfare since the prices above marginal costs here are meant to provide incentives for future innovation and determining the appropriate magnitude of such incentives is a herculean task. Determining the wrong price, or allowing political pressure to push prices lower, could decrease welfare by decreasing the dynamic incentives to develop new products. As we consider the benefits of a policy solution relying on a greater use of price regulation, we must consider the potential inefficiency introduced by high prices against the lost welfare from the forced bundling.

Another potential option is to allow for some decoupling of the components of the insurance bundle. This would decrease the ability of monopoly providers to capture a portion of the value created by the more competitively priced core medical products. Such a policy solution, however, would involve trading off concerns about the introduction of adverse selection with the efficiency benefits of a more targeted insurance product. The primary concern is that to the extent that individuals have private information about their need for specific medical services, tailored insurance products could introduce meaningful adverse selection. This would obviously be exacerbated in situations where individuals can purchase insurance on a relatively frequent basis and/or for conditions where the demand for the product is not immediate (Cabral, 2017).

Therefore, solutions to address the reconstruction of bundles should strive to create separate bundles, one which contains a relatively broad set of monopoly products and one that contains medical goods and services that are sold in a more competitive environment. Ideally, a broad enough set of monopoly products would limit the ability of individuals to exploit private information about their underlying health. For example, one could imagine separating out brand name and generic medications into separate bundles. Individuals would be free to purchase just the generic insurance package, but if they developed a condition that could only be covered by a patent protected medication they would either have to pay out of pocket or wait until the intellectual property protection expired. In addition, such a policy would need to be paired with restrictions on how often individuals could move in and out the monopoly bundle. If individuals could easily move into the monopoly bundle as their health state changed such an insurance product would face meaningful difficulties with adverse selection.

It therefore becomes clear that policy solutions which rely on changing the nature of the insurance bundle require some degree of commitment on the part of society to deny access to goods and services that are not covered by an individual's chosen package and exceed the individual's ability to pay. There are meaningful questions as to whether society will actually enforce such restrictions. Absent convincing people that society is willing to deny access to treatment, such policies will suffer from adverse selection.[41] To the extent that the costs of such adverse selection outweigh the benefits of a more efficient insurance package such a policy solution would not improve welfare over the status quo. In addition, depending on the type of unbundling that is undertaken there could be additional complexities introduced. For example, creating a separate bundle for branded and generic products could create a problem of coordinating benefits for individuals with conditions that could be treated by either a branded or generic product.

5. Conclusions

While health insurance bears similarities to more traditional financial insurance products, it also conveys the important and relatively unique benefit of breaking the liquidity constraint that many consumers face when attempting to purchase costly but valuable medical services, such as high-priced pharmaceutical treatments. Previous work—most notably Nyman (1998, 1999, 2003)—has shown that when prices are exogenous, this

[41] This can be thought of as a form of the “samaritan's dilemma” (Coates, 1995).

increase in access benefits consumers. However, our analysis—which endogenizes the prices charged by manufacturers of high-value products—demonstrates that while the Nyman access benefit does indeed operate, the welfare gains can be wiped out if the producers of the high-value products have sufficient monopoly pricing power. We show that consumer surplus may actually be lower when a high-value product is covered by insurance. This is not just a manifestation of the traditional intuition in a moral hazard model. In such models, increased cost sharing lowers the price of covered services and can benefit consumers. In our model, increased cost sharing can raise price of the covered treatment and can make consumers worse off. The difference between our model and the standard intuition arises because in standard models, cost sharing makes the demand for the covered product more price elastic, whereas when liquidity constraints matter this need not be the case.

We also find that the traditional bundling of pharmaceutical insurance with insurance for other medical services allows monopolist drug makers to charge prices that exceed the value created by their products. The manufacturers of innovative products must still price according to a downward sloping demand curve; however, they make their pricing decision with the knowledge that they will be bundled with products that provide a large amount of consumer surplus. Bundling allows the innovators to capture value that is created by more competitively-priced services, an effect that becomes even stronger as the number of monopoly treatments increases. This may explain why the prices of some drugs seem to exceed the economic value of their health benefits. While regulations establishing minimum coverage standards for insurance and stipulations about what pharmaceutical products must be included on a formulary can plausibly create consumer surplus gains through a more efficient and accessible insurance market, our results and discussion about potential policy solutions show that these gains must be weighed against losses that come from these pricing dynamics.

It is important to note that the vast majority of a drug maker's profits accrue in the period before patent expiration. Any policy that would address the pricing dynamics that we discuss, such as policies to unbundle drug coverage, could therefore have a dramatic impact on profits, and this, in turn, might have unintended consequences for innovative activity (Acemoglu and Linn, 2004; Finkelstein, 2004; Blume-Kohout and Sood, 2013). In addition, while there may be worries that attempts at price controls in pharmaceuticals will be welfare reducing through reduced innovation, our results suggest that some existing regulations may provide incentives for innovation that are themselves not welfare maximizing because they are based on prices that exceed value.

Finally, we note that a full characterization of pharmaceutical pricing would incorporate several mechanisms that affect equilibrium prices. It would model how the pharmaceutical firm's pricing decision takes account of the spectrum of consumer valuations for its product in light of cost sharing, the drug's intrinsic value, and liquidity constraints. It would also take into account the possibility that insurance plan covers a multitude of drugs, leading to a potential complementary monopoly problem. Our paper encompasses both of these mechanisms. However, there is a third mechanism we do not consider: possible imperfect competition among pharmaceutical companies to attain access to formularies. Such competition could lead pharmaceutical firms to trade off lower prices for access to more customers, access that could come in the form of lower cost sharing or a decreased use of other utilization management techniques such as prior authorization (Chandra and Garthwaite, 2017). As a result of this third mechanism, lower prices could be associated with decreased cost sharing, which is the same association that can arise in the model of monopoly pricing in this paper. A useful direction of future research would be to develop a theory of drug pricing that includes all three mechanisms and an empirical methodology that distinguishes among them.

Appendix

Derivation of expected utilities from purchasing and not purchasing insurance at an arbitrary premium ϕ:
Consider, first, a consumer who does not purchase insurance.

- If $s = 0$, the consumer sets $x = 0$ and $z = W$, resulting in utility $u^{NI}(s = 0) = W$.
- If $s = 1$, the consumer's utility maximization problem is $\max_{x\in\{0,1\},z\geq 0} V(W)x + z$, subject to: $Px + z \leq W$. If $W < P$, the individual is liquidity constrained from purchasing the treatment, so $x = 0$ and $z = W$, and the individual's utility is W. If $W \geq P$, the individual is not liquidity constrained from purchasing the treatment, and thus the individual could potentially purchase the treatment, in which case $x = 1$, $z = W - P$, and utility is thus $W + V(W) - P$. If the individual does not purchase the treatment, utility is just W. Purchasing the treatment is thus optimal if $V(W) \geq P$. Now, by Assumption 2, $V(W) > W$, so $W \geq P$ implies that $V(W) - P > 0$, so if an uninsured individual can afford the treatment, that individual will, in fact, purchase it, resulting in utility $W + V(W) - P$. Thus, when the individual becomes ill, his utility is

$$u^{NI}(s = 1) = \begin{cases} W & \text{if} \quad W < P \\ W + V(W) - P & \text{if} \quad W \geq P \end{cases}.$$

For a consumer with wealth W the expected utility from not purchasing insurance is thus $EU^{NI}(W) = (1 - \rho_I)u^{NI}(s = 0) + \rho_I u^{NI}(s = 1)$, or

$$EU^{NI}(W) = \begin{cases} W & \text{if} \quad W < P \\ W + \rho_I\,[V(W) - P] & \text{if} \quad W \geq P \end{cases}. \tag{13}$$

Now consider a consumer who has purchased insurance at an arbitrary premium $\phi > 0$. To do so, the consumer must have enough wealth to afford the premium, i.e., $W \geq \phi$.

- If $s = 0$, the individual sets $x = 0$ and $z = W - \phi$, resulting in utility $u^I(s = 0) = W - \phi$.

- If $s = 1$, the individual's utility maximization problem is $\max_{x\in\{0,1\},z\geq 0} V(W)x + z$, subject to: $\sigma Px + z \leq W - \phi$. Suppose, first, that $W \in [\phi, \phi + \sigma P)$. The individual cannot afford to purchase the treatment given the required copayment σP, and thus, $x = 0$, $z = W - \phi$, resulting in $u^I(s = 1) = W - \phi$. Suppose, next, that $W \geq \phi + \sigma P$. The individual can afford to purchase the treatment at the coinsurance price of σP. If does so, it sets $x = 1$, $z = W - \phi - \sigma P$ and attains utility $W + V(W) - \phi - \sigma P$. If the individual does not purchase the treatment, it gets utility $W - \phi$. Purchasing the treatment is better than not purchasing the treatment if $V(W) \geq \sigma P$. This holds because $V(W) > W \geq \phi + \sigma P \geq \sigma P$, where the first inequality follows from Assumption 2; the second inequality follows because this analysis pertains to the case in which $W \geq \phi + \sigma P$, and the third inequality follows because $\phi > 0$. Thus, when $W \geq \phi + \sigma P$, $u^I(s = 1) = W + V(W) - \phi - \sigma P$. Summarizing

$$u^I(s=1) = \begin{cases} W - \phi & \text{if } \ \phi \leq W < \phi + \sigma P. \\ W + V(W) - \phi - \sigma P & \text{if } \ W \geq \phi + \sigma P \end{cases}. \tag{14}$$

Now, $EU^I(W, \phi, \sigma) = (1 - \rho_I)u^I(s = 0) + \rho_I u^I(s = 1)$. For $W \in [\phi, \phi + \sigma P)$, we have $EU^I(W, \phi, \sigma) = W - \phi$. For $W \geq \phi + \sigma P$, we have

$$\begin{aligned} EU^I(W, \phi, \sigma) &= (1 - \rho_I)[W - \phi] + \rho_I\,[W + V(W) - \phi - \sigma P] \\ &= W - \phi + \rho_I V(W) - \rho_I \sigma P. \end{aligned}$$

Thus

$$EU^I(W,\phi,\sigma) = \begin{cases} W - \phi & \text{if } \ \phi \leq W < \phi + \sigma P \\ W - \phi + \rho_I\,[V(W) - \sigma P] & \text{if } \ W \geq \phi + \sigma P \end{cases}. \tag{15}$$

■

Proof of Lemma 1:

Proof of part (a): If $W \in [\phi, \phi + \sigma P)$, $EU^I(W) = W - \phi$. Now, if for this person, $W < P$, then from (1), $EU^{NI}(W) = W$, so clearly $EU^I(W) < EU^{NI}(W)$. Suppose, on the other hand, $W \geq P$. In that case $V(W) > W \geq P$, where the first inequality follows from Assumption 2. Thus, $EU^{NI}(W) > W > EU^I(W)$.

Proof of part (b): We actually established this as part of the above derivation: if $W \geq \phi + \sigma P$, then we have $V(W) > W \geq \phi + \sigma P \geq \sigma P$.■

Proof that the liquidity constraint is less price elastic than the value constraint:

The price elasticity along the liquidity constraint is

$$\begin{aligned} \varepsilon^L_{QP}(P) &= \frac{d(1 - F([(1-\sigma)\rho_I + \sigma]P)}{dP}\frac{P}{1 - F([(1-\sigma)\rho_I + \sigma]P)} \\ &= -[(1-\sigma)\rho_I + \sigma]P\frac{f([(1-\sigma)\rho_I + \sigma]P)}{1 - F([(1-\sigma)\rho_I + \sigma]P)}. \end{aligned}$$

The price elasticity along the value constraint is

$$\begin{aligned} \varepsilon^V_{QP}(P) &= \frac{d(1 - F(V^{-1}(P)))}{dP}\frac{P}{1 - F(V^{-1}(P))} \\ &= -\frac{dV^{-1}(P)}{dP}P\frac{f(V^{-1}(P))}{1 - F(V^{-1}(P))}. \end{aligned}$$

Let $(P_K(\sigma), Q_K(\sigma))$ be the point at which the two constraints cross for a given σ. Therefore $[(1-\sigma)\rho_I + \sigma]P_K(\sigma) = V^{-1}(P_K(\sigma))$. Also, let $W_K(\sigma) = V^{-1}(P_K(\sigma)) = [(1-\sigma)\rho_I + \sigma]P_K(\sigma)$ be the wealth of the consumer type who is just indifferent between purchasing insurance (and thus the treatment) and not purchasing when the price of the treatment is $P_K(\sigma)$. Thus, we have $P_K(\sigma) = V(W_K(\sigma))$.

Noting that $\frac{d(V^{-1}(P))}{dP} = \frac{1}{V'(V^{-1}(P))}$, we can write $\varepsilon^V_{QP}(P_K(\sigma))$ as

$$\varepsilon^V_{QP}(P_K(\sigma)) = -\frac{V(W_K(\sigma))}{V'(W_K(\sigma))}\frac{f(W_K(\sigma))}{1 - F(W_K(\sigma))}.$$

We can write $\varepsilon^L_{QP}(P_K(\sigma))$ as

$$\begin{aligned} \varepsilon^L_{QP}(P_K(\sigma)) &= -[(1-\sigma)\rho_I + \sigma]P_K(\sigma) \\ &\quad \frac{f([(1-\sigma)\rho_I + \sigma]P_K(\sigma))}{1 - F([(1-\sigma)\rho_I + \sigma]P_K(\sigma)))} \\ &= -W_K(\sigma)\frac{f(W_K(\sigma))}{1 - F(W_K(\sigma))}. \end{aligned}$$

Thus

$$\frac{\left|\varepsilon^L_{QP}(P_K(\sigma))\right|}{\left|\varepsilon^V_{QP}(P_K(\sigma))\right|} = V'(W_K(\sigma))\frac{W_K(\sigma)}{V(W_K(\sigma))} < 1.$$

■

Proof that with no insurance ($\sigma = 1$), the liquidity constraint lies below the value constraint for all $Q \in [0, 1]$:

When $\sigma = 1$, $[(1-\sigma)\rho_I + \sigma] = 1$, so we want to establish for all P

$$\begin{aligned} D_L(P, 1) - D_V(P) &= 1 - F(P) - \left(1 - F(V^{-1}(P))\right) \\ &= F(V^{-1}(P)) - F(P) < 0. \end{aligned}$$

Now from Assumption 2, for any value of x, $V(x) > x$. Because $V(\cdot)$ is strictly increasing by Assumption 1, $V^{-1}(\cdot)$ is also strictly increasing. Thus $P = V^{-1}(V(P)) > V^{-1}(P)$, and $F(P) > F(V^{-1}(P))$ since $F(\cdot)$ is increasing. ■

Proof that $\frac{dQ^*_K(\sigma)}{d\sigma} < 0$ and $\frac{dP^*_K(\sigma)}{d\sigma} > 0$:

Let $W_K(\sigma)$ be the wealth level of the consumer who, when the coinsurance rate is σ and the price is $P_K(\sigma)$, is just indifferent between purchasing insurance (and thus the treatment) and not purchasing insurance. Because $P_K(\sigma)$ is the price at which the value constraint and the liquidity constrain intersect for this given σ, this wealth level satisfies

$$V(W_K(\sigma)) = \qquad P_K(\sigma). \tag{16}$$

$$W_K(\sigma) = \qquad [(1-\sigma)\rho_I + \sigma]P_K(\sigma), \tag{17}$$

i.e., this consumer is just able to afford insurance and is just indifferent between purchasing it and not purchasing it. Differentiating each side of (16) and (17) with respect to σ gives us

$$V'(W_K(\sigma))\frac{dW_K(\sigma)}{d\sigma} = \qquad \frac{dP_K(\sigma)}{d\sigma}. \tag{18}$$

$$\frac{dW_K(\sigma)}{d\sigma} = [(1-\sigma)\rho_I + \sigma]\frac{dP_K(\sigma)}{d\sigma} + (1-\rho_I)P_K(\sigma). \tag{19}$$

Using (16), (17), and (18), we can substitute for $P_K(\sigma)$, $\frac{dP_K(\sigma)}{d\sigma}$, and $[(1-\sigma)\rho_I+\sigma]$ in (19) to get

$$\frac{dW_K(\sigma)}{d\sigma}\left\{1 - \frac{V'(W_K(\sigma))W_K(\sigma)}{V(W_K(\sigma))}\right\} = (1-\rho_I)V(W_K(\sigma)).$$

Given Assumption 1, this implies $\frac{dW_K(\sigma)}{d\sigma} > 0$, which given (18), implies $\frac{dP_K(\sigma)}{d\sigma} > 0$. Because $Q_K(\sigma) = 1 - F(W_K(\sigma))$, it follows immediately that $\frac{dQ_K(\sigma)}{d\sigma} < 0.$■

Condition under which, with full insurance ($\sigma = 0$), the liquidity constraint lies above the value constraint for all $Q \in [0,1)$:

A sufficient condition is $\rho_I < \frac{\underline{W}}{V(\underline{W})}$. This is a stronger condition than Assumption 3.

To establish the result, we want to show that $D_L(P,0) > D_V(P)$ except for P such that $D_L(P,0) = D_V(P) = 1$. When $\sigma = 0$, $[(1-\sigma)\rho_I + \sigma] = \rho_I$, so we want to show for all P

$$D_L(P,0) - D_V(P) = 1 - F(\rho_I P) - \left(1 - F(V^{-1}(P))\right)$$

$$= F(V^{-1}(P)) - F(\rho_I P) \geq 0,$$

and with strict equality for prices greater than those at which $D_L(P,0) = D_V(P) = 1$. As a first step, let $\underline{P} = V(\underline{W})$ and let $\hat{P} = \frac{\underline{W}}{\rho_I}$. For all $P \leq \underline{P}$, $D_V(P) = 1$ and for all $P \leq \hat{P}$, $D_L(P,0) = 1$. Now, $\frac{\hat{P}}{\underline{P}} = \frac{1}{\rho_I}\frac{\underline{W}}{V(\underline{W})} > 1$ by our assumption $\rho_I < \frac{\underline{W}}{V(\underline{W})}$. Thus, for $P \in [\underline{P}, \hat{P}]$, $D_L(P,0) = 1 > D_V(P)$, and for $P < \underline{P}$, $D_L(P,0) = D_V(P) = 1$. Now consider $P \geq \hat{P}$, and note that $D_L(P,0) > D_V(P)$ if and only if $V^{-1}(P) > \rho_I P$. This is equivalent to $\rho_I < \frac{V^{-1}(P)}{P}$ for all $P \geq \hat{P}$. Now, differentiating the right-hand side of this inequality gives us

$$\begin{aligned}\frac{d(\frac{V^{-1}(P)}{P})}{dP} &= \frac{\frac{dV^{-1}(P)}{dP}P - V^{-1}(P)}{P^2}\\ &= \frac{1}{P^2}\left[\frac{P}{V'(V^{-1}(P))} - V^{-1}(P)\right]\\ &= \frac{1}{P^2}\left[\frac{V(V^{-1}(P))}{V'(V^{-1}(P))} - V^{-1}(P)\right]\\ &= \frac{1}{P^2}\frac{V(V^{-1}(P))}{V'(V^{-1}(P))}\left[1 - \frac{V'(V^{-1}(P))V^{-1}(P)}{V(V^{-1}(P))}\right] > 0,\end{aligned}$$

where the inequality follows because from Assumption 1, $\frac{V'(x)x}{V(x)} < 1$ for any x. This implies that in order for $\rho_I < \frac{V^{-1}(P)}{P}$ for all $P \geq \hat{P}$, it must be the case that it holds at the lowest relevant value of P, or in other words, $\rho_I < \frac{V^{-1}(\hat{P})}{\hat{P}}$. Given the definition of $\hat{P}$, this can be written as $\underline{W} < V^{-1}(\frac{\underline{W}}{\rho_I})$, which can be rearranged as $\rho_I < \frac{\underline{W}}{V(\underline{W})}$, which is what we have assumed. Thus, $D_L(P,0) > D_V(P)$ for all $P \geq \hat{P}$, and given the earlier result, it holds for all P except those less than $\underline{P}$ over which $D_L(P,0) = D_V(P) = 1.$■

Derivation of profit-maximizing quantity and price of the innovator:

We can express the innovator's optimization problem as

$$\max_{W,P} P[1 - F(W)] \tag{20}$$

$$\text{subject to} \quad W \geq [(1-\sigma)\rho_I + \sigma]P. \tag{21}$$

$$V(W) \geq P. \tag{22}$$

$$\underline{W} \leq W \leq \overline{W}. \tag{23}$$

where W is the wealth of the marginal consumer who purchases insurance and thus quantity is $Q = 1 - F(W)$.

To derive the optimal solution, we proceed in three steps. In step 1, we consider a "relaxed" version of the optimization problem above in which we ignore constraint (22). We identify a range of σ in which the solution to that relaxed optimization problem satisfies (22), which necessarily implies that the solution to the relaxed problem also solves the full optimization problem for this range of σ. In step 2, we follow the same logic but consider a relaxed version of the optimization problem ignoring constraint (21). We show that there is a range of σ in which the solution to that relaxed optimization problem indeed satisfies (21), which necessarily implies that the solution to this relaxed problem solves the full optimization problem for this range of σ. In step 3, we show that over the remaining range of σ, the solution to the full optimization problem occurs where both (21) and (22) bind.

Step1—Analysis of relaxed problem that ignores (22): Note that in this relaxed problem the constraint (21) must bind. This is because the objective function strictly increases in P as long as $W < \overline{W}$, and $W = \overline{W}$ cannot be part of the optimal solution since it would make the objective function equal to 0, while other values of W would make it positive. Thus, we can state the relaxed problem as

$$\max_{W \geq \underline{W}} \frac{W[1 - F(W)]}{[(1-\sigma)\rho_I + \sigma]} \Leftrightarrow \max_{W \geq \underline{W}} \ln(W) + \ln[1 - F(W)]$$
$$- \ln[(1-\sigma)\rho_I + \sigma]. \tag{24}$$

This reflects that, as just noted, the constraint $W \leq \overline{W}$ must be slack. Let W_L^* denote the solution to this problem. Given Assumption 4 the objective function $\ln(W) + \ln[1 - F(W)]$ can readily be shown to be strictly concave. Thus the solution W_L^* is given by

$$\begin{aligned}&\text{If } \frac{1}{\underline{W}} - f(\underline{W}) > 0 \ , \quad \frac{f(W_L^*)W_L^*}{1 - F(W_L^*)} = 1.\\ &\text{If } \frac{1}{\underline{W}} - f(\underline{W}) \leq 0 \ , \quad W_L^* = \underline{W},\end{aligned} \tag{25}$$

where $\frac{1}{\underline{W}} - f(\underline{W}) = \frac{d[\ln(W) + \ln[1-F(W)]]}{dW}|_{W=\underline{W}}$. Let $P_L^*(\sigma) = \frac{W_L^*}{[(1-\sigma)\rho_I + \sigma]}$. Thus $(W_L^*, P_L^*(\sigma))$ is the solution to the innovator's relaxed problem when we ignore (22).

We now claim that $(W_L^*, P_L^*(\sigma))$ is the solution to the full problem when over the range $\sigma \geq \sigma_b \equiv \frac{\frac{W_L^*}{V(W_L^*)} - \rho_I}{1 - \rho_I}$. Because $V(W) > W$ from Assumption 2, it follows that $\sigma_b < 1$, and momentarily, we will verify that $\sigma_b > 0$. Thus, this range is a proper subset of [0, 1]. To verify that it solves the full optimization problem, we need to show that

$$V(W_L^*) \geq P_L^*(\sigma) \quad \text{for} \quad \sigma \geq \sigma_b = \frac{\frac{W_L^*}{V(W_L^*)} - \rho_I}{1-\rho_I}.$$

Using the expression for $P_L^*(\sigma)$ and rearranging terms, it is clear that this inequality indeed holds.

Step 2—Analysis of relaxed problem that ignores (21): Analogous to step 1, we note that in this relaxed problem constraint (22) must bind and the constraint $W \leq \overline{W}$ must be slack. Thus, we can state the relaxed problem as

$$\max_{W \geq \underline{W}} V(W)[1-F(W)] \Leftrightarrow \max_{W \geq \underline{W}} \ln(V(W)) + \ln[1-F(W)]\,. \tag{26}$$

Let W_V^* denote the solution to this problem. Given Assumption 5 the objective function $\ln(V(W)) + \ln[1-F(W)]$ is strictly concave. Thus the solution W_V^* is given by

$$\begin{array}{ll} \text{If } \dfrac{V'(\underline{W})}{V(\underline{W})} - f(\underline{W}) > 0 \ , & \dfrac{f(W_V^*)W_V^*}{1-F(W_V^*)} = \dfrac{V'(W_V^*)W_V^*}{V(W_V^*)}. \\ \text{If } \dfrac{V'(\underline{W})}{V(\underline{W})} - f(\underline{W}) \leq 0 \ , & W_V^* = \underline{W}, \end{array} \tag{27}$$

where $\frac{V'(\underline{W})}{V(\underline{W})} - f(\underline{W}) = \frac{d[\ln(V(W))+\ln[1-F(W)]]}{dW}|_{W=\underline{W}}$. Let $P_V^* = V(W_V^*)$. Thus W_V^*, P_V^* is the solution to the innovator's relaxed problem when we ignore (21).

We now claim that W_V^*, P_V^* is the solution to the full problem over the range $\sigma \leq \sigma_a \equiv \frac{\frac{W_V^*}{V(W_V^*)}-\rho_I}{1-\rho_I}$. By Assumption 3, $\sigma_a > 0$. To show that $\sigma_a \leq \sigma_b$, we first establish that if

$$\frac{1}{\underline{W}} - f(\underline{W}) > 0 \ , \quad \underline{W} \leq W_V^* < W_L^*.$$
$$\frac{1}{\underline{W}} - f(\underline{W}) \leq 0 \ , \quad W_V^* = W_L^* = \underline{W}.$$

Consider, first, the case in which $\frac{1}{\underline{W}} - f(\underline{W}) > 0$. As shown above, in this case we have an interior solution in which $\underline{W} < W_L^*$. Now as indicated in (27) there are two possibilities for W_V^*. If $W_V^* = \underline{W}$, then clearly $W_V^* < W_L^*$. If $\underline{W} < W_V^*$, then from (27) we have

$$\frac{f(W_V^*)W_V^*}{1-F(W_V^*)} = \frac{V'(W_V^*)W_V^*}{V(W_V^*)} < 1 = \frac{f(W_L^*)W_L^*}{1-F(W_L^*)}, \tag{28}$$

where the inequality follows from Assumption 1, and the equality is from (25). Now, because Assumption 4 implies $\frac{f(W)W}{1-F(W)}$ is increasing in W, (28) implies $W_V^* < W_L^*$.

Consider next the case in which $\frac{1}{\underline{W}} - f(\underline{W}) \leq 0$, so $W_L^* = \underline{W}$. By Assumption 1, $\frac{V'(\underline{W})}{V(\underline{W})} < \frac{1}{\underline{W}}$, so if $\frac{1}{\underline{W}} - f(\underline{W}) \leq 0$, then, too, $\frac{V'(\underline{W})}{V(\underline{W})} - f(\underline{W}) \leq 0$, and from (27), $W_V^* = \underline{W}$.

Now, because Assumption 1 implies $\frac{V'(W)W}{V(W)} < 1$, the function $\frac{W}{V(W)}$ is increasing in W. If $\frac{1}{\underline{W}} - f(\underline{W}) > 0$, we have

$$0 < \sigma_a = \frac{\frac{W_V^*}{V(W_V^*)} - \rho_I}{1-\rho_I} < \frac{\frac{W_L^*}{V(W_L^*)} - \rho_I}{1-\rho_I} = \sigma_b < 1.$$

Therefore the range $[0, \sigma_a)$ is contained with [0, 1] and does not overlap with $[\sigma_b, 1]$. (Note that this confirms the claim above that $\sigma_b > 0$.)

If, by contrast, $\frac{1}{\underline{W}} - f(\underline{W}) \leq 0$, then we have $0 < \sigma_a = \frac{\frac{W_V^*}{V(W_V^*)}-\rho_I}{1-\rho_I} = \frac{\frac{\underline{W}}{V(\underline{W})}-\rho_I}{1-\rho_I} = \frac{\frac{W_L^*}{V(W_L^*)}-\rho_I}{1-\rho_I} = \sigma_b < 1$.

Finally to show that $\left(P_V^*, W_V^*\right)$ solves the full optimization problem, we need to show that

$$W_V^* \geq [(1-\sigma)\rho_I + \sigma]P_V^* \quad \text{for} \quad \sigma \leq \sigma_a = \frac{\frac{W_V^*}{V(W_V^*)} - \rho_I}{1-\rho_I}.$$

Because $P_V^* = V(W_V^*)$, by rearranging terms on the left-hand side of the above inequality we can see that the indeed holds for $\sigma \leq \sigma_a$.

Step 3: We now consider the range $\sigma \in (\sigma_a, \sigma_b)$, and we show that both (21) and (22) in the full optimization problem must bind. Note that this case pertains only to $\frac{1}{\underline{W}} - f(\underline{W}) > 0$; otherwise (σ_a, σ_b) is empty.

First, note that it cannot be the case that neither (21) and (22) are non-binding because the objective function is strictly increasing in P. This means that either (i) (21) is binding and (22) is slack; (ii) (21) is slack and (22) is binding; or (iii) both are binding. But we have just shown that when case (i) arises $W = W_L^*$ and the only way for (22) to be slack is if $\sigma \geq \sigma_b$. By analogous logic, case (ii) can arise only if $\sigma \leq \sigma_a$. Thus, for $\sigma \in (\sigma_a, \sigma_b)$, the only possible solution to the innovator's optimization problem is for both constraints to bind. That then implies that for $\sigma \in (\sigma_a, \sigma_b)$, the solution is $P_K(\sigma), W_K(\sigma)$ given by (16) and (17) above.

We now show that $W_V^* < W_K(\sigma) < W_L^*$ over the range $\sigma \in (\sigma_a, \sigma_b)$ (again recognizing that with $\frac{1}{\underline{W}} - f(\underline{W}) > 0$, we have $\underline{W} < W_L^*$). From (16) and (17)

$$\frac{V_K(\sigma)}{W_K(\sigma)} = [(1-\sigma)\rho_I + \sigma]\,.$$

Straightforward algebra establishes that $W_K(\sigma_a) = W_V^*$ and $W_K(\sigma_b) = W_L^*$. Moreover, previously in this appendix, it was established that $\frac{dW_K(\sigma)}{d\sigma} > 0$. Therefore, $W_K(\sigma) \in (W_V^*, W_L^*)$ for $\sigma \in (\sigma_a, \sigma_b)$.

Summing up, the solution to the innovators problem is $W^*(\sigma), P^*(\sigma), Q^*(\sigma)$ is as follows:

If $\frac{1}{\underline{W}} - f(\underline{W}) > 0$, or equivalently $\underline{W}f(\underline{W}) < 1$,

$$\begin{array}{rcl} W^*(\sigma) & = & \begin{cases} W_V^* & \sigma \in [0, \sigma_a] \\ W_K(\sigma) & \sigma \in [\sigma_a, \sigma_b] \\ W_L^* & \sigma \in [\sigma_b, 1] \end{cases} \\ P^*(\sigma) & = & \begin{cases} V(W_V^*) & \sigma \in [0, \sigma_a] \\ P_K(\sigma) & \sigma \in [\sigma_a, \sigma_b] \\ \dfrac{W_L^*}{(1-\sigma)\rho_I + \sigma} & \sigma \in [\sigma_b, 1] \end{cases} \\ Q^*(\sigma) & = & \begin{cases} 1-F(W_V^*) & \sigma \in [0, \sigma_a] \\ Q_K(\sigma) & \sigma \in [\sigma_a, \sigma_b] \\ 1-F(W_L^*) & \sigma \in [\sigma_b, 1] \end{cases} , \end{array} \tag{29}$$

where $W_K(\sigma)$ and $P_K(\sigma)$ are given by (16) and (17), $\sigma_a = \frac{\frac{W_V^*}{V(W_V^*)} - \rho_I}{1-\rho_I} < \frac{\frac{W_L^*}{V(W_L^*)} - \rho_I}{1-\rho_I} = \sigma_b$, and $Q_K(\sigma) = 1 - F(W_K(\sigma))$.

If $\frac{1}{\underline{W}} - f(\underline{W}) \leq 0$, or equivalently, $\underline{W}f(\underline{W}) \geq 1$,

$$\begin{aligned} W^*(\sigma) &= \underline{W} \\ P^*(\sigma) &= \begin{cases} V(W_V^*) & \sigma \in [0, \sigma_a] \\ \dfrac{W_L^*}{(1-\sigma)\rho_I + \sigma} & \sigma \in [\sigma_b, 1] \end{cases} . \\ Q^*(\sigma) &= 1, \end{aligned} \tag{30}$$

where $\sigma_a = \frac{\frac{\underline{W}}{V(\underline{W})} - \rho_I}{1-\rho_I} = \sigma_b$.

■

Proof of Proposition 2:

Let $W^*(\sigma) = W(P^*(\sigma), \sigma)$ be the marginal consumer at the profit-maximizing price. Thus,

$$TS^*(\sigma) = W_M + \rho_I \int_{W^*(\sigma)}^{\overline{W}} V(t)f(t)dt \tag{31}$$

If $f(\underline{W})\underline{W} \geq 1$, then $W_L^* = W_V^* = \underline{W}$, and $TS^*(\sigma)$ is a constant $W_M + \rho_I \int_{\underline{W}}^{\overline{W}} V(t)f(t)dt$ for all $\sigma \in [0, 1]$. It weakly attains its maximum at $\sigma = 0$.

Consider, then, $f(\underline{W})\underline{W} < 1$, so that $W_L^* > W_V^* \geq \underline{W}$. Now $Q^*(\sigma) = 1 - F(W^*(\sigma))$, so $W^*(\sigma) = F^{-1}(1 - Q^*(\sigma))$, where $F^{-1}(\cdot)$ is the inverse of $F(\cdot)$. From (29),

$$Q^*(0) \geq Q^*(\sigma) \quad \text{for} \quad \sigma \in (0, 1] \quad ,$$

and the inequality is strict for $\sigma \in (\sigma_a, 1]$. Thus,

$$W^*(0) \leq W^*(\sigma) \quad \text{for} \quad \sigma \in (0, 1],$$

with strict inequality for $\sigma \in (\sigma_a, 1]$. From (31), it follows that

$$TS^*(0) \geq TS^*(\sigma) \quad \text{for} \quad \sigma \in (0, 1],$$

and in particular, the inequality for is strict for $\sigma \in (\sigma_a, 1]$.■

Proof of Lemma 2:

Consumer surplus $CS^*(\sigma)$ can be written as

$$CS^*(\sigma) = W_M + \rho_I \left[\int_{W^*(\sigma)}^{\overline{W}} V(t)f(t)dt - P^*(\sigma)[1 - F(W^*(\sigma)] \right],$$

where, as in the previous proof, $W^*(\sigma)$ is the marginal consumer at the profit-maximizing price.

$$\frac{dCS^*(\sigma)}{d\sigma} = -\rho_I \left\{ [V(W^*(\sigma)) - P^*(\sigma)]f(W^*(\sigma))\frac{dW^*(\sigma)}{d\sigma} - [1 - F(W^*(\sigma)]\frac{dP^*(\sigma)}{d\sigma} \right\}. \tag{32}$$

Now, if $f(\underline{W})\underline{W} < 1$, we have

$$\sigma \in [0, \sigma_a], \frac{dW^*(\sigma)}{d\sigma} = 0 \text{ and } \frac{dP^*(\sigma)}{d\sigma} = 0 \Rightarrow \frac{dCS^*(\sigma)}{d\sigma} = 0.$$

$$\sigma \in (\sigma_a, \sigma_b), \begin{bmatrix} V(W^*(\sigma)) - P^*(\sigma) = V(W_K(\sigma)) - P_K(\sigma) = 0 \\ \text{and } \dfrac{dP^*(\sigma)}{d\sigma} = \dfrac{dP_K(\sigma)}{d\sigma} > 0 \end{bmatrix} \Rightarrow \frac{dCS^*(\sigma)}{d\sigma} < 0.$$

$$\sigma \in [\sigma_b, 1], \frac{dW^*(\sigma)}{d\sigma} = 0 \text{ and } \frac{dP^*(\sigma)}{d\sigma} < 0 \Rightarrow \frac{dCS^*(\sigma)}{d\sigma} > 0.$$

Thus, as we move from full insurance to no insurance, consumer surplus as a function of σ is initially flat, then decreases and eventually increases. It must attain its maximum at either $\sigma = 0$ or $\sigma = 1$, establishing part (b) of the lemma. If $f(\underline{W})\underline{W} \geq 1$, then as shown above, $\sigma_a = \frac{\frac{\underline{W}}{V(\underline{W})} - \rho_I}{1-\rho_I} = \sigma_b$, so

$$\sigma \in \left[0, \frac{\frac{\underline{W}}{V(\underline{W})} - \rho_I}{1 - \rho_I}\right], \frac{dCS^*(\sigma)}{d\sigma} = 0$$

$$\sigma \in \left[\frac{\frac{\underline{W}}{V(\underline{W})} - \rho_I}{1 - \rho_I}, 1\right], \frac{dCS^*(\sigma)}{d\sigma} > 0.$$

Thus, consumer surplus is initially flat and then eventually increases as $\sigma \to 1$, establishing part (a) of the lemma.■

Proof of Proposition 3: Since $\underline{W}f(\underline{W}) < 1$, and $\lim_{\theta \to \overline{\theta}} \varepsilon_{WV}(W, \theta) = 1$, then in a neighborhood of $\theta < \overline{\theta}$, we must have $\underline{W}f(\underline{W}) < \varepsilon_{WV}(\underline{W}, \theta)$. Rearranging this yields $\frac{V'(\underline{W})}{V(\underline{W})} - f(\underline{W}) > 0$. From (27) it follows that in this neighborhood $W_V^*(\theta)$ is given by

$$\frac{f(W_V^*(\theta))W_V^*(\theta)}{1 - F(W_V^*(\theta))} = \frac{\partial V(W_V^*(\theta), \theta)}{\partial W} \frac{W_V^*(\theta)}{V(W_V^*(\theta), \theta)}.$$

Cancelling $W_V^*(\theta)$ from each side of this expression and totally differentiating with respect to θ gives us

$$\frac{\partial[\frac{f(W_V^*(\theta))}{1-F(W_V^*(\theta))} - \frac{\frac{\partial V(W_V^*(\theta),\theta)}{\partial W}}{V(W_V^*(\theta),\theta)}]}{\partial W} \frac{dW_V^*}{d\theta} - \frac{\partial\left[\frac{\partial V(W_V^*(\theta),\theta)}{\partial W} V(W_V^*(\theta), \theta\right]}{\partial \theta} = 0 \quad .$$

The log-concavity of $V(W)[1 - F(W)]$ from Assumption 5 implies that the left-most term above is positive, and by the premise of the proposition $\frac{\partial \varepsilon_{WV}(W,\theta)}{\partial \theta} = \frac{\partial\left[\frac{\partial V(W,\theta)}{\partial W} \frac{W}{V(W,\theta)}\right]}{\partial \theta} > 0$, so it follows that $\frac{\partial\left[\frac{\partial V(W,\theta)}{\partial W} V(W,\theta)\right]}{\partial \theta} > 0$ as well. Thus we must have $\frac{dW_V^*}{d\theta} > 0$. Since $\lim_{\theta \to \overline{\theta}} \varepsilon_{WV}(W, \theta) = 1$, from (27), $\lim_{\theta \to \overline{\theta}} W_V^*(\overline{\theta}) = W_L^*$. It follows that $\lim_{\theta \to \overline{\theta}} P^*(0, \theta) = V(W_L^*) > W_L^* = P^*(1)$, where the inequality holds from Assumption 2. Moreover, from (5), $\lim_{\theta \to \overline{\theta}} \left[CS^*(0, \theta) - CS^*(1)\right] < 0$

because $\lim_{\theta \to \overline{\theta}} \int_{W_V^*(\theta)}^{W_L^*} \left[V(t) - V(W_V^*(\theta))\right] = 0$ and as just noted $\lim_{\theta \to \overline{\theta}} V(W_V^*(\theta)) = V(W_L^*) > W_L^*$. By continuity, there exists a neighborhood θ below $\overline{\theta}$ such that $P^*(0, \theta) > P^*(1)$ and $CS^*(0, \theta) < CS^*(1)$.■

Proof that the demand curve for the treatment under expanded insurance does not include insured purchasers:

Let $D_i(P)$ be the demand curve for the treatment from consumers who have insurance and let $D_u(P)$ denote potential demand for the treatment from consumers who do not have insurance.

$$D_i(P) = 1 - F\left(\max\{\rho_C C + \rho_I P, V^{-1}\left(P - \frac{\rho_C}{\rho_I}(B - C)\right)\right), \tag{33}$$

$$D_u(P) = \int_{\{W \in [\underline{W}, \underline{W}+1] | W < \rho_C C + \rho_I \text{ and } W \geq P\}} dt$$

We now prove that given Assumption 7, which is equivalent to $\underline{W} - \rho_C C \geq \frac{\rho_C \rho_I}{1-\rho_I} C$, $D_u(P) = 0$. We first note that a necessary condition for $D_u(P) > 0$ for some price P' is that at that price the treatment is less expensive than health insurance, i.e., $P' < \rho_I P' + \rho_C C$. Otherwise, all consumers could afford health insurance and would be "located" on demand curve $D_i(P)$. This can be rewritten as $P' < \frac{\rho_C}{1-\rho_I} C$, where $\frac{\rho_C}{1-\rho_I} < 1$. A second necessary condition is that there be at least some consumers who could not afford health insurance when the treatment price is P', i.e., $\underline{W} - \rho_C C < \rho_I P'$. Otherwise, all consumers would be "located" on $D_i(P)$. This can be rewritten as $P' > \frac{\underline{W} - \rho_C C}{\rho_I}$. These two necessary conditions imply $\frac{\underline{W} - \rho_C C}{\rho_I} < \frac{\rho_C}{1-\rho_I} C$, or equivalently $\underline{W} < \frac{\rho_C}{1-\rho_I} C$. However, this contradicts Assumption 7. Given our assumptions, then, the demand curve for the treatment is $D_i(P)$. ■

Proof of Proposition 4:

Note that

$$P^* = \hat{\underline{V}} = \underline{V} + \frac{\rho_C}{\rho_I}(B - C). \tag{34}$$

Given Assumption 8, we have $\frac{\rho_C}{\rho_I}(B - C) > \underline{V} + \alpha > \alpha$, which in conjunction with (34) implies $P^* > \underline{V} + \alpha > V(W)$ for all $W \in [\underline{W}, \underline{W} + 1]$, where the second inequality follows because $\alpha > 0$, and $\underline{V} + \alpha$ is the highest possible consumer valuation.■

Proof of Proposition 5:

We have two possibilities. First, if $\underline{W} - \rho_C C \geq 1$ we have corner solutions when the treatment is not covered by health insurance and when it is covered: $P^0 = \underline{W} - \rho_C C$, $P^* = \hat{\underline{V}}$; $Q^0 = Q^* = 1$. Thus, the quantity of the treatment is the same in either case, but because $\hat{\underline{V}} > \underline{V} > \underline{W} > \underline{W} - \rho_C C$, the price is higher when the treatment is covered by insurance than when it is not. Second, if $\underline{W} - \rho_C C < 1$ we have a corner solution when the treatment is covered by insurance but an interior solution when it is not covered: $Q^0 = \frac{\underline{W} - \rho_C C + 1}{2} < 1 = Q^*$ i.e., the quantity of the treatment is greater when insurance covers the treatment. Comparing the prices we have,

$$P^* - P^0 = \hat{\underline{V}} - \left(\frac{\underline{W} + 1}{2}\right). \tag{35}$$

Because $V(\underline{W} + 1) = \underline{V} + \alpha > \underline{W} + 1$ (from Assumption 1) and $\hat{\underline{V}} > \underline{V} > \alpha$ (from Assumption 9) it follows that $2\hat{\underline{V}} > \underline{W} + 1$, which given (35) implies that $P^* - P^0 > 0$.■

Proof of Proposition 6:

Under either basic and expanded insurance, consumer surplus when $D(P)$ consumers demand the treatment at price P, and Y consumers obtain core medical services can be shown to equal

$$CS(P, Y) = \underline{W} + \frac{1}{2} + \rho_C(B - C)Y$$
$$+\rho_I\left[(\underline{V} + \alpha - P)D(P) - \frac{\alpha}{2}D(P)^2\right]. \tag{36}$$

In what follows it will be convenient to use a linear transformation of (36), constructed by dividing it by ρ_I and dropping the constant term $\frac{\underline{W} + \frac{1}{2}}{\rho_I}$:

$$cs(P, Y) = \frac{\rho_C(B - C)}{\rho_I}Y + (\underline{V} + \alpha - P)D(P) - \frac{\alpha}{2}D(P)^2. \tag{37}$$

Under expanded insurance, $P = P^* = \hat{\underline{V}}$, $D(P^*) = 1$, and $Y = 1$, so we have

$$cs^* = cs(\hat{\underline{V}}, 1) = \frac{\alpha}{2}, \tag{38}$$

and under basic insurance we have

$$P^0 = \begin{cases} \underline{W} - \rho_C C & \text{if } \underline{W} - \rho_C C \geq 1 \\ \dfrac{\underline{W} - \rho_C C + 1}{2} & \text{if } \underline{W} - \rho_C C \leq 1 \end{cases},$$

$$D^0(P^0) = \begin{cases} 1 & \text{if } \underline{W} - \rho_C C \geq 1 \\ \dfrac{\underline{W} + 1 - \rho_C C}{2} & \text{if } \underline{W} - \rho_C C \leq 1 \end{cases},$$

and $Y = 1$; thus

$$cs^0 = \frac{\rho_C(B-C)}{\rho_I} + \begin{cases} \underline{V} - (\underline{W} - \rho_C C) + \dfrac{\alpha}{2} & \underline{W} - \rho_C C \geq 1 \\ \left(\dfrac{\underline{W} - \rho_C C + 1}{2}\right) \begin{pmatrix} \underline{V} + \dfrac{3\alpha}{4} \\ -\dfrac{\alpha(\underline{W} - \rho_C C)}{4} - \dfrac{\underline{W} - \rho_C C}{2} - \dfrac{1}{2} \end{pmatrix} & \underline{W} - \rho_C C \leq 1 \end{cases} . \tag{39}$$

If $\underline{W} - \rho_C C \geq 1$, $Q^0 = Q^* = 1$, so $cs^0 = \underline{V} - (\underline{W} - \rho_C C) + \frac{\alpha}{2} > \frac{\alpha}{2} = cs^*$. The inequality follows because by Assumption 2, $V(\underline{W}) = \underline{V} > \underline{W}$.

When $\underline{W} - \rho_C C < 1$,

$$cs^0 - cs^* = \frac{\rho_C(B-C)}{\rho_I} + \frac{(\underline{W} - \rho_C C) + 1}{2}\left(\underline{V} + \frac{3\alpha}{4} - \frac{\alpha(\underline{W} - \rho_C C)}{4} - \frac{(\underline{W} - \rho_C C)}{2} - \frac{1}{2}\right) - \frac{\alpha}{2}. \tag{40}$$

To show that this expression is positive, we note that

$$\begin{aligned} \frac{\partial[cs^0 - cs^*]}{\partial\alpha} &= (\frac{(\underline{W} - \rho_C C) + 1}{2})(\frac{3}{4} - \frac{(\underline{W} - \rho_C C)}{4}) - \frac{1}{2} \\ &= -\frac{1}{8}[(\underline{W} - \rho_C C) - 1]^2 < 0, \end{aligned}$$

since we are in the case in which $\underline{W} - \rho_C C < 1$. To establish $cs^0 - cs^* > 0$, it suffices to show that $cs^0 - cs^*|_{\alpha=\underline{V}'} > 0$. This is because $\alpha < 1$, and as we will prove below, $\underline{\hat{V}} > 1$.

Substituting $\underline{\hat{V}}$ for α in (40) gives us

$$cs^0 - cs^*|_{\alpha=\underline{\hat{V}}} = \frac{\rho_C(B-C)}{\rho_I} + \frac{(\underline{W} - \rho_C C) + 1}{2} \begin{pmatrix} \underline{V} + \dfrac{3\underline{\hat{V}}}{4} - \dfrac{(\underline{W} - \rho_C C)\,\underline{\hat{V}}}{4} \\ -\dfrac{(\underline{W} - \rho_C C)}{2} - \dfrac{1}{2} \end{pmatrix} - \frac{\underline{\hat{V}}}{2} \tag{41}$$

Because $\underline{V} = \underline{\hat{V}} - \frac{\rho_C(B-C)}{\rho_I}$, we can write the second term on the right-hand side of (41) as

$$\frac{(\underline{W} - \rho_C C) + 1}{2} \begin{pmatrix} \underline{\hat{V}} - \dfrac{\rho_C(B-C)}{\rho_I} + \dfrac{3\underline{\hat{V}}}{4} - \dfrac{(\underline{W} - \rho_C C)\,\underline{\hat{V}}}{4} \\ -\dfrac{(\underline{W} - \rho_C C)}{2} - \dfrac{1}{2} \end{pmatrix} - \frac{1}{2}\left(\underline{\hat{V}} - \frac{\rho_C(B-C)}{\rho_I}\right).$$

Substituting this into (41) and rearranging terms gives us

$$\begin{aligned} cs^0 - cs^*|_{\alpha=\underline{\hat{V}}} \quad = \quad & \left(\frac{\rho_C(B-C)}{\rho_I}\right)\left(\frac{3}{2} - \frac{(\underline{W} - \rho_C C) + 1}{2}\right) \\ & + \frac{(\underline{W} - \rho_C C) + 1}{2} \begin{pmatrix} \underline{\hat{V}} + \dfrac{3\underline{\hat{V}}}{4} - \dfrac{(\underline{W} - \rho_C C)\underline{\hat{V}}}{4} \\ -\dfrac{(\underline{W} - \rho_C C)}{2} - \dfrac{1}{2} \end{pmatrix} - \frac{\underline{\hat{V}}}{2}. \end{aligned}$$

Now recall that we are considering the case in which we have an interior solution under basic insurance, i.e., $Q^0 = \frac{\underline{W}+1-\rho_C C}{2} < 1$.Thus $\frac{3}{2} - \frac{(\underline{W}-\rho_C C)+1}{2} > 0$, and so

$$\begin{aligned}
cs^0 - cs^*|_{\alpha=\underline{\hat{V}}} &> \frac{(\underline{W}-\rho_C C)+1}{2}\left(\underline{\hat{V}} + \frac{3\underline{\hat{V}}}{4} - \frac{(\underline{W}-\rho_C C)\underline{\hat{V}}}{4} - \frac{(\underline{W}-\rho_C C)}{2} - \frac{1}{2}\right) - \frac{\underline{\hat{V}}}{2} \\
&= \frac{3(\underline{W}-\rho_C C)\underline{\hat{V}}}{4} - \frac{(\underline{W}-\rho_C C)^2\underline{\hat{V}}}{8} - \frac{(\underline{W}-\rho_C C)^2}{4} - \frac{(\underline{W}-\rho_C C)}{2} + \frac{3\underline{\hat{V}}}{8} - \frac{1}{4} \\
&\geq \frac{3(\underline{W}-\rho_C C)\underline{\hat{V}}}{4} - \frac{(\underline{W}-\rho_C C)\underline{\hat{V}}}{8} - \frac{(\underline{W}-\rho_C C)}{4} - \frac{(\underline{W}-\rho_C C)}{2} + \frac{3\underline{\hat{V}}}{8} - \frac{1}{4} \\
&= \frac{5(\underline{W}-\rho_C C)\underline{\hat{V}}}{8} - \frac{3(\underline{W}-\rho_C C)}{4} + \frac{3\underline{\hat{V}}}{8} - \frac{1}{4} \\
&\geq \frac{5(\underline{W}-\rho_C C)}{8} - \frac{3(\underline{W}-\rho_C C)}{4} + \frac{3}{8} - \frac{1}{4} \\
&= -\frac{(\underline{W}-\rho_C C-1)}{8} > 0.
\end{aligned} \tag{42}$$

The second and fourth inequalities follows because $\underline{W} - \rho_C C < 1$ and thus $(\underline{W} - \rho_C C)^2 < \underline{W} - \rho_C C$. The third inequality follows because, as we now show, $\underline{\hat{V}} > 1$.

By Assumption 2, $\underline{V} + \alpha > \underline{W} + 1$, so it follows that

$$\underline{\hat{V}} > \underline{W} + 1 - \alpha + \frac{\rho_C}{\rho_I}(B - C). \tag{43}$$

By Assumption 8, $\frac{\rho_C}{\rho_I}(B - C) > \underline{V} + \alpha$, or equivalently, $\frac{\rho_C}{\rho_I}(B - C) - \alpha > \underline{V}$, which from (43) implies $\underline{\hat{V}} > \underline{W} + \underline{V} + 1 > 1$,since $\underline{V} > \underline{W} \geq 0$. Given the implications in (42), it follows that consumer surplus under basic insurance is greater than consumer surplus under expanded insurance.■

Derivation of the demand curve for treatment under expanded insurance and multiple innovators

When actuarially fair expanded insurance covers each of the N treatments, the net value of purchasing expanded insurance for a consumer with wealth W can be shown to be

$$EU^I(W) = W + \rho_C(B - C) + \rho_I(V(W) - P_A), \tag{44}$$

where $P_A \equiv \frac{\sum_{i=1}^{N}}{N}$ is the average price of all treatments. The expression in (44) is derived in the same way we derived the expected utility from purchasing expanded insurance that covered a single treatment.

To simplify the analysis, we focus on a plausible special case: there is no demand for any of the treatments from consumers who are uninsured. We discuss sufficient conditions for this special case momentarily. For now we note that it implies $EU^{NI}(W) = W$. It follows that a consumer purchases a treatment in the relevant state only if he or she is insured, which occurs if (a) insurance is affordable; (b) insurance is valuable, i.e., $W \geq \rho_C C + \rho_I P_A$, and $EU^I(W) \geq EU^{NI}(W)$, or $V(W) \geq P_A - \frac{\rho_C}{\rho_I}(B - C)$. The demand curve for any individual treatment is the measure of consumers such that

$$\frac{W - \rho_C C}{\rho_I} \geq P_A \quad \text{and} \quad V(W) \geq P_A - \frac{\rho_C}{\rho_I}(B - C) \quad .$$

This is given by the expression in (33) with P_A substituted for P. If the value constraint lies everywhere below the liquidity constraint—and recall a sufficient condition for this is (11)—the demand curve for each treatment is:

$$D(P_A) = \min\left\{\max\left\{\frac{1}{\alpha}\left[\underline{\hat{V}} + \alpha - P_A\right], 0\right\}, 1\right\}. \tag{45}$$

Given this demand curve, the smallest price that any innovator would charge in a symmetric Nash equilibrium would be $\underline{\hat{V}}$. To see why, suppose there was a symmetric Nash equilibrium in which all firms charged a price less than $\underline{\hat{V}}$. This would imply $P_A < \underline{\hat{V}}$. But this would imply that a single innovator could slightly raise its price while keeping $P_A < \underline{\hat{V}}$. This would be a profitable deviation, contradicting the supposition that the Nash equilibrium is less than $\underline{\hat{V}}$. Thus a sufficient condition for the earlier-stated assumption that there is no demand for a treatment from an uninsured consumer would be $\underline{\hat{V}} > \overline{W} + 1$, which ensures that at the lowest price an innovator would set in equilibrium, an uninsured consumer with the highest possible wealth would be unable to afford the treatment. (Note that this condition is satisfied in our numerical example in Section 3.1.)

Proof of Proposition 7:

When the demand curve for the treatment is the value constraint, with a single innovator we have a corner solution in which the innovator serves the entire market, i.e., $Q^* = D(P^*) = D(\underline{\hat{V}}) = 1$. A necessary condition for profit maximization in this case is that the innovator could not increase profits by raising price, i.e., $D(P^*) + P^* \frac{dD(P^*)}{dP} \leq 0$,or $1 - \frac{\underline{\hat{V}}}{\underline{\alpha}} \leq 0$, which holds because $\underline{\hat{V}} > \underline{V} > \alpha$.

Innovator $k \in \{1, \ldots, N\}$ has profit $\pi(P_1,\ldots P_k,\ldots,P_N) = P_k D(P_A)$. Thus,

$$\begin{aligned}
\frac{d\pi}{dP_k} &= \frac{1}{\alpha}\left[\underline{\hat{V}} + \alpha - P_A - \frac{P_k}{N}\right] \\
&= \frac{1}{\alpha}\left[\underline{\hat{V}} + \alpha - \frac{\sum_{i \neq k} P_i}{N} - \frac{2P_k}{N}\right].
\end{aligned}$$

Suppose, now, all innovators but k set a price $\underline{\hat{V}}$. Innovator k has no incentive to set a lower price than $\underline{\hat{V}}$ since the quantity demanded when it sets $\underline{\hat{V}}$ equals 1, and it would continue to be 1 if innovator k's price was lower than $\underline{\hat{V}}$. Would innovator k set a higher price than $\underline{\hat{V}}$? If not, then the Nash equilibrium price is $\underline{\hat{V}}$. Now note that $\frac{d\pi}{dP_k}$ decreases in P_k, so if this expression is non-positive at $P_k = \underline{\hat{V}}$ it will be negative for all $P_k > \underline{\hat{V}}$. Thus, it suffices to evaluate

$$\frac{d\pi}{dP_k}|_{P_1=\ldots=P_N=\underline{V}} = 1 - \frac{\underline{\hat{V}}}{N\alpha}.$$

Since $\underline{\hat{V}} > \underline{V} > \alpha$, this expression will be non-positive for $N \leq N^* \equiv \frac{\underline{\hat{V}}}{\alpha}$. For such N, the symmetric Nash equilibrium price $P_N^* = \underline{\hat{V}}$. However, for $N > N^*$,

$$1 - \frac{\underline{\hat{V}}}{\alpha}\frac{1}{N} > 0.$$

This implies that the corner solution price that was optimal for a single innovator would no longer be an equilibrium for a group of innovators because each one would have an incentive to raise price. In this case, we have an interior equilibrium given by $\frac{d\pi}{dP_k}|_{P_1=\ldots=P_N=P_N^*} = 0$, or

$$P_N^* = \frac{N}{N+1}\left(\underline{\hat{V}} + \alpha\right) > \underline{\hat{V}} = P^*,$$

where the inequality follows because $N > \frac{\underline{\hat{V}}}{\alpha}$.■

Proof of Proposition 8:

If $N < \frac{\underline{\hat{V}}}{\alpha}$, we have a corner solution in which $Q_N^* = 1$ and $P_N^* = \underline{\hat{V}}$, just as in the single-firm case. Letting CS_N^* denote equilibrium consumer surplus with multiple innovators under expanded insurance and CS^0 denote equilibrium consumer surplus under basic insurance (which is independent of the number of the innovators given the arguments in the main text), we have $CS_N^* = CS_1^* < CS^0$, where the inequality follows from Proposition 6.

If $N \geq \frac{\underline{\hat{V}}}{\alpha}$, we have an interior solution with N innovators, and $P_N^* = \frac{N}{N+1}\left(\underline{\hat{V}} + \alpha\right) > \underline{\hat{V}} = P_1^*$. Having already proved that $CS^0 > CS_1^*$, we just need to prove $CS_1^* > CS_N^*$ to establish the result.

Under expanded insurance $Q = Y = D(P_N^*)$, so given (36)

$$\begin{aligned} CS_N^* &= \underline{W} + \frac{1}{2} + \rho_C(B-C)D(P_N^*) + \rho_I\left[(\underline{V} + \alpha - P_N^*)D(P_N^*) - \frac{\alpha}{2}D(P_N^*)^2\right] \\ &= \underline{W} + \frac{1}{2} + \rho_I\left[(\underline{\hat{V}} + \alpha - P_N^*)D(P_N^*) - \frac{\alpha}{2}D(P_N^*)^2\right]. \end{aligned} \tag{46}$$

Because we are in the case in which the demand curve for the treatment is the value constraint, $D(P)$ is given by (12).

Now, for $P \in (\underline{\hat{V}}, \underline{\hat{V}} + \alpha)$ let's express the term in square brackets in (46) as a function of P:

$$cs(P) = (\underline{\hat{V}} + \alpha - P)D(P) - \frac{\alpha}{2}D(P)^2.$$

Differentiating with respect to P and noting that $D(P) = \frac{1}{\alpha}\left[\underline{\hat{V}} + \alpha - P\right]$ for $P \in (\underline{\hat{V}}, \underline{\hat{V}} + \alpha)$ implies that

$$\frac{dcs(P)}{dP} = -D(P) < 0, \quad P \in (\underline{\hat{V}}, \underline{\hat{V}} + \alpha). \tag{47}$$

Since $P_1^* < P_N^*$ when $N \geq \frac{\underline{\hat{V}}}{\alpha}$, we have, light of (47),

$$CS_1^* = \underline{W} + \frac{1}{2} + \rho_I cs(P_1^*) > \underline{W} + \frac{1}{2} + \rho_I cs(P_N^*) = CS_N^*.$$

Thus, $CS^0 > CS_N^*$.■

Proof of Proposition 9:

The $\hat{N}$ we refer to in the statement of the proposition is $\hat{N} = \max\left\{N^*, \frac{2}{\alpha}\frac{\underline{\hat{V}}+\alpha}{\underline{W}+1-\rho_C C} - 1\right\}$, where, recall, $N^* = \frac{\underline{\hat{V}}}{\alpha}$.

Using (36), we can express total surplus as

$$\begin{aligned} TS(P, Y) = {} & \underline{W} + \frac{1}{2} + \rho_C(B-C)Y \\ & + \rho_I\left[(\underline{V} + \alpha)D(P) - \frac{\alpha}{2}D(P)^2\right], \end{aligned}$$

or equivalently,

$$TS(Q, Y) = \underline{W} + \frac{1}{2} + \rho_C(B-C)Y + \rho_I\left[(\underline{V} + \alpha)Q - \frac{\alpha}{2}Q^2\right].$$

Further, note that

$$TS(Q, Q) = \underline{W} + \frac{1}{2} + \rho_I\left[\left(\underline{\hat{V}} + \alpha\right)Q - \frac{\alpha}{2}(Q)^2\right].$$

We note that $TS(Q, Y)$ increases in Q (holding Y fixed) for all $Q \leq 1$, and $TS(Q, Q)$ also increases in Q for all $Q \leq 1$. Finally, let TS^0 and TS_N^* be total surplus under basic insurance and expanded insurance with N innovators, respectively. We have

$$TS^0 = TS(Q^0, 1) \geq TS(Q^0, Q^0) = \left(\underline{\hat{V}} + \alpha\right)Q^0 - \frac{\alpha}{2}\left(Q^0\right)^2,$$

and

$$TS_N^* = TS(Q_N^*, Q_N^*) = \left(\underline{\hat{V}} + \alpha\right)Q_N^* - \frac{\alpha}{2}\left(Q_N^*\right)^2.$$

To establish $TS_N^* < TS^0$, it suffices to show $Q_N^* < Q^0$. If $\underline{W} - \rho_C C \geq 1$, $Q^0 = 1$, and as we saw in the proof of Proposition 7, if $N \geq N^* = \frac{\underline{\hat{V}}}{\alpha}$, then $Q_N^* < 1$. Since, by assumption, $N > \hat{N}$, then $N > N^*$, so indeed, $Q_N^* < Q^0 = 1$. If $\underline{W} - \rho_C C < 1$, $Q^0 = \frac{\underline{W}+1-\rho_C C}{2}$. Since $N > \hat{N}$, then $N > N^*$ and $N > \frac{2}{\alpha}\frac{\underline{\hat{V}}+\alpha}{\underline{W}+1-\rho_C C} - 1$. Together, these imply $Q_N^* = \frac{1}{\alpha}\frac{\underline{\hat{V}}+\alpha}{N+1} < \frac{\underline{W}+1-\rho_C C}{2} = Q^0$.■

References

Acemoglu, Daron, Linn, Joshua, 2004. Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry. Q. J. Econ. 119 (3), 1049–1090.

Becker, Gary, 2007. Health as Human Capital: Synthesis and Extensions. Oxford Econ. Papers 59 (3), 379–410.

Besanko, David, Dranove, David, Garthwaite, Craig, 2016. Insurance and the High Prices of Pharmaceuticals. NBER Working Paper No. 22353.

Blume-Kohout, Margaret, Sood, Neeraj, 2013. Market Size and Innovation: effects of Medicare Part D on Pharmaceutical Research and Development. J. Public Econ. 97, 327–336.

Coate, Stephen, 1995. Altruism, the Samaritan's Dilemma and Government Transfer Policy. Am. Econ. Rev. 85 (1), 46–57.

Cabral, Marika, 2017. Claim Timing and Ex Post Adverse Selection. Rev. Econ. Stud. 84 (1), 1–44.

Chandra, Amitabh, Garthwaite, Craig, 2017. The Economics of Indication-Based Pricing. N. Engl. J. Med. 377 (2), 103–106.

Cournot, Augustin, 1927. In: Bacon, Nathaniel T. (Ed.), Researches into the Mathematical Principles of the Theory of Wealth. Macmillan, New York.

Einav, Liran, Finkelstein, Amy, Ryan, Stephen P., Schrimpf, Paul, Cullen, Mark R., 2013. Selection on Moral Hazard in Health Insurance. Am. Econ. Rev. 103 (1), 178–219.

Feldman, Roger, Dowd, Bryan, 1991. A New Estimate of the Welfare Loss of Excess Health Insurance. Am. Econ. Rev. 81 (1), 531–537.

Feldstein, Martin, 1973. The welfare loss of excess health insurance. J. Polit. Econ. 61, 251–280.

Finkelstein, Amy, 2004. Static and Dynamic Effects of Health Policy: Evidence From the Vaccine Industry. Q. J. Econ. 119 (2), 527–564.

Friedman, Bernad, 1974. Risk Aversion and the Consumer Choice of Health Insurance Option. Rev. Econ. Stat. 56 (2), 209–214.

Geruso, Michael, Layton, Timothy, Prinz, Daniel, 2017. Screening in Contract Design: Evidence from the ACA Health Insurance Exchanges. NBER Working Paper No. 22832.

Jena, Anupam, Philipson, Tomas, 2013. Endogenous Cost-effectiveness Analysis and Health Care Technology Adoption. J. Health Econ. 32, 172–180.

Kniesner, Thomas, Viscusi, W. Kip, 2019. The Value of a Statistical Life. Oxford Research Encyclopedia of Economics and Finance, forthcoming; Vanderbilt Law Research Paper No. 19-15. Available at SSRN: https://ssrn.com/abstract=3379967.

Lakdawalla, Darius, Sood, Neeraj, 2009. Health Insurance as a Two-Part Pricing Contract. J. Public Econ. 102, 1–12.

Lakdawalla, Darius, Anup Malani, and Julian Reif (2017). The Insurance Value of Medical Innovation, Journal of Public Economics, vol. 145, pp. 94–102.

Loftus, Peter, 2015. How Much Should Cancer Drugs Cost. Wall Street J., June 18.

Managan, Dan, 2015. Pricey New Cholesterol Rx Covered by Big Drug Plan, but..., CNBC.com, http://www.cnbc.com/2015/10/06/pricey-new-cholesterol-rx-covered-by-big-drug-plan-but.html (accessed April 4, 2016).

Manning, Will M., Marquis, Susan, 1996. Health Insurance: the Tradeoff Between Risk Pooling and Moral Hazard. J. Health Econ. 15 (5), 609–640.

Newhouse, Joseph, 1993. Free for All? Lessons from the RAND Health Insurance Experiment. Harvard University Press, Cambridge, MA.

Nyman, John A., 1998. Theory of Health Insurance. J. Health Adm. Educ. 16 (1), 41–66.

Nyman, John A., 1999. Health Insurance: the Access Motive. J. Health Econ. 18 (2), 141–152.

Nyman, John A., 2003. The Theory of Demand for Health Insurance. Stanford University Press, Stanford, CA.

Pauly, Mark, 1968. The Economics of Moral Hazard: Comment. Am. Econ. Rev. 58 (3), 531–537.

Viscusi, W. Kip, Aldy, Joseph E., 2003. The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World. J. Risk Uncertainty 27 (1), 5–76.

Walker, Joseph, 2015. High Prices for Drugs Attacked at Meeting. Wall Street J. (June 1).