

NBER WORKING PAPER SERIES

INSURANCE AND THE HIGH PRICES OF PHARMACEUTICALS

David Besanko
David Dranove
Craig Garthwaite

Working Paper 22353
<http://www.nber.org/papers/w22353>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2016

We would like to thank Dan Barron, Amitabh Chandra, Niko Matouschek, Fiona Scott Morton, Matthew Notodiwidgo, and Michael Powell for their helpful comments and suggestions. We are grateful to Rena Conti for her comments and for giving us permission to use the data discussed in Section The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2016 by David Besanko, David Dranove, and Craig Garthwaite. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Insurance and the High Prices of Pharmaceuticals
David Besanko, David Dranove, and Craig Garthwaite
NBER Working Paper No. 22353
June 2016
JEL No. H0,H51,I0,I1,I11,L1,L13

ABSTRACT

We present a model in which prospective patients are liquidity constrained, and thus health insurance allows patients access to treatments and services that they otherwise would have been unable to afford. Consistent with large expansions of insurance in the U.S. (e.g., the Affordable Care Act), we assume that policies expand the set of services that must be covered by insurance. We show that the profit-maximizing price for an innovative treatment is greater in the presence of health insurance than it would be for an uninsured population. We also show that consumer surplus is less than it would be if the innovation was not covered. These results show that even in the absence of moral hazard, there are channels through which insurance can negatively affect consumer welfare. Our model also provides an economic rationale for the claim that pharmaceutical firms set prices that exceed the value their products create. We empirically examine our model's predictions by studying the pricing of oncology drugs following the 2003 passage of Medicare Part D. Prior to 2003, drugs covered under Medicare Part B had higher prices than those that would eventually be covered under Part D. In general, the trends in pricing across these categories were similar. However, after 2003 there was a far greater increase in prices for products covered under Part D, and as result, products covered by both programs were sold at similar prices. In addition, these prices were quite high compared to the value created by the products---suggesting that the forced bundle of Part D might have allowed firms to capture more value than their products created.

David Besanko
Department of Strategy
Kellogg School of Management
Northwestern University
2001 Sheridan Road
Evanston, IL 60208
d-besanko@kellogg.northwestern.edu

Craig Garthwaite
Department of Strategy
Kellogg School of Management
Northwestern University
2001 Sheridan Road
Evanston, IL 60208
and NBER
c-garthwaite@kellogg.northwestern.edu

David Dranove
Department of Strategy
Kellogg School of Management
Northwestern University
2001 Sheridan Road
Evanston, IL 60208
d-dranove@kellogg.northwestern.edu

Insurance and the High Prices of Pharmaceuticals*

David Besanko[†]

David Dranove[‡]

Craig Garthwaite[§]

June 9, 2016

Abstract

We present a model in which prospective patients are liquidity constrained, and thus health insurance allows patients access to treatments and services that they otherwise would have been unable to afford. Consistent with large expansions of insurance in the U.S. (e.g., the Affordable Care Act), we assume that policies expand the set of services that must be covered by insurance. We show that the profit-maximizing price for an innovative treatment is greater in the presence of health insurance than it would be for an uninsured population. We also show that consumer surplus is less than it would be if the innovation was not covered. These results show that even in the absence of moral hazard, there are channels through which insurance can negatively affect consumer welfare. Our model also provides an economic rationale for the claim that pharmaceutical firms set prices that exceed the value their products create. We empirically examine our model's predictions by studying the pricing of oncology drugs following the 2003 passage of Medicare Part D. Prior to 2003, drugs covered under Medicare Part B had higher prices than those that would eventually be covered under Part D. In general, the trends in pricing across these categories were similar. However, after 2003 there was a far greater increase in prices for products covered under Part D, and as result, products covered by both programs were sold at similar prices. In addition, these prices were quite high compared to the value created by the products—suggesting that the forced bundle of Part D might have allowed firms to capture more value than their products created.

Keywords: pharmaceutical prices, health insurance access, liquidity constraints, complementary monopoly

JEL classification: I11, I13, L13

1 Introduction

The welfare economics of insurance traditionally involves a trade-off between consumption smoothing and moral hazard. The potential welfare effects of these two features of health insurance have been extensively studied (Pauly, 1968; Feldstein, 1973; Friedman, 1974; Feldman and Dowd, 1991; Manning and Marquis, 1989, 1996; Newhouse, 1993; Einav *et al.*, 2013). Nyman (1999) identifies the elimination of liquidity constraints for the purchasing of costly medical services as a third feature of health insurance. While eliminating liquidity constraints likely improves access to expensive medical products, little is known about the welfare implications of this increased access.

On the one hand, Nyman (1999) proposes that the access benefits of health insurance result in a large increase in consumer surplus in the following manner: “if insurance is the only way to gain access to expensive health care, then the value of insurance for that care is the expected consumer surplus from the health care services that would otherwise be inaccessible. This value may be large, especially if the procedures that insurance made possible were life-saving.”

*We would like to thank Dan Barron, Amitabh Chandra, Niko Matouschek, Fiona Scott Morton, Matthew Notodiwigdo, and Michael Powell for their helpful comments and suggestions. We are grateful to Rena Conti for her comments and for giving us permission to use the data discussed in Section 6 of the paper. We would also like to acknowledge Matthew Schmidt for his helpful research assistance in the early stages of this project.

[†]Kellogg School of Management, Northwestern University

[‡]Kellogg School of Management, Northwestern University

[§]Kellogg School of Management, Northwestern University and NBER

This proposed benefit of health insurance has not been lost on policymakers. Two large and recent federal expansions of social and private health insurance in the United States, 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, because they would allow patients to access treatments and services that they otherwise would have been unable to afford.¹ In addition to providing access by breaking liquidity constraints, these expansions also expanded and solidified regulations regarding the set of products and services that must be bundled together into a minimally acceptable insurance product.²

In this paper we examine the welfare implications of increased access to expensive medical services by providing a demand model that incorporates two salient features of the United States pharmaceutical market: (1) liquidity constrained customers and (2) firms selling medical goods and services in both competitive and monopoly markets that are covered by a single insurance bundle. We use this model to examine pricing and consumer welfare in the pharmaceutical sector—a market that is characterized by many products with high and rising prices that sometimes exceed the value directly created by the product. These unprecedented high prices have recently received considerable attention from both policymakers and economists including calls for both direct (e.g. explicit price controls) and indirect (e.g. increased bargaining rights for government purchasers) policies to decrease prices. 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’s prices and their effect on welfare.

Broadly speaking, the expensive branded drugs attracting attention fall into two categories.³ The first category includes speciality medications that offer treatments for life threatening and/or severely debilitating conditions. These drugs carry hefty price tags but the prices are lower than estimates of the value of the statistical lives saved by the products. Despite the positive benefit/cost ratio, the profits generated by these new products have attracted a great deal of negative attention from policymakers and the general public.⁴ The poster children for this first category are Sovaldi and Harvoni, two drugs developed by Gilead for the treatment of hepatitis C. These drugs offer a marked improvement over existing treatments and initially were sold at a price of \$80,000 for a 12 week course of treatment. Given their efficacy and price, they have attracted many customers and generated approximately \$10 billion in profits in 2014 (Slatko, 2015). Despite these high prices, Sovaldi represents such a leap forward in the treatment of hepatitis C that it is estimated to be cost effective (Liu et al., 2015; Chhatwal et al., 2015; Najafzadeh et al., 2015).

The second category includes drugs that provide some benefit to patients but are sold at prices that appear to exceed the value that they create. This category includes many recently introduced oncology products as well products that treat a wide number of other conditions. For example, when used to treat

¹For example, a white paper prepared by Senate Democrats in support of the Patient Protection and Affordable Care Act opens by stating: "The Patient Protection and Affordable Care Act will ensure that all Americans have access to quality, affordable health care." Responsible Reform for the Middle Class www.dpc.senate.gov/healthreformbill/healthbill04.pdfdemocrats.senate.gov/reform (accessed 6/25/2015).

²These regulations are both explicit and implicit. For Medicare Part D, the regulations cover what types of products must be covered and many private insurance companies have adopted these standards. For the ACA, the legislation and subsequent regulations established a set of services that must be covered for qualify as insurance for the purposes of the individual mandates to purchase insurance.

³In addition there have been recent increases in the prices for generic drugs serving small markets. This primarily relates to the minimum efficient scale relative to the market size for these small patient populations and are generally beyond the scope of this paper (except to note that the profit maximizing price for these generic medications is likely far lower for an uninsured population).

⁴For example, Gilead was forced to justify to Congress the high prices it charged for its hepatitis C cure and was subsequently the target of an 18 month investigation by the Senate Finance committee (Pollack, 2014; Berkrot, 2015). This attention to prices is driven, in part, by the belief that the total potential liability insurers face from specialty pharmaceutical products will lead to higher premiums and reduced access in the private market, as well as create budget problems for public programs. Steve Miller, the Chief Medical Officer of the Pharmacy Benefit Manager Express Scripts, said about Sovaldi, "What they [Gilead] have done with this particular drug will break the country ... [i]t will make pharmacy benefits no longer sustainable. Companies just aren't going to be able to handle paying for this drug."

metastatic colon cancer, the current price per quality adjusted life year for Stivarga is over \$700,000 – far more than the value suggested by most economic studies (Goldstein et al., 2015). Similarly, a price based on efficacy and the standard value of life years gained for the recently developed PCSK9 Inhibitors such as Praluent and Repatha (which treat high cholesterol) would have been 75 percent lower than the list price (Mangan, 2015).⁵

Our model helps explain both the high prices of recently developed valuable pharmaceuticals and the puzzling ability of firms to sell products at prices that exceed the value created by their products. We also examine the welfare consequences of the access to expensive pharmaceuticals that insurance provides. We begin by noting that the average uninsured customer appears to be liquidity constrained in the market for expensive branded pharmaceuticals. In 2011, the median United States household had a net worth of approximately \$68,000, with nearly \$50,000 of that wealth being home equity.⁶ In addition, survey evidence reports that roughly 75 percent of people with incomes less than \$50,000 and 67 percent of individuals with incomes between \$50,000 and \$100,000 would have difficulty coming up with \$1,000 to pay an unexpected expense (Sweet and Swanson, 2016). Since most banks would not provide an unsecured loan for medical services, uninsured individuals find themselves without the financial means to afford a large fraction of these new expensive treatments. It therefore seems probable that at current prices, many potential patients would have to forego treatment with many of these new high priced drugs, were it not for insurance.

Given these facts, Nyman’s conclusion that insurance creates an access benefit for enrollees seems to be correct, especially for new high priced drugs.⁷ However, understanding the welfare implications of this benefit requires carefully considering the optimal pricing response of pharmaceutical firms.⁸ These potential pricing decisions can be quite meaningful and failing to consider the supply side response to health insurance therefore can have important welfare implications. For example, it is difficult to imagine that Gilead would charge anywhere near \$84,000 for Sovaldi were it not covered by insurance and therefore the existence of insurance for this product has clear implications for access and pricing.

To illustrate this point we develop a simple demand model, where individuals may purchase insurance for a "core" set of medical services (i.e. existing hospital and physician services as well as more competitively priced pharmaceuticals) and a potentially life-saving innovative product (e.g. a new pharmaceutical treatment for cancer). With this model we address five distinct questions that have heretofore been unaddressed by the literature: (1) To what extent might the unprecedented high prices of new prescription drugs depend on the liquidity benefits of insurance? (2) Would consumers be better off if these drugs were not insured? (3) How does bundling innovative products sold by patent-holding firms and “core” medical services in a single product affect pricing and welfare? (4) How can innovating firms sell products for prices that exceed their value? (5) How does the growing fragmentation of the pharmaceutical marketplace affect pricing?

While our model is too stylized to offer precise numerical predictions, a canonical example that broadly fits real world data confirms some of the model’s surprising implications. First, we show that the profit-maximizing price for products is far greater in the presence of health insurance than it would be for an uninsured population, and consumer surplus is lower. The decrease in consumer surplus is quite large and exceeds the consumption smoothing benefits of health insurance for these products. We further show that

⁵While list prices are not often paid by insurers, the nations largest pharmacy benefit manager noted that they did not receive the cost effective price following their negotiations related to these drugs (Managan, 2015).

⁶Data available at: http://www.census.gov/people/wealth/files/Wealth_Tables_2011.xlsx

⁷We also note that the fact that liquidity constraints may place the willingness to pay of customers below the social value of a good is not limited to the area of health care or health insurance. Fuchs, Green, and Levine (2013) develop a model describing the optimal subsidies for mosquito bednets in developing countries where liquidity constraints result in the willingness to pay for these products to be below their social value.

⁸Nyman’s model makes the strong assumption of fixed prices and this shortcoming drives the erroneous conclusion that the access motive for insurance increases consumer surplus.

this negative effect on consumer welfare grows as the number of innovative products increases. In fact, as the number of innovative products covered by insurance increases, the high prices they charge for their products eventually cause some consumers to not purchase insurance, resulting in a decrease in total surplus. We note that these results do not rely on moral hazard in that it exists even when customers never buy a product at a price that exceeds its value. Therefore, these results show that even in a situation where insurers could eliminate moral hazard (either through well designed cost sharing or effective managed care organizations) there are still other channels through which insurance can reduce welfare.

In addition, our model provides an economic explanation for the common claim that firms in the pharmaceutical sector are able to successfully set prices that far exceed the value created by their products (Loftus, 2015; Walker, 2015). It is true that insured consumers don't directly pay the cost of drugs, and this might lead to moral hazard in their purchasing decisions. However, it is also true that if all manufacturers priced above the value created, the resulting insurance premiums would push consumers out of the insurance market. This effect on insurance premiums means that even if insured customers exhibit moral hazard, the monopoly providers' pricing should be constrained by the value created by their products. However, our model demonstrates that requiring insurers to provide a single bundle including existing basic medical services and innovative treatments allows the monopoly manufacturers of innovative products to set prices that capture some of the value created in the more competitive market for other health services, value that was previously consumer surplus. This would explain recent examples of oncology and cardiovascular products that are covered by health insurance and have been found to not be cost effective, but whose prices are accepted by insurers (Johnson, 2015; Magnan, 2015). We note that the ability of firms to capture more value than they create solely results from regulations requiring an insurance bundle and that without these regulations, even in the presence of moral hazard, a monopolist selling an innovative product would be constrained by the effect that its price would have on insurance premiums.

To empirically examine our model's predictions we examine the pricing of oncology drugs following the 2003 passage of Medicare Part D. While the widespread adoption of insurance in the American market makes finding an ideal empirical setting for testing our model difficult, there are many features of the oncology market and Medicare Part D that are useful for this purpose. First, Part D required that oncology products had to be covered which allows the producers of newly launched products in this category to exert a large amount of market power when setting prices. Second, a number of oncology products were previously covered by Medicare because those products were primarily administered in a physician's office or outpatient setting (and thus are covered under Medicare's Part B medical benefit rather than the Part D pharmaceutical benefit).

We exploit this fact to identify the effect of insurance on pharmaceutical pricing by first examining the general trends occurring in the oncology market and then separately examining products based on insurance coverage. While there are a relatively small number of oncology products, we show that prices were generally stable until 2003 and then increased rapidly in the years following. This increase would be consistent with both the greater share of the population with insurance and the bundling of new products under patent with those that are sold in a more competitive market. To further examine these prices, we next compare the change in prices for products based on whether they are covered under Part D and Part B. We find that prior to 2003, drugs covered under Part B were far more expensive than those that would eventually be covered under Part D. In general, the trends in pricing across these categories were flat and similar before 2003. However, after that year there was a far greater increase in prices for products covered under Part D. As a result, in the years following 2003, products covered by both programs were sold at similar prices. In addition, these prices were quite high compared to the value created by the products – suggesting that the

forced bundle of Part D might have allowed firms to capture more value than their products created.

Finally, our model provides suggestive explanations for the recent increase in new high priced pharmaceutical products and the increases for existing products. We show that firms must consider at least two factors when setting the price of their pharmaceutical: (1) the value created by their product; (2) the externality of their price on the premiums for insurance and therefore the purchase of other pharmaceuticals. Our results with multiple innovative products, sold by different firms, suggest that as the number of innovators increase there is a commensurate increase in prices and above a certain level a decrease in total surplus. This result stems from the Cournot complementary monopolist problem where each innovator is pricing without internalizing its effect on the market. However, firms marketing a portfolio of innovative products must internalize this negative externality to a greater degree than firms that are marketing a single product. We document that the time period of recent price increases has been coincident with a marked reduction in the concentration of sales in the pharmaceutical sector—providing an economic rationale for the recent price increases. We note that this is not simply a conventional market structure-to-price relationship, as many of the drugs in question held monopoly or near monopoly status over this time period. Rather, the reduction in concentration encourages firms to ignore the externality that their high prices impose on the rest of the industry through their impact on insurance premiums.

The remainder of this paper is organized as follows. Section 2 presents and analyzes a model of pricing a pharmaceutical product in the face of consumer liquidity constraints. We first consider the case in which insurance does not cover the treatment, and then we consider what happens when the treatment is covered. For both cases, 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 encompass multiple innovators. Section 4 incorporates a consumption-smoothing motive for insurance into our model and discusses how it affects our results. Section 5 studies the impact of minimum insurance coverage standards. Section 6 contains the empirical analysis of Medicare Part D and oncology products, and Section 7 applies the model to provide insight into the role played by insurance in recent instances of high drug prices. Section 8 summarizes and concludes. Proofs of all propositions and derivations of some key expressions are in the Appendix.

2 The model

The focus of our analysis is the pricing of a single innovative treatment for a serious but rare illness. In Section 3, we extend the model to consider treatments for several different innovators. As in Nyman (1998, 1999), consumers are assumed to be risk neutral, and so the consumption-smoothing motive for insurance is absent.⁹ Potential consumers of the innovative treatment (hereafter the treatment) differ only in the amount of their liquid wealth \widehat{W}_i . We assume that $\widehat{W}_i \sim \text{UNIFORM}[\underline{\widehat{W}}, \overline{\widehat{W}}]$, and we set $\overline{\widehat{W}} = \underline{\widehat{W}} + 1$.¹⁰

A consumer can be in one of three possible mutually exclusive states, $s = \{0, 1, 2\}$:

- $s = 0$: The consumer is healthy and does not need medical services.
- $s = 1$: The consumer experiences a medical condition for which a package of well-established “core” medical services is required.
- $s = 2$: The consumer experiences the illness for which an innovative treatment is required.

⁹We discuss consumption smoothing in Section 4.

¹⁰The choice of the uniform density allows us to compute exact formulas for prices, quantities, and surpluses. None of our key intuitions depends on the uniform distribution.

The probability of state 1 is $\rho_C \in (0, 1)$; the probability of state 2 is $\rho_I \in (0, 1)$, and the probability of state 0 is $1 - \rho_C - \rho_I > 0$. Think of state 1 as the occurrence of a medical condition (e.g., a heart attack) for which procedures, protocols, and treatments are well understood. The cost of treating these conditions may be high, but their benefits are likely to be higher, and because they are well established, these core medical services would typically be covered under conventional health insurance plans. State 2, on the other hand, is a serious but relatively rare illness for which an innovative, high value-added pharmaceutical treatment has only recently emerged.¹¹ Throughout the analysis, we will treat ρ_C as “somewhat small” and ρ_I as “very small” (though we emphasize that none of our result depend on any particular ordering of ρ_C and ρ_I).

In light of the above discussion, consumer utility is

$$U_i(x_i, y_i, z_i, s) = \begin{cases} z_i & s = 0 \\ By_i + z_i - F_1 & s = 1 \\ V_i x_i + z_i - F_2 & s = 2 \end{cases},$$

where $z_i \in [0, \infty)$ is the quantity of a numeraire consumption good; $y_i \in \{0, 1\}$ denotes the decision to purchase the package of core medical services (where $y_i = 1$ constitutes buying the package); $x_i \in \{0, 1\}$ denotes the decision to purchase the treatment (where $x_i = 1$ constitutes purchasing the treatment); B is the value of core medical services (for simplicity, assumed to be the same for all consumers); F_s is a fixed utility loss if the individual needs medical services, and $V_i = v(\widehat{W}_i) > 0$ is the incremental value of the treatment. Note that the treatment has value only when the consumer develops the serious illness, and core medical services only have value when the consumer develops the medical condition for which those services are appropriate.¹² The utility losses plays no essential role in our analysis, so we set $F_s = 0$.

The price of core medical services is assumed to be determined competitively and equals $C \in [0, B)$. By contrast, the price P of the treatment is determined monopolistically by the innovator of the treatment. For simplicity, the innovator’s marginal cost is assumed to be 0. This innovator’s price is endogenous, and its determination is a key focus of our analysis.

Consumers have the opportunity to purchase health insurance. We consider two alternative forms of health insurance: *basic health insurance*, which covers just core medical services but not the treatment and *expanded health insurance*, which covers both core medical services and the treatment. Under the basic health insurance plan, the individual pays a price of 0 for core medical services but the full price P for the treatment. Under the expanded health insurance plan, the price for both core medical services and the treatment is 0. A consumer who does not purchase health insurance pays the full price for both core medical services and the treatment. When the treatment is included on the health plan’s formulary, we assume that consumers cannot purchase *a la carte*. That is, a consumer cannot purchase coverage only for the treatment or only for core medical services.¹³ This is consistent with rules of what constitutes insurance under the Affordable Care Act.¹⁴

Health insurance is sold in a competitive market, so it has an actuarially fair premium. When the treatment is not covered, that premium is $\rho_C C$. When the treatment is covered, the plan’s premium is $\rho_C C + \rho_I P$.

¹¹Below we extend the model to include multiple serious illnesses, each having its own innovative treatment.

¹²It 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.

¹³As we discuss below, the basic economics of the case when insurance for the treatment is decoupled from insurance can be considered by setting $\rho_C = 0$ in our analysis of treatment pricing.

¹⁴Note that insurers play an entirely passive role in this model. For example, if the price of the treatment increases, this is entirely translated into higher premiums. As an alternative to premium increases, insurers might instead increase cost sharing requirements. While the qualitative implications for demand and pricing strategies are similar, the mechanism through which affects drug purchases is more complex. We therefore do not consider this alternative to premium increases

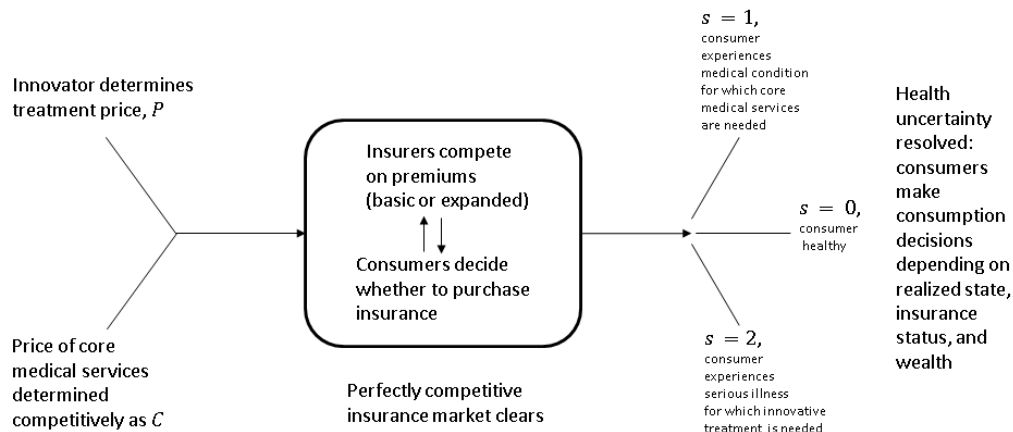


Figure 1: The timing of consumer choice and market activity.

A consumer’s choice problem proceeds in two stages. Before learning the state s , an individual chooses whether to purchase a health insurance plan. Then, depending on the individual’s first-stage choice and the realized state, the individual chooses x_i, y_i , and z_i . Figure 1 summarizes the timing of both consumer choice and market activity.

Throughout the analysis, we make the following assumptions:

Assumption 1 *The incremental value of the treatment $V_i = v(\widehat{W}_i)$ is increasing in wealth and is such that $v(\widehat{W}_i) > \widehat{W}_i$ for all $\widehat{W}_i \in [\underline{\widehat{W}}, \overline{\widehat{W}}]$.*

Assumption 2 *$\underline{\widehat{W}} + 1 < C$, i.e., no consumer can afford core medical services.*

Assumption 3 *$\underline{\widehat{W}} > \rho_C C$, i.e., every consumer can afford basic health insurance.*

Assumption 4 *$\rho_I v(\underline{\widehat{W}} + 1) < \rho_C (B - C)$, i.e., the expected social value of the treatment is less than the expected social value of core medical services.*

Assumption 1 is consistent with evidence that the income elasticity of the value of life is positive (Viscusi and Aldy, 2003).¹⁵ However, the assumption that $V_i > \widehat{W}_i$ 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

¹⁵Assumption 1 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_i in our model—is “rising in initial wealth.”

that treatment at less than \$2,000. However, this revealed preference logic ignores the reality of liquidity constraints. Assumption 1 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.

With respect to the other assumptions, Assumption 2 is plausible in a setting characterized by significant medical cost inflation, as in the U.S.¹⁶ A justification for Assumption 3 is that wealth levels \widehat{W}_i are inclusive of subsidies for the purchase of health insurance, as under the Affordable Care Act. Assumptions 2 and 3 create a potential access motive to obtain basic health insurance in the spirit of Nyman (1998, 1999).¹⁷ Finally, Assumption 4 is plausible in light of the notion that the serious illness is rare, while the core medical services cover an array of established treatments that have sufficiently high value B when needed. We discuss the roles played by these assumptions below. Later in the paper we present numerical calculations that illustrate that these assumptions can be satisfied by empirically plausible parameter values.

In the Appendix we derive the expected utility of a consumer from *not* purchasing insurance, EU_i^{NI} , the expected utility from purchasing the basic health insurance plan, EU_i^{BI} , and (when available) the expected utility from purchasing expanded health insurance, EU_i^I :

$$EU_i^{NI} = \widehat{W}_i + \rho_I (V_i - P) 1(\widehat{W}_i \geq P). \quad (1)$$

$$EU_i^{BI} = \widehat{W}_i + \rho_C (B - C) + \rho_I (V_i - P) 1[\widehat{W}_i - \rho_C C \geq P]. \quad (2)$$

$$EU_i^I = \widehat{W}_i + \rho_C (B - C) + \rho_I (V_i - P), \quad (3)$$

where in (1) and (2), $1(t)$ is the indicator function: $1(t) = 1$ if $t = TRUE$, and 0 otherwise.

2.1 Case 1: Treatment is not covered by health insurance

2.1.1 Demand for the treatment under basic insurance

We begin by characterizing both consumer choice for the basic health insurance plan and consumer demand for the treatment.

Proposition 1 *All consumers purchase basic health insurance. Only consumers for whom $\widehat{W}_i - \rho_C C \geq P$ purchase the treatment if they become ill. Thus, when health insurance does not cover the treatment, the demand for the treatment is fully determined by consumer liquidity constraints.*

The result that all consumers purchase basic health insurance is *not just because* of Nyman’s access motive. For consumers whose wealth is in the interval $[P, P + \rho_C C)$ not purchasing basic insurance would free up wealth that breaks the liquidity constraint in the event that the consumer becomes ill. That is, the possibility of the serious illness creates the potential for a “reverse Nyman effect” wherein a consumer may want to save money that would have been spent on basic insurance to free up resources to treat the serious illness should it arise. Thus, it is conceivable not all consumers would opt for basic coverage.

¹⁶Here and elsewhere, we implicitly ignore the wealthiest tail of the wealth distribution. This would be problematic for us if the innovator 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.

¹⁷And indeed, as we show below, all consumers will, in fact, opt to choose basic health insurance to gain access to services that otherwise would be unaffordable.

However, in our model Assumption 4 ensures that this “reverse Nyman effect” does not arise. Thus, Proposition 1 arises in part because of Nyman’s access motive but also in part because core medical services are sufficiently valuable that no consumer finds it worthwhile to skimp on basic insurance in order to free up wealth to afford the treatment in the event that it is needed. Given the variety of valuable products and services that are included in the package of basic medical services, we believe that this is as reasonable assumption.

Now, define $W_i \equiv \widehat{W}_i - \rho_C C$ to be the net-of-basic-insurance wealth of individual i (or simply, wealth) and correspondingly, let $\underline{W} = \widehat{W} - \rho_C C$ be the lower bound on wealth. Thus, $W_i \sim \text{UNIFORM}[\underline{W}, \underline{W} + 1]$. Note that Assumption 3 implies that $\underline{W} > 0$. Now, we make a further assumption:

Assumption 5 $V_i = v(W_i) = \underline{V} + \alpha(W_i - \underline{W})$, where $\alpha \in (0, 1)$, and $\underline{V} + z\alpha > \underline{W} + z$ for all $z \in [0, 1]$.

Given the distribution of W_i , it follows that $V_i \sim \text{UNIFORM}[\underline{V}, \underline{V} + \alpha]$. The parameter α captures the extent to which the valuation of the treatment increases in wealth. Given Viscusi and Aldy’s (2003) finding that the income elasticity of the value of a statistical life is approximately 0.5 to 0.6, $\alpha < 1$ is a natural assumption.¹⁸

In light of Proposition 1, the wealth distribution determines the demand curve for the treatment, which in inverse form is

$$P_0(Q) = \underline{W} + 1 - Q, \quad Q \in [0, 1]. \quad (4)$$

In the absence of a liquidity constraint on out-of-pocket purchases of the treatment, a consumer would purchase the treatment as long as $V_i \geq P$. Thus, the (inverse) non-liquidity-constrained demand curve for the treatment is

$$P_1(Q) = \underline{V} + \alpha - \alpha Q, \quad Q \in [0, 1]. \quad (5)$$

Figure 2 illustrates that the liquidity constrained demand curve D_0 lies everywhere to the left of the non-liquidity constrained schedule D_1 .

2.1.2 Pricing the treatment: treatment is not covered by health insurance

The innovator’s profit-maximization problem is

$$\max_{Q \in [0, 1]} \pi_0(Q) = P_0(Q)Q,$$

and its solution is

$$Q^0 = \begin{cases} 1 & \text{if } \underline{W} \geq 1 \\ \frac{\underline{W} + 1}{2} & \text{if } \underline{W} \leq 1 \end{cases}. \quad (6)$$

$$P^0 = \begin{cases} \underline{W} & \text{if } \underline{W} \geq 1 \\ \frac{\underline{W} + 1}{2} & \text{if } \underline{W} \leq 1 \end{cases}. \quad (7)$$

Under both basic and expanded insurance, *ex ante* total surplus when Q consumers obtain the treatment (when needed) and Y consumers obtain core medical services (when needed) can be shown to equal

$$\widehat{W} + \frac{1}{2} + \rho_C(B - C)Y + \rho_I \left[(\underline{V} + \alpha)Q - \frac{\alpha}{2}Q^2 \right]. \quad (8)$$

¹⁸However, as we demonstrate in the Online Appendix to this paper, our results on how insurance coverage of the treatment affects price and consumer welfare continue to hold if we allow $\alpha \geq 1$.

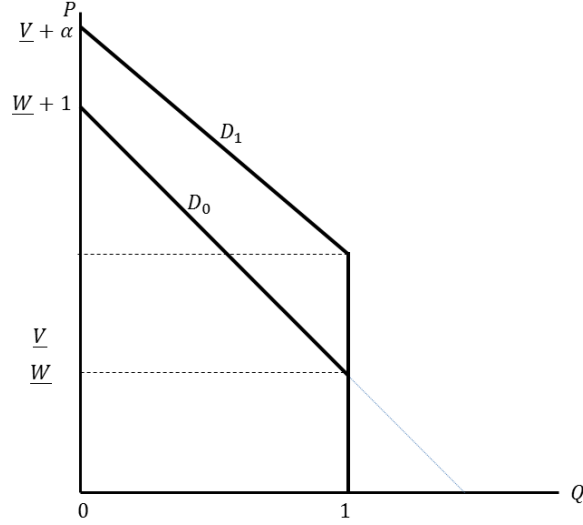


Figure 2: Demand curves with (D_0) and without (D_1) liquidity constraints

In what follows it will be convenient to measure surplus using a linear transformation of the expression in (8), constructed by dividing it by ρ_I and dropping the constant term $\frac{\widehat{W} + \frac{1}{2}}{\rho_I}$:

$$TS(Q, Y) = \frac{\rho_C(B - C)}{\rho_I} Y + (\underline{V} + \alpha) Q - \frac{\alpha}{2} Q^2. \quad (9)$$

The expression in (9) is the *increment* to total surplus from the consumption of health care services, but *reweighted* from the perspective of the state of the world in which a consumer becomes seriously ill. For simplicity, we refer to this simply as total surplus. The corresponding expression for consumer surplus when the price of the treatment is P is

$$CS(Q, Y, P) = \frac{\rho_C(B - C)}{\rho_I} Y + (\underline{V} + \alpha - P) Q - \frac{\alpha}{2} Q^2, \quad (10)$$

where it is understood that Q is the quantity of the treatment demanded at price P .

The first-best outcome maximizes $TS(Q, Y)$ for $Q \in [0, 1]$ and $Y \in [0, 1]$ and is given by $Y^F = Q^F = 1$. From Proposition 1, all consumers purchase basic health insurance and thus $Y^0 = 1$. When the innovator serves the entire market, $Q^0 = 1$, the first-best social surplus is attained. By contrast, when $Q^0 < 1$ —which arises when consumer wealth levels are sufficiently small—we have a deadweight loss (DWL) represented by area $ABCE$ in Figure 3.

2.2 Case 2: Treatment is covered by health insurance

2.2.1 Consumer choice and the demand for the treatment under expanded health insurance

Expanded health insurance covers a bundle consisting of core medical services and the treatment. Because the price of the expanded plan, $\rho_C C + \rho_I P$, is actuarially fair and consumers are risk neutral, there is no economic difference between an insured consumer who pre-pays $\rho_I P$ for the treatment portion of his health

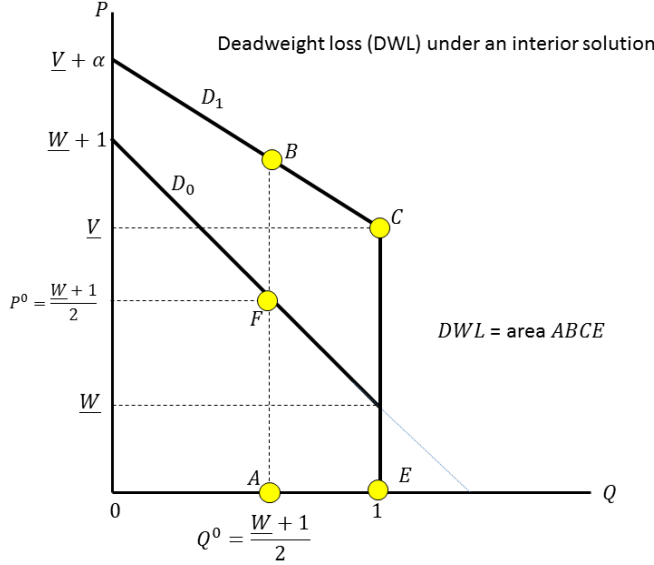


Figure 3: When the innovator’s profit-maximization problem has an interior solution, the deadweight loss under basic insurance is area $ABCE$.

insurance plan, which only delivers a benefit when he becomes ill, and an uninsured consumer who pays P out-of-pocket only when he becomes ill.¹⁹ Thus, we will talk as if P is the price of the treatment paid by insured consumers.²⁰

In the Appendix we derive the demand curve for the treatment. One component of treatment demand, which we label $D_u(P)$, captures the demand from individuals who cannot afford insurance but could afford the treatment, i.e., $P - \rho_C C < W_i < \rho_I P$. In the Appendix we show that if $\underline{W} \geq \frac{\rho_C \rho_I}{1 - \rho_I} C$, there can be no uninsured buyers of the treatment at any price, i.e., $D_u(P) = 0$ for all P . In other words, demand for the treatment and demand for expanded insurance are one in the same. Given that $\frac{\rho_C \rho_I}{1 - \rho_I} < 1$, this is a weak condition and, indeed as $\rho_I \rightarrow 0$ it reduces to our earlier assumption that $\underline{W} \geq 0$. For now we assume that $\underline{W} \geq \frac{\rho_C \rho_I}{1 - \rho_I} C$ and we ignore this component in what follows.²¹

The main component of treatment demand, which we label $D^*(P)$, is the demand for the treatment that comes from consumers who purchase expanded insurance. This component is made up of two parts which are depicted in Figure 4. The first part, labeled D_L^* in panel A of Figure 4, represents the demand for insurance if the treatment had infinite value, so that the only factor limiting insurance demand is liquidity. We call this the *liquidity constraint schedule*. Along this schedule, consumers purchase insurance (and thus the treatment) only if $P < \frac{W_i}{\rho_I}$, and thus D_L^* has the equation (in inverse demand form)

$$\frac{\underline{W} + 1 - Q}{\rho_I}.$$

The second part, labeled D_I^* in panel B of Figure 4, represents the demand for the insurance if there were no

¹⁹Of course—and this a key point of our analysis—liquidity constraints may affect the feasibility of either of these choices.

²⁰This economic equivalence between pre-paying ρP for the treatment and paying P out-of-pocket for it only in the state of the world when it is needed implies that in our welfare analysis below we can compute the consumer surplus from purchasing the treatment exactly as we did when the treatment is not covered, i.e., using (10).

²¹In the Online Appendix, we discuss what happens when $\underline{W} < \frac{\rho_C \rho_I}{1 - \rho_I} C$ and thus $D_u(P) > 0$ for some P . Using numerical analysis, we find that for these parameterizations, the main results reported below continue to hold.

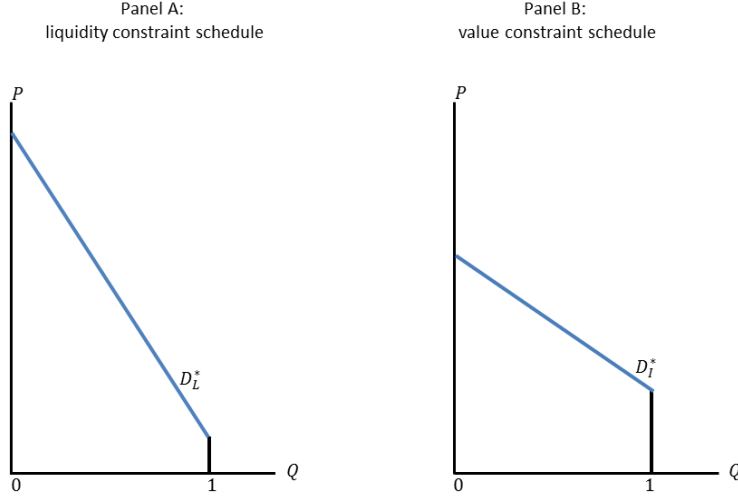


Figure 4: The demand curve for treatment under expanded insurance has two parts. The first part, in panel A, is the liquidity constraint. The second part, in panel B, is the value constraint.

liquidity constraint, so that the only limiting factor is the value of the bundle covered by insurance. We call this the *value constraint schedule*. This schedule traces out consumers' inclusive valuations of the insurance bundle, $V_i + \frac{\rho_C(B-C)}{\rho_I}$. Along this schedule, consumers purchase insurance (and thus the treatment) only if the treatment price P is less than this inclusive value, and thus D_I^* is given by

$$\underline{V}' + \alpha - \alpha Q,$$

where

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

Figure 5 depicts the three potential resulting demand curves for the treatment. In panel A, D_L^* lies everywhere below D_I^* on the quantity interval between 0 and 1, and therefore $D^*(P) = D_L^*$. In panel B, D_I^* lies everywhere below D_L^* and therefore $D^*(P) = D_I^*$. In panel C, D_L^* intersects D_I^* at a quantity strictly between 0 and 1.²² 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, $D^*(P)$ represents the lower envelope of D_L^* and D_I^* . Expressing $D^*(P)$ in inverse form, the demand curve in panel C is given by:²³

$$P^*(Q) = \min \left\{ \frac{W + 1 - Q}{\rho_I}, \underline{V}' + \alpha - \alpha Q \right\}, Q \in [0, 1],$$

²²Expressions for the price P_K and quantity Q_K corresponding to the kink in D^* are derived in the Appendix. Note that panel C of Figure 5 depicts the case in which $D^*(P)$ consists of D_L^* for low prices and D_I^* for high prices. It is straightforward to establish that given $\alpha < 1$ this must be the case. If, by contrast, $\alpha \geq 1$ (the case we consider in the Online Appendix), $D^*(P)$ could consist of D_L^* for high prices and D_I^* for low prices. However, this case would arise only if $\rho_I > \frac{1}{\alpha}$, which would require that α is very large (which seems empirically implausible) or that ρ_I is large (which is at odds with the focus in this paper on treatments of fairly rare medical conditions).

²³In the Appendix, we present the non-inverse formula for $D^*(P)$.

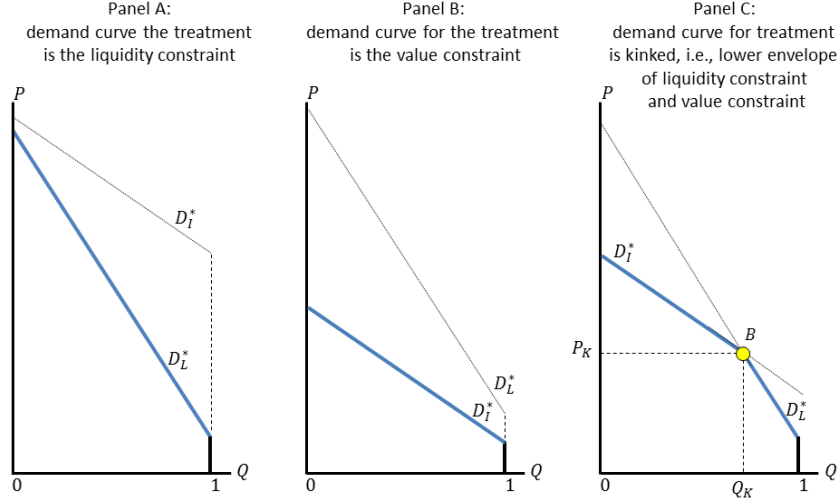


Figure 5: The demand curve for the treatment could be D_L^* , D_I^* , or a kinked demand curve given by the lower envelope of D_L^* and D_I^* .

In general, we cannot rule out any of the cases depicted in Figure 5. However, our model is intended to be empirically relevant for treatments for life threatening but relatively rare medical conditions, such as Hepatitis C, and we now argue that in such cases the most plausible scenario is that the demand curve for the treatment coincides with the value constraint.²⁴ Specifically, note that the vertical distance between the liquidity constraint and the value constraint is

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

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,

$$0 < \rho_I < \frac{W - \rho_C(B-C)}{V}. \quad (11)$$

Thus, when the illness requiring the treatment is sufficiently rare and wealth is sufficiently large (so that $W - \rho_C(B-C) > 0$ and thus (11) is not vacuous), consumer liquidity constraints will not limit the purchase of expanded insurance.²⁵ This in turn implies that, in contrast to the case of basic insurance, consumer liquidity constraints will not limit the purchase of the treatment either, i.e., *inclusion of the treatment in*

²⁴The incidence of Hepatitis C virus (HCV) in the U.S. population is quite low, with the Centers for Disease Control and Prevention estimating that 29,713 new cases occurred in 2013, with 2.7 million Americans suffering from chronic HCV infection. Centers for Disease Control and Prevention, "Hepatitis C FAQs for Health Professionals," <http://www.cdc.gov/hepatitis/hcv/hcvfaq.htm> (accessed July 7, 2015). As another example, the incidence of lung cancer, though about 7 times larger than HCV, is also quite low, with 207,339 diagnosed with lung cancer in 2011, an incidence of about 0.006 percent. Centers for Disease Control and Prevention, "Lung Cancer Statistics," <http://www.cdc.gov/cancer/lung/statistics/> (accessed July 7, 2015).

²⁵As we will see in the numerical example below, the condition $W - \rho_C(B-C) > 0$ —which requires that consumer liquid wealth be sufficiently large or the likelihood of requiring core medical services not be "too high"—does not contradict either Assumption 2 or Assumption 4. Moreover, if it holds then Assumption 3 holds as well.

the health insurance plan “breaks” consumer liquidity constraints. In effect, we are focusing on a setting in which liquidity *does not* constrain the payment of *insurance premiums* (either basic or expanded), but it would constrain out-of-pocket payments for the treatment when the treatment is not covered. Though this is not the only scenario of potential interest, it is one that is both empirically plausible and (importantly for the focus of this paper) consistent with Nyman’s (1999) model of health insurance.²⁶

2.2.2 Pricing the treatment under expanded health insurance

The innovator’s pricing problem is

$$\max_{Q \in [0,1]} \pi^*(Q) = (\underline{V}' + \alpha - \alpha Q) Q. \quad (12)$$

In the Appendix we establish that, given Assumptions 1, 4, and 5, the solution to this problem—denoted by Q^* —is a corner solution in which the treatment is sold to all consumers, i.e., $Q^* = 1$. The corresponding profit-maximizing price in this case is thus $P^* = \underline{V}'$. Because there are no uninsured consumers of the treatment, the innovator’s profit-maximization solution implies that all consumers purchase insurance and thus all consumers obtain core medical services when needed, i.e., $Y^* = 1$.

Why, given our parameter assumptions, does the innovator price the treatment so as to serve the entire market? To answer this question, let’s contrast the innovator’s problem under expanded insurance to that under basic insurance. When the treatment is covered, its demand curve $P^*(Q)$ lies everywhere to the right of the demand curve $P_0(Q)$ when the treatment is not covered. (Formally, this is implied by Assumption 1). That, coupled with $\alpha < 1$ (via Assumption 5), ensures that $P^*(Q)$ is more price elastic than $P_0(Q)$ for any $Q \in [0, 1]$. Thus, in general, coverage of the treatment must result in a higher profit-maximizing quantity of the treatment than if it is not covered. Whenever the profit-maximizing solution in the latter case is to serve the entire market, the same will necessarily be true when the treatment is covered. Beyond this, though, Assumption 4—which played a critical role in driving the result of full insurance coverage under basic insurance (i.e., Proposition 1)—ensures that the inclusive value of the insurance bundle for a consumer with even the lowest level of wealth is fairly “large.” The upshot is that the coverage of the treatment shifts its demand curve sufficiently rightward that it makes sense for the innovator to serve the entire market even in circumstances in which it would not have had the treatment not been covered.

Because \underline{V}' equals 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 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 2 *Suppose (11) holds so that inclusion of the treatment in the health insurance plan breaks consumers’ liquidity constraints. 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., $P^* > v(\widehat{W}_i)$ for all $\widehat{W}_i \in [\widehat{W}, \widehat{W} + 1]$.*

²⁶In the Online Appendix, we consider the case in which $D^* = D_L$ for all $Q \in [0, 1]$, as well as the more general case in which the kink in the demand curve becomes potentially relevant. We demonstrate that our results on how insurance coverage of the treatment affects consumer welfare continue to hold in these cases (both with $\alpha < 1$ and $\alpha \geq 1$ but sufficiently small so that $\alpha < \frac{1}{\rho_I}$).

	Basic insurance (\$ 00,000)	Expanded insurance (\$ 00,000)	% 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%

Table 1: Comparison of outcomes under basic and expanded insurance

Proposition 2 provides an economic rationale for the common claim that firms in the pharmaceutical sector set prices that exceed the value created by their products (Loftus, 2015; Walker, 2015). In Section 6 we present empirical evidence consistent with the implication of Proposition 2.

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:

- $C = 1.50$ (This is roughly the midpoint of estimates of the cost of coronary artery bypass surgery in the U.S. in the early 2010s.)²⁷
- $B = 2.25$ (This implies a benefit-cost ratio for core medical services of 1.5 to 1.)
- $\widehat{W} = 0.41$ (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.)
- $\widehat{W} + 1 = 1.41$ (This is roughly the 90th percentile for household income in the U.S. in the 2010s.)²⁸
- $\underline{V} = 1.025$ (This assumes that the value of the treatment for the highest-income household is 2.5 times as large as that household’s wealth.)
- $\rho_C = 0.06$ 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.)²⁹
- $\alpha = 0.89$ (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 $\frac{dv(\widehat{W}_i)}{d\widehat{W}_i} \frac{\widehat{W}_i}{v(\widehat{W}_i)}$ equal to 0.55 when evaluated at the mid-point of the wealth distribution, i.e., about \$92,000).

These parameter values satisfy Assumptions 1-4. They also satisfy $\underline{W} \geq \frac{\rho_C \rho_I}{1 - \rho_I} C$ (so there are no uninsured buyers of the treatment) and condition (11) (so demand for health insurance is value constrained).

Table 1 summarizes the outcomes under basic and expanded insurance for this example. 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 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

²⁷ Coronary Artery Bypass Patient Education, Frequently Asked Questions, <http://cabggroupproject.web.unc.edu/frequently-asked-questions/> (accessed August 25, 2015).

²⁸ Capping the income data in this way implicitly 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.

²⁹ 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.

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 3 *Suppose (11) holds so that inclusion of the treatment in the health insurance plan breaks consumers' liquidity constraints. Then $P^* > P^0$ i.e., the price of the treatment when it is covered by insurance is greater than the price when it is not covered. When $\underline{W} \geq 1$, $Q^* = Q^0$, and when $\underline{W} < 1$, $Q^* > Q^0$, 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 under when consumer wealth is sufficiently low.*

As noted above, under expanded insurance all consumers receive the treatment and core medical services when needed. For this reason, total surplus TS^* under expanded insurance must be at least as large as total surplus TS^0 under basic insurance:

$$\begin{aligned} TS^* - TS^0 &= TS(Q^*, Y^*) - TS(Q^0, Y^0) \\ &= TS(1, 1) - TS(Q^0, 1) \geq 0, \end{aligned}$$

since $Q^0 \leq Q^* = 1$, and $\frac{\partial TS(Q, Y)}{\partial Q} = \underline{V} + \alpha - \alpha Q > 0$.

However, the impact of expanded insurance on consumer welfare is less obvious: while the quantity of the treatment is at least as large under expanded insurance as it is under basic insurance, the price of the treatment is strictly higher. Substituting (6) and (7) into (10) yields the consumer surplus when the treatment is not covered:

$$\begin{aligned} CS^0 &= CS(Q^0, Y^0, P^0) \\ &= \frac{\rho_C(B - C)}{\rho_I} + \begin{cases} \underline{V} + \frac{\alpha}{2} - \underline{W} & \underline{W} \geq 1 \\ \left(\frac{\underline{W}+1}{2}\right) \left(\frac{\underline{V} + \frac{3\alpha}{4}}{-\frac{\alpha\underline{W}}{4} - \frac{\underline{W}}{2} - \frac{1}{2}}\right) & \underline{W} \leq 1 \end{cases}, \end{aligned} \quad (13)$$

By contrast, consumer surplus when the treatment is covered is

$$CS^* = CS(1, 1, \underline{V}') = \frac{\alpha}{2}. \quad (14)$$

We can establish that consumers fare worse under expanded insurance than under basic insurance:

Proposition 4 *Suppose (11) holds so that inclusion of the treatment in the health insurance plan breaks consumers' liquidity constraints. Then $CS^* < CS^0$, i.e., consumer surplus is lower when the treatment is covered by insurance than when it is not.*

Table 2 summarizes the welfare effects of expanded insurance compared to basic insurance. When $\underline{W} < 1$ the quantity of the treatment purchased is higher under expanded insurance, but this offset by the negative effect on consumers of a higher price. As we see from Table 1 this impact can be quite large: consumer surplus falls by 85 percent under expanded insurance, driven by a 396 percent increase in the price of the treatment.

Just because aggregate consumer surplus falls does not mean that *all* consumers are worse off. When $\underline{W} < 1$, there will be consumers who do not purchase the treatment were it not covered. These consumers get health insurance in both cases, so they obtain core medical services when needed, and when the treatment

	Quantity of treatment, Q	Price, P	Quantity of core medical service, Y	Total surplus, TS	Consumer surplus, CS
$\underline{W} \geq 1$	$Q^0 = Q^* = 1$	Higher	Same	Same	Lower
$\underline{W} < 1$	$Q^0 < Q^* = 1$	Higher	Same	Higher	Lower

Table 2: Comparison of price, quantity of treatment, quantity of basic medical services, consumer surplus, and total surplus under basic and expanded insurance

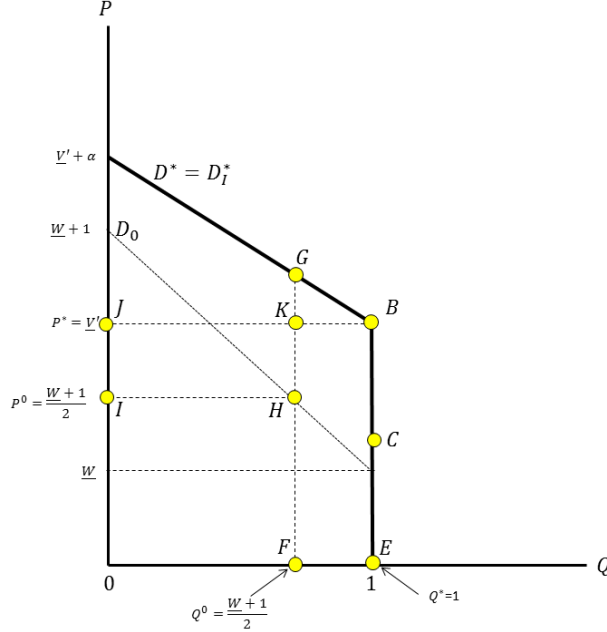


Figure 6: Consumers with inclusive valuations located along line segment GB are better off when the treatment is covered by insurance, even though consumers are worse off overall.

is covered, they also obtain the treatment when needed. Figure 6 illustrates this. Consumers with inclusive valuations on line segment GB would not have bought the treatment were it not insured, but when the treatment is insured they can afford to purchase health insurance and accordingly obtain the treatment when needed. The additional consumer surplus to these consumers is area KGB . However, this gain is more than offset by the loss of surplus $IJKH$ by consumers who would have purchased the treatment out-of-pocket if it was not insured. Still, it *is* possible that coverage of the treatment could make all consumers worse off. As noted in Proposition 3, when $\underline{W} \geq 1$ the number of consumers obtaining the treatment is the same as it is when the treatment is not covered, but each consumer pays a higher price.

This discussion brings into sharp focus the distributional consequences of covering the treatment: it transfers surplus from consumers with the highest valuations of the treatment—i.e., those with the highest wealth—to the innovator, and potentially—when consumer wealth is sufficiently low—to consumers with lowest valuations of the treatment and the lowest wealth.

3 Multiple innovators

Health insurance typically covers a host of services, including many innovative pharmaceutical products. In this section, we consider the implications of multiple innovators, and we show that our analysis assuming a single innovator puts 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, expanded insurance covering multiple innovative treatments can result in lower total surplus than basic insurance.

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_i to consumer i should that state occur (and no value should any other state occur). Given this specification, the case of $N = 1$ reduces to the model above. The assumption that the treatment provides consumer i 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_i is the value to consumer i of prolonging his or her life.

Suppose the basic insurance plan covers none of the N treatments. In that case, the consumer 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 for the single-innovator case studied in the previous section, and as a result, under the basic insurance plan, each treatment's innovator sets the same price P^0 that a single innovator would set.

If, on the other hand, expanded insurance covers each of the N treatments, then given actuarially fair insurance pricing, the net value of purchasing expanded insurance for consumer i is

$$EU_i^I = \widehat{W}_i + \rho_C(B - C) + \rho_I(V_i - P_A),$$

where $P_A = \frac{\sum_{i=1}^N P_i}{N}$ is the *average price* of all treatments, $P_A = \frac{\sum_{i=1}^N P_i}{N}$. Continuing to assume that there are no uninsured buyers of treatments, the demand curve for each individual treatment is $D^*(P_A)$, which is, as above, the lower envelope of the liquidity constraint and the value constraint. Throughout this section, we continue to focus on the case in which (11) holds, and thus the demand curve for each individual treatment is the value constraint.

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 $P_1^* = \underline{V}'$.) 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.³⁰

Proposition 5 *Suppose (11) holds so that inclusion of the treatment in the health insurance plan breaks consumers' liquidity constraints. There exists a critical number of innovators $N^* > 1$ such that for $N \leq N^*$, $P_N^* = P_1^* > P^0$, and for $N > N^*$, $P_N^* = \frac{N}{N+1}(\underline{V}' + \alpha) > P_1^* > P^0$, i.e., when expanded 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 they are covered by insurance is greater than the price when they 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., $Q_N^* = 1$. However, when $N > N^*$, the corresponding quantity is

$$Q_N^* = D^*(P_N^*) = \frac{1}{\alpha} \frac{\underline{V}' + \alpha}{N + 1} < 1,$$

³⁰As just noted, this proposition pertains to the case in which the demand curve for the treatment coincides with the value constraint. In the Online Appendix, 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.

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 expanded 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 $N \rightarrow \infty$, 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 has essentially given rise to Cournot’s complementary monopoly problem (Cournot 1927, Chapter IX). Though the model just sketched is special, the intuition is broader. Each innovator, when considering raising the price of its treatment, raises the price of (actuarially fair) insurance by just a fraction of the increment to the treatment price, but it decreases the demand not just for its own treatment for all other services that the insurance bundle comprises. Thus, each innovator perceives itself facing significant “pricing room,” and the bigger the insurance bundle the more “room” it perceives. By raising 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.³¹ Consequently, in light of Proposition 4, expanded insurance covering N treatments results in a lower consumer surplus than basic insurance.

Proposition 6 *Suppose (11) holds so that inclusion of the treatment in the health insurance plan breaks consumers’ liquidity constraints. Then $CS_N^* < CS^0$, i.e., consumer surplus under expanded insurance that covers multiple treatments is less than consumer surplus under basic insurance.*

Propositions 5 and 6 imply that our main 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 of 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. In addition, these results suggest that as the concentration of firms in the innovative portion of the pharmaceutical market changes this should directly affect the prices charged for these products, with a lower concentration resulting in higher prices.

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 7 *Suppose (11) holds so that inclusion of the treatment in the health insurance plan breaks consumers’ liquidity constraints. 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 expanded insurance that covers multiple treatments is less than total surplus under basic insurance.*

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. 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 so much that (for N large enough) some individuals decide to forego expanded

³¹This is formally demonstrated in the proof of Proposition 6.

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 Do the consumption smoothing benefits of insurance outweigh the drop in consumer surplus under expanded insurance?

We have shown that consumer surplus under expanded insurance is less than consumer surplus under basic insurance. But so far we have neglected the consumption smoothing benefits of insurance. Perhaps the ability to better smooth consumption offsets the higher price that consumers effectively pay for the treatment under expanded insurance.

To explore this issue, we return to the single-innovator model and consider a simple specification in which the benefit of consumption smoothing is given by the extent to which insurance reduces the variance in consumption spending $z_i = z(\widehat{W}_i)$. Under expanded insurance, the variance of consumption spending for an individual consumer is zero: either a consumer obtains insurance and has a consumption level of $\widehat{W}_i - \rho P^*$ in each state or does not obtain insurance and has consumption spending \widehat{W}_i in each state.³² By contrast, under basic insurance it is straightforward to show

$$Var(z_i) = \begin{cases} 0 & x_i = 0 \\ (1 - \rho_I) \rho_I (P^0)^2 & x_i = 1 \end{cases} .$$

That is, those who do not purchase the treatment have no variation in consumption spending (it is $W_i - \rho_C C$ in all states), but for those that do purchase the treatment, the variance in consumption spending is $(1 - \rho_I) \rho_I (P^0)^2$. Aggregating across all consumers, we have

$$\int Var(z_i(\widehat{W})) d\widehat{W} = (1 - \rho_I) \rho_I (P^0)^2 Q^0 .$$

Letting γ denote a coefficient of risk aversion, the consumption smoothing benefit of expanded insurance is thus $\gamma (1 - \rho_I) \rho_I (P^0)^2 Q^0$.

Using the numerical example in Table 1 as an illustrative example, the consumption smoothing benefit of expanded insurance is 0.0056γ , while the reduction in consumer surplus is approximately 2.44. This implies that coefficient of risk aversion would need to be about 436 for the consumption-smoothing benefits of expanded insurance to outweigh the reduction in consumer surplus, well outside the range of empirical estimates of coefficients of risk aversion.

Though the example is specific, the intuition is robust. The reduction in consumer surplus under expanded insurance tends to be greatest when the probability ρ_I of the serious illness is smallest, for it is under these conditions that the innovator enjoys the greatest degree of pricing power. But extending insurance to a rare event smooths consumption only a little. This can be seen explicitly in the calculations above, but the insight would hold for a more general concave utility function.³³

³²Recall in our model the implication of not having insurance in our model is the *loss* of a benefit B in state 1 and V_i in state 2. These can be thought of as avoided losses of income, and they are already counted as part of our consumer surplus calculation.

³³A more complete analysis would take into account that the risk aversion of consumers would alter the profit-maximizing price of insurance. But including that effect in our analysis would almost certainly strengthen the claim because risk aversion would strengthen the demand for expanded insurance, thus tending to increase P^* without changing P^0 , which would still be determined by liquidity constraints.

5 Minimum coverage standards for health insurance

In the preceding analysis, we showed that coverage of the innovative treatment increases price and reduces consumer surplus relative to the case in which the treatment is not covered. We can think of this analysis as pertaining to a setting in which a treatment has recently been discovered. Our analysis sheds light on whether the inclusion of a newly discovered treatment in health insurance formularies makes consumers better off when the innovator can exercise market power in the pricing of the treatment.

But we can use our model to study an additional issue. Suppose that coverage of the treatment arises because insurers are mandated to do so through a minimum coverage standard (MCS). To study the impact of these standards on consumers, the appropriate benchmark is not the case in which the treatment is not covered by insurance. Rather, it is a setting in which consumers have the choice between a basic insurance plan that does not cover the treatment or an expanded plan that does. A minimum coverage standard that requires coverage of the treatment would rule out the choice of the basic plan, which means that insurance bundles together the treatment and core medical services. By contrast, in the absence of a minimum standard, the treatment and core medical services are unbundled.

Having already studied pricing of the treatment under expanded insurance, we now consider pricing of the treatment when consumers can choose between the basic health insurance plan, the expanded health insurance plan, or no insurance at all. Throughout we assume that the innovator sets a uniform price P —i.e., it does not discriminate between the price it charges a basic plan or an expanded plan, and the innovator is assumed to set that price before consumers make their plan choice. We also focus on the single-innovator case and maintain Assumptions 1-5.

Recall that in Proposition 1, we showed that when faced with a choice between the basic health insurance plan and no insurance, all consumers choose the basic health insurance plan. Thus, the choice between the basic health insurance plan, the expanded health insurance plan, or no insurance boils down to a choice between the first two. For any given price P of the treatment, a consumer will choose the expanded plan if $EU_i^I \geq EU_i^{BI}$, where EU_i^{BI} and EU_i^I are given in (2) and (3).³⁴ Since

$$EU_i^I - EU_i^{BI} = \begin{cases} 0 & W_i \geq P \\ V_i - P & W_i < P \end{cases}, \quad (15)$$

consumers who could have afforded the treatment with just basic insurance are indifferent between basic and expanded insurance, and consumers who could not have afforded the treatment when covered by basic insurance but who value it strictly prefer expanded insurance to basic insurance. Demand for expanded insurance will come from these groups of consumers provided that they can afford it. That is, demand for expanded health insurance comes from consumers whose (W_i, V_i) is such that

$$W_i \geq P \text{ or } \{W_i < P \text{ and } V_i - P \geq 0\}$$

and

$$W_i \geq \rho_I P.$$

Because $V_i > W_i$, the set of consumers for whom

$$W_i \geq P \text{ or } \{W_i < P \text{ and } V_i - P \geq 0\}$$

³⁴For simplicity, we assume that if a consumer is indifferent between the two options, it chooses expanded insurance. This is actually without loss of generality. The demand for the treatment ends up being the same if indifferent consumers purchase basic rather than expanded insurance.

is merely the set of consumers for whom $V_i - P \geq 0$.³⁵ Thus, the demand for expanded insurance comes from those for whom

$$V_i - P \geq 0 \text{ and } W_i \geq \rho_I P. \quad (16)$$

The demand for the *treatment* in this case must be the same as the demand for expanded insurance. This is because a consumer does not purchase expanded insurance because either $V_i < P$ —the consumer does not value the treatment—or because $W_i < \rho_I P$ —the consumer cannot afford expanded insurance, which in turn implies that $W_i < P$, so the consumer cannot afford the treatment given that basic insurance has been purchased.

The demand curve for the treatment is thus the set of consumers for whom (16) holds. This is just the lower envelope of the liquidity constraint D_L^* schedule and the value constraint D_I^* schedule but with the latter evaluated at $\rho_C = 0$ (or, in other words, the demand curve D_1 defined in (5) in Section 2.1.2). In inverse form, the demand curve for the treatment is given by

$$\widehat{P}(Q) = \min \left\{ \frac{W + 1 - Q}{\rho_I}, V + \alpha - \alpha Q \right\}, Q \in [0, 1].$$

The treatment demand curve has the same form as it does when the consumers only had the choice between expanded insurance and no insurance, except that the “piece” corresponding to the value constraint contains reflects just the incremental value of the treatment as opposed to the inclusive value of the expanded insurance bundle. This makes sense: when we unbundle insurance, the demand for the treatment is effectively determined at the margin, i.e., by its own incremental value.

The analysis of treatment pricing when minimum coverage standards are absent is identical to our earlier analysis, but with $\rho_C = 0$. Reducing ρ_C in our earlier analysis has the effect of shifting the value constraint D_I^* downward, which in turn makes it more likely that treatment pricing is value constrained rather than liquidity constrained. This enables us to prove:

Proposition 8 *Let \widehat{P} be the profit-maximizing price of the treatment when consumers can choose from a basic health insurance plan and an expanded health insurance plan, and let P^* be the profit-maximizing price of the treatment under MCS when consumers’ only insurance option is the expanded plan. Finally, suppose (as above), $\rho_I < \frac{1}{\alpha}$. Then $\widehat{P} \leq P^*$, with the inequality holding strictly if under the MCS the profit-maximizing price occurs along the value constraint. That is, a MCS mandating that the treatment be covered cannot decrease the price of the treatment and may increase it.*

Because consumer choice is more limited under MCS and because the price of the treatment cannot go down (and may go up), an immediate implication is that these features of an MCS reduces consumer welfare. That being said, attempting to sell unbundled insurance products that each cover particular conditions or sets of conditions may lead to a series of adversely selected risk pools. Therefore, when considering the welfare effects of an MCS we must weigh the potential efficiency losses from this adverse selection against the lost consumer welfare and higher treatment prices that exist in the bundled market. To our knowledge, our paper is the first highlight a mechanism for consumer surplus losses from these coverage standards.

³⁵The proof is as follows. If (W_i, V_i) is such that

$$W_i \geq P \text{ or } \{W_i < P \text{ and } V_i - P \geq 0\},$$

then it follows that $V_i - P \geq 0$. Conversely, if $V_i - P \geq 0$, then there are one of two possibilities: either $W_i \geq P$ or $W_i < P$. If the latter holds then for that set of consumers we have $\{W_i < P \text{ and } V_i - P \geq 0\}$.

6 The Role of Expanded Insurance on Drug Prices: Evidence from Medicare Part D

Two important predictions of our model are that the existence of health insurance increases the profit-maximizing price and that the bundling of products into a single insurance bundle allows the monopoly producer of an innovative product to capture more value than its product creates. Due to the widespread adoption of health insurance, empirically documenting these effects is difficult. We attempt to overcome these difficulties by focusing on the creation of Medicare Part D, which was the first widespread social pharmaceutical insurance program for the elderly in the United States. Prior to this program, the only social insurance for pharmaceuticals for seniors covered products administered in a medical setting such as a physician’s office or outpatient facility. These drugs were covered under Medicare Part B and included many intravenous products.

To examine the implications of our model, we estimate the effect of Medicare Part D on the launch prices of oncology products—a category that has seen a relatively large number of innovative products released in the years before and after the creation of Part D. In addition, oncology is a category that contains many intravenous products covered by Part B prior to the creation of Part D. The enabling legislation for the creation of Medicare Part D was passed in late 2003 and seniors were first able to sign up for coverage as of 2006. Unlike the traditional Medicare program, coverage under Part D was provided by private insurers whose expenses were subsidized. Given that private insurers can make determinations about which drugs will be included on their formularies, the program contains a number of regulations regarding what drugs and drug classes must be covered by a plan. For example, Part D plans must provide coverage for all diseases for which there are existing treatments. In addition, for each drug class, an insurer must provide at least two options that are chemically distinct (unless there is only one existing option a class). For six particular protected classes (immunosuppressant, antidepressant, antipsychotic, anticonvulsant, antiretroviral, and antineoplastic products) the insurer must cover “all or substantially all” drugs in the class. In addition, they may not use any step therapy or prior authorization measures for patients that are already taking the drug with the burden of proof being on the insurer to show a patient was not previously taking the drug (Center for Medicare and Medicaid Services, 2014). Finally, products that were covered under Part B were intended to be broadly unaffected by the new coverage offered under Part D.

Several features of Part D provide an opportunity for empirically examining the implications of our model with respect to oncology (i.e. antineoplastic) products. First, the expansion increased the share of elderly Americans with drug coverage by 10 percentage points (Engelhart and Gruber, 2010). For these newly insured individuals, Part D broke the liquidity constraint for expensive drugs such as oncology products. Second, the creation of protected classes guaranteed that all new oncology products would be included in a bundle with products that are sold in more competitive markets. This is explicitly true for enrollees in Part D, but these new regulations also likely influenced coverage decisions in the private market where insurers often follow Medicare’s policies because of a variety of factors such as litigation risk (Bagley, Chandra, and Frakt, 2015).³⁶ This forced bundle allows us to examine the prediction that a firm producing an innovative product that is bundled with more competitively priced products can capture more than the value created by its product.

Duggan and Scott Morton (2010) examined the effect of Medicare Part D on all drug prices and found that the newly insured paid lower prices for all medications. The mechanism underlying this result is that private

³⁶For example, in May 2015 Anthem Blue Cross-Blue Shield was sued for denying a claim for an expensive hepatitis-C treatment that it deemed was not medically necessary given the sickness of the patient (Pfeiffer, 2015). Similarly, state Medicaid systems in several states face legal liability for their efforts to limit access to the same treatment (Kardish, 2014).

insurers could negotiate price discounts that the uninsured were unable to previously obtain. However, this price effect was quite heterogeneous and, consistent with the predictions of our model, in situations where insurers were either required (because of the protected class regulations) or effectively required (because of a small number of products in the class) to cover all products, the passage of Part D was estimated to either increase or have no effect on the average prices paid by newly insured consumers. We focus our analysis on the launch prices charged in one of these protected classes and implement a different identification strategy that exploits the fact that many oncology products previously had insurance coverage under Medicare’s Part B medical benefit. In contrast to previous work we also consider the survival-adjusted price of products which serves as an indicator of the amount of value captured by producers.

Given our interest in how the presence and structure of insurance programs impacts the prices charged by innovators (i.e. those products for which there are limited opportunities for consumer or insurer substitution) we focus our attention on oncology products that were required to be covered under Part D. In accordance with our model, we expect that the passage of Part D should have two effects on this market. First, there should be an increase in prices for products, such as oncology, that are used by individuals that are covered by Part D. Second, because of the forced bundling and protected class status the producers of oncology products should be better able to sell their products for more than the value that they create.

We begin by noting that prices have been rising in oncology in general over the last two decades. For example, Howard *et al.* (2015) documents that the inflation and survival-adjusted prices for newly released oncology products increased by 10 percent per year from 1995 – 2013. Using the same data from this study, we examine the relationship between the passage of Medicare Part D and rising survival-adjusted oncology prices.³⁷ These data contain the launch prices for 58 oncology products that were approved for sale by the FDA between 1995 and 2013 that were initially approved for cancer treatment and were intended for affecting survival rather than alleviating symptoms. In addition, drugs are only included in these data if there are clear estimates of their survival benefit – which is important for determining the value created by the product as well as a quality adjusted measure of prices. The launch price is “episode treatment price” which is the monthly price paid by Medicare multiplied by the typical treatment duration. More detail on these pricing data are available in Howard *et al.* (2015).

Figure 7 contains the resulting scatter plot and fitted curve from a locally weighted regression of price on the year of FDA approval for the oncology products in our sample. This shows that the price increase over the entire sample identified in Howard *et al.* (2015) is composed of relatively flat and stable trend from 1995 – 2003 and rapidly increasing price increases in the years following.

This pattern is consistent with our model’s prediction that profit-maximizing prices are higher when a greater share of potential customers have insurance. To further examine this relationship, we exploit the fact that not all oncology products were covered by the passage of Part D. Unlike other pharmaceutical categories, many oncology products are administered in a physician’s office or hospital setting. As a result, these products are covered (and have always been covered) under Medicare Part B.³⁸ Figure 8 contains the average launch price per year for oncology products based on whether they were covered by either Medicare Part B or Part D. The solid line represents the log of the average price for products that were covered (or eventually covered) by Part D. Between 1995 and the passage of Part D in 2003, there were 8 oncology products that were generally not covered under Part B but would eventually be covered under Part D.

³⁷We are grateful to Rena Conti for granting us permission to use these data.

³⁸Previous work has used the share spending for the disease category that is eligible for Medicare coverage. For most drugs, this is a good approximation of the effect of expanded Part D coverage. However, a disproportionately large fraction of oncology products are eligible for coverage under Part B. One strategy would be to estimate a triple difference model that exploits both the share of patients over 65 and the Part B coverage status. Given that there are only 58 oncology products released over the 18 year period in our sample, splitting the data along both of these dimensions is too demanding.

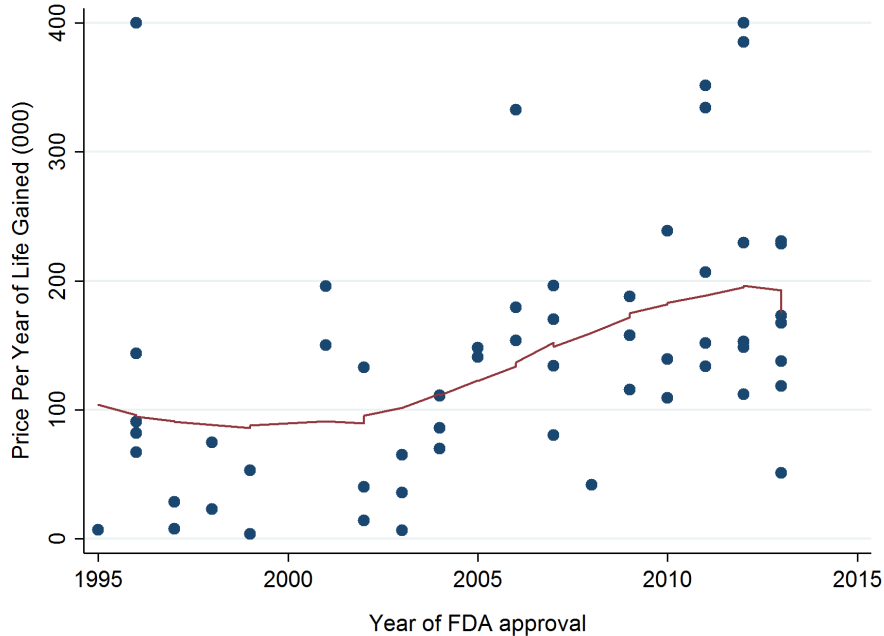


Figure 7: Prices of new oncology products, 1995-2015.

Overall, the prices for these products were far below the average prices of drugs of the 12 products released over this time that were covered by Part B—represented by the dashed line. In addition, prior to the passage of Part D the trend in prices is generally flat and similar across the two groups of drugs. The only exception to this pattern is 2001, the year in which Gleevec received FDA approval. Following the passage of Part D, there is an immediate sudden increase in the average launch price of Part D oncology products. As a result, after the 2006 implementation of Part D the two sets of products had very similar prices.

There could be a concern that the higher prices for products that were released after the passage of Part D simply reflect a technological shift in which oral chemotherapy products became more effective and thus the higher prices were a reflection of this greater value. Figure 9 contains the average prices per life year gained by Part D coverage status and year. Prior to 2003, the Part D products had an average (median) price per life year gained of approximately \$52,500 (\$11,000). In comparison, the intravenous oncology products covered by Part B that were released prior to 2004 had an average (median) life year gained of \$100,000 (\$70,700). Following the passage of Part D there was a marked increase in the prices for both sets of products. However, this increase was far greater for products that were covered by Part D and resulting in near parity in their survival adjusted launch prices. After 2004 the average (median) price per life year gained for Part D products was \$176,000 (\$153,000), and for Part B products it was \$172,000 (\$147,000).

These figures suggests two reactions to the passage of Medicare Part D. First, the manufacturers of oral chemotherapy products increased their prices. Second, these manufacturers were now able to command prices that exceeded many estimates of the value that they create. To more systematically examine these outcomes, we estimate the following difference-in-differences regression model:

$$y_{it} = \alpha + \gamma_t + \lambda \cdot X_i + \beta_1 \cdot I\{Part\ D\}_i + \beta_2 \cdot I\{Part\ D_i\} \cdot I\{PostPartD\}_t + \varepsilon_{it},$$

where y_{it} are a set of different measures of launch prices, γ_t are a series of year fixed effects, X_i are a series of product specific market characteristics, $I\{PartD\}_i$ is an indicator variable that is equal to 1 if the oncology

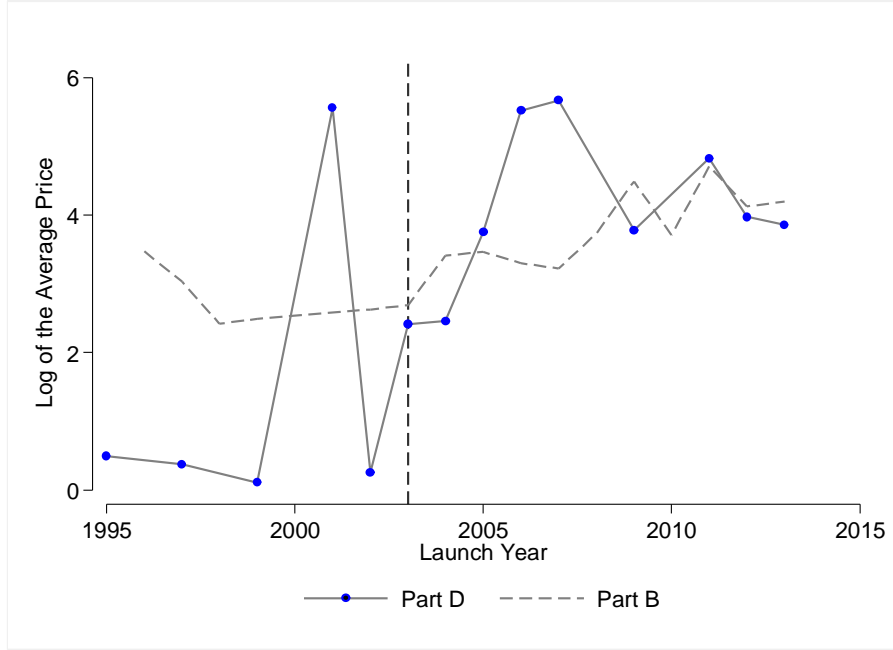


Figure 8: Average Launch Prices for Oncology Products by Medicare Coverage Status and Year, 1995-2013

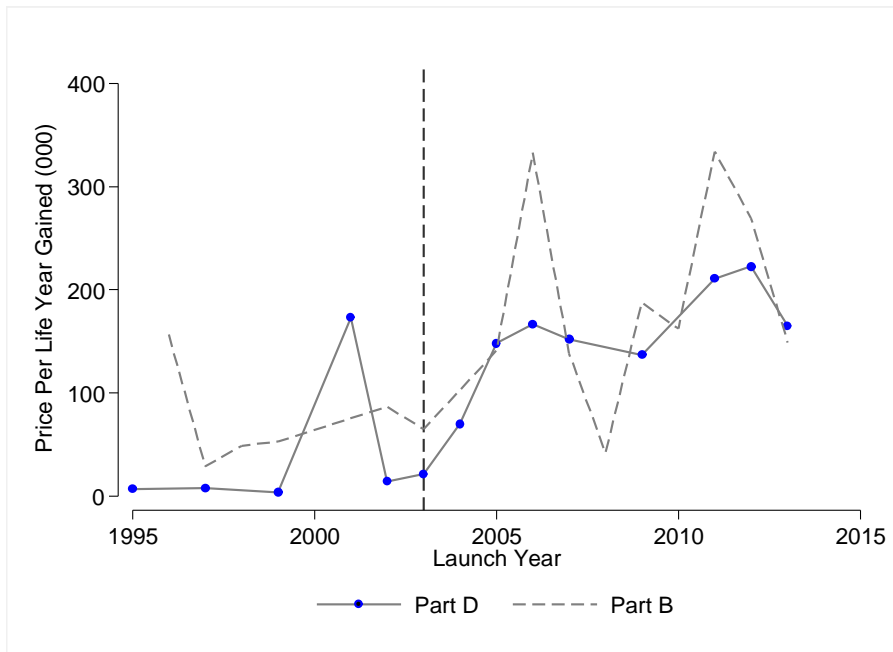


Figure 9: Average Launch Price Per Life Year Gained for Oncology Products by Medicare Coverage Status and Year, 1995-2013

product is covered by Part D, $I\{Post\ Part\ D\}_t$ is an indicator variable equal to 1 for years after 2003, and ε_{it} is an idiosyncratic error term. The coefficient of interest is β_2 which represents the change in the average prices for oncology products covered by Part D after the passage of the insurance expansion compared to oncology products covered by Part B.

Table 3 contains the estimates for a variety of spending outcomes. For each outcome we present the results from a model without any covariates, those from a model with product characteristics, and finally one with both product characteristics and disease fixed effects. The first three columns of Table 3 contain the estimates for the outcome of the log of the average price. Across all three specifications, the passage of Part D is associated with a large increase in the average launch price of oncology products that were covered by that program compared to drugs that were covered under Part B. For example, the coefficient in column (3) implies that the passage of Part D is associated with an approximately 420 percent increase in the average launch price of affected oncology products.

The coefficients in columns (4)-(6) of Table 3 estimate the change in price per life year gained in the years following the passage of Part D. The estimate in column (4) without any product or market covariates is roughly similar in magnitude to the estimate for the change in the log price in column (3). The coefficient in column (6) suggests that the passage of Part D was responsible for a 175 percent increase (p-value = 0.062) in the price per life year gained for covered products compare to products that were covered under Medicare Part B. While the estimated changes in both the average price and average price per life year gained are large, note that they are broadly in line with the estimate in our canonical example in Table 1.

We next examine the question of whether the Passage of Part D allowed manufacturers to capture more value than their products create. Conducting such an analysis requires defining the value created by these products. During the debate about the passage of Part D there was considerable discussion about the cost of the program and many suggested that the government should implement a cost effectiveness measure. At the time, a commonly discussed threshold was between \$50,000 and \$100,000 per quality adjusted life year, with this number based on treatments that were currently being provided to dialysis patients (Millick, 2004). More recent work has estimated that dialysis treatment is valued at approximately \$129,000 per quality adjusted life year and \$62,000 per life year (Lee, Chertow, and Zenios, 2009). The estimates in Table 4 examine the change in the share of products that are above these two thresholds. The estimates related to the threshold of \$50,000 per life year suggest a large increase in the share of products over this threshold for those products covered by Part D compared to oncology products covered under Part B. The estimated effect ranges from an approximately 64 percent increase (p-value = 0.004) in the model with no covariates to an increase of approximately 43 percent (p-value = 0.128) in the model with both controls for product characteristics and disease fixed effects. The estimates related to the share of products over \$100,000 are large in magnitude but statistically insignificant across all three specifications.

While these estimated effects are for only one product class and change in coverage, they provide strong suggestive evidence of the effects predicted by our model. Following the insurance expansion and the forced bundle, the monopoly producers of oncology products were able to increase prices and capture far more value than they could when facing a market with less insurance and more flexibility regarding the composition of the insurance bundle.

7 The Role of Insurance in Recent High Drug Prices

While our results regarding the prices of oncology products following the passage of Part D show a sharp increase in those prices following 2003, a second clear pattern is the rising prices of all oncology products in

	Log Price		Log Price		Log Price	
	per Life Year		per Life Year		Gained	
I{Part D}	-1.966*** (0.462)	-1.691** (0.587)	-1.988*** (0.452)	-1.746*** (0.332)	-1.605*** (0.386)	-1.232** (0.44)
I{Part D}*I{Year > 2003}	2.07*** (0.6)	1.883** (0.69)	1.663** (0.559)	1.638*** (0.383)	1.59*** (0.407)	1.009* (0.505)
Deaths		-0.0004** (0.004)	0.008 (0.006)		0.003* (0.002)	-0.002 (0.006)
I{Combination Regimen}		0.407 (0.436)	0.855 (0.507)		0.594** (0.261)	0.794* (0.454)
Line		-0.023 (0.288)	-0.027 (0.436)		0.238 (0.174)	0.449 (0.315)
Orphan		0.27 (0.57)	-0.097 (0.558)		-0.133 (0.23)	-0.273 (0.488)
No. of competitors		-0.017 (0.028)	-0.01 (0.121)		-0.037* (0.011)	0.174 (0.136)
Priority		0.008 (0.418)	0.057 (0.301)		-0.117 (0.306)	-0.015 (0.295)
Biologic		-0.084 (0.278)	0.013 (0.39)		-0.075 (0.254)	-0.0392 (0.329)
r^2	0.732	0.744	0.941	0.811	0.863	0.93
N	58	58	58	58	58	58
Disease FE?	N	N	Y	N	N	Y

Table 3: Difference-in-Differences Estimates of the Change in Prices of Oncology Products After the Passage of Medicare Part D. *p-value < 0.1, **p-value < 0.05, ***p-value < 0.001.

	Price over \$50k per Life Year	Price over \$100k per Life Year
I{Part D}	-0.643** (0.209)	-0.143 (0.162)
I{Part D}*I{Year > 2003}	0.643** (0.209)	0.22 (0.21)
Deaths	-0.0004 (0.0007)	-0.0005 (0.001)
I{Combination Regimen}	0.29** (0.127)	0.48** (0.175)
Line	0.019 (0.053)	0.082 (0.137)
Orphan	0.145* (0.071)	-0.06 (0.133)
No. of competitors	-0.008* (0.005)	-0.005 (0.008)
Priority	0.125* (0.068)	0.057 (0.112)
Biologic	0.042 (0.061)	-0.062 (0.117)
r^2	0.835	0.666
N	58	58
Disease FE?	N	N
	Y	Y
	58	58
	0.899	0.899
	58	58

Table 4: Difference-in-Differences Estimates of the Change in Share of Oncology Products Above Price Thresholds After the Passage of Medicare Part D. * p-value < 0.1, ** p-value < 0.05, *** p-value < 0.001.

recent years. An open question is why has there been such a large increase in recent years for products covered by Part B. In addition, we note that these high prices have not been limited to oncology products. Over the last decade there has been increased attention on the high prices charged for new pharmaceuticals, peaking with the release of Sovaldi, Gilead’s cure for hepatitis C. Sovaldi was initially priced at approximately \$84,000 per course of treatment and was part of a wave of new specialty pharmaceutical products that cost upwards and in excess of \$100,000 per course of treatment. In addition to these highly priced new drugs, several small companies have recently acquired generic drugs and dramatically increased their prices. Perhaps the most highly discussed of these is Turing Pharmaceuticals, which bought the production facilities for Daraprim in August, 2015 and immediately raised the price from \$13.50 to \$750 per pill (Pollack, 2015). This strategy of price increases following the acquisition of existing products has also been seen for other companies such as Valeant Pharmaceuticals and Mallinckrodt Plc (Rockoff and Silverman, 2015).

It is difficult to find a simple explanation for these unprecedented prices and recent price increases. For example, one might attribute price increases to Part D, but the price data on oncology products show recent high prices for drugs covered by Part B. And while the Affordable Care Act was passed in 2009 and implemented fully in 2014, many of these high prices preceded this insurance expansion. Therefore, it seems unlikely that these insurance expansions, on their own, could explain the recent rise in drug prices.

During this same time period there have also been changes in market structure that, in combination with the changes in insurance programs, provide a possible explanation for these price increases. As recently as a decade ago, a large fraction of the highest grossing drug products in the United States was marketed by a relatively small number of large firms. This has changed dramatically over the last decade. Today, many firms, including several relative newcomers, sell just a handful of drugs. Figure 10 contains the HHI based on worldwide sales for the top 200 pharmaceuticals in every five years from 1998-2013. From 1998 to 2003, the market concentration among these products remains markedly similar. However, over the next five years there is a steep drop in the HHI which remains approximately the same through 2013. A similar pattern can be seen in Figure 11, which contains the percentage of worldwide sales in the top 200 for the average firm with a product that is among the top 10 grossing worldwide.

This shift has been driven both by expiring patents for high-selling drugs from large manufacturers and the introduction of new drugs by firms with a far smaller portfolio of products. Consider that in 1998, Pfizer had the 4th, 7th, and 8th highest grossing drugs while Merck had the 2nd and 5th. In that year Merck accounted for 11.8 percent of top 200 sales, Pfizer accounted for 9.7 percent, and the average market share for a firm with a top 10 grossing pharmaceutical was 6.7 percent. By 2013, each of these firms had only one top 10 drug and only Roche had more than one drug in the top 10. As a result, in 2013 the top 200 market share of these firms had fallen to less than 6 percent, Roche had the highest share of the top 200 that year accounted at 6.2 percent, and the average market share of a firm with a top 10 grossing drug was only 4.6 percent. As large firms left the top 10, they were often replaced by far smaller competitors.

These data document the increased fragmentation of drug sales. In addition, a shift in the research and development strategy of pharmaceutical firms has increasingly meant that drugs under production at the largest pharmaceutical firms began their life at smaller companies. Rockoff (2014) notes that from 2002-2012, the percentage of drugs under development by the top 10 pharmaceutical firms that began at another company increased from 16 percent to 33 percent. Smaller companies that receive royalties from their discoveries would be unlikely to worry about any negative externalities from high prices. This effectively further fragments the industry.

Our model of pharmaceutical pricing for an insured population suggests that the simultaneous occurrence of these two phenomena—higher prices and smaller firms—is no mere coincidence. Nor is the timing of the

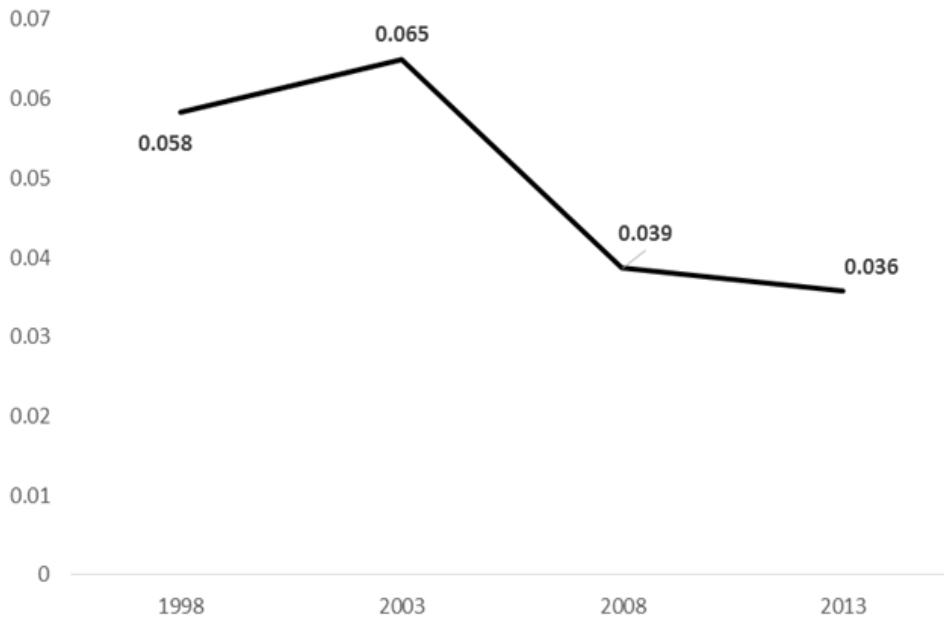


Figure 10: HHI for the top 200 highest selling drugs. Note: drug pricing data are taken from the MedAd News annual list of top 200 selling pharmaceuticals. Sales data are included for the firm with marketing rights (either sole or partnership) in the United States.

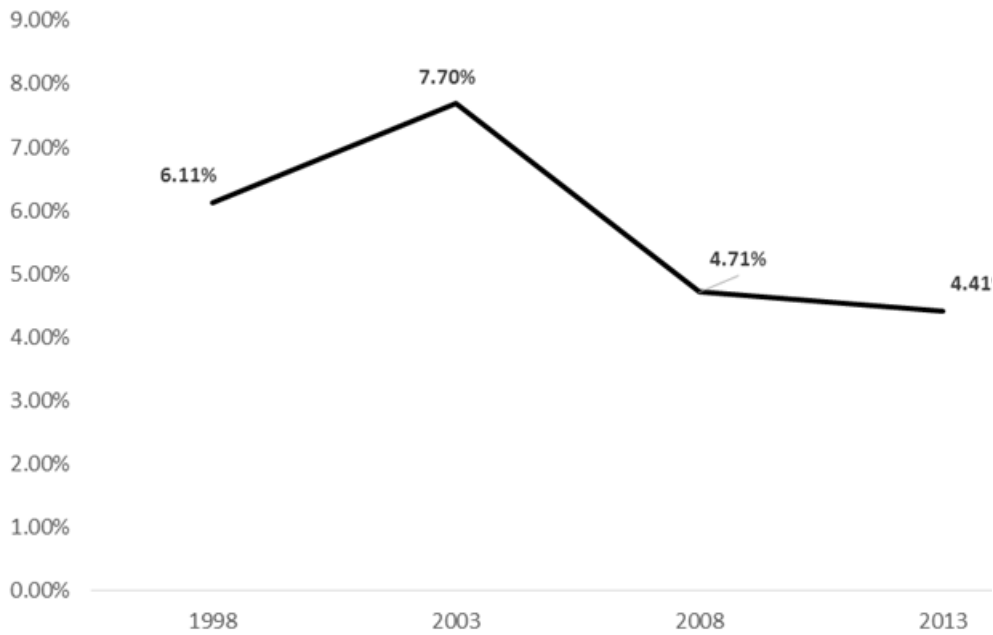


Figure 11: Average market share of the top 200 for a firm with a top 10 drug. Note: drug pricing data are taken from the MedAd News annual list of top 200 selling pharmaceuticals. Sales data are included for the firm with marketing rights (either sole or partnership) in the United States.

price spikes, coming shortly after the implementation of Medicare Part D. Recall that our model identifies two demand constraints on drug pricing. First, drug makers must consider the value that is created by their new drug. Second, they must consider the degree to which their price will make the bundled insurance product too expensive for some customers. This second constraint suggests that high prices charged by one firm can generate a negative externality on the sales of other drugs. This occurs through a decrease in the potential market of insured customers who, in the absence of insurance, would be liquidity constrained with respect to their purchases of high value drugs. As we note in Section 3, when multiple innovative products are sold by different firms, as is the case in the current market for speciality pharmaceuticals, prices will be weakly higher and if there are enough products total surplus will decrease as consumers decline to purchase the insurance bundle. This results from firms failing to internalize the effect of their price on the demand for the insurance bundle. As a result, drug makers acting in their own self-interest may set individual prices that exceed the prices that maximize industry profits. However, our model suggests that a market dominated by a few large firms with extensive product portfolios should have internalized more of the effect of their pricing decision for new products on the affordability of the insurance bundle and/or perhaps even behaved in an implicitly collusive fashion to hold the line on large price increases (Ellison and Wolfram, 2006).³⁹ In contrast, the increased fragmentation of the drug market should result in higher prices, even if we hold the size of the insured population constant.

Consistent with this model, note that Gilead, the marketer of Sovaldi, had only 3.6 percent of top 200 sales in 2014 (compared to a mean for a top 10 grossing firm of approximately 4.6 percent). Most of these sales were for Sovaldi, and the firm averaged approximately 1 percent of top 200 sales per year in preceding five years. Similarly, Turing Pharmaceuticals, the firm that acquired and raised the price of Daraprim has only been in existence for one year and has only two FDA approved products in its portfolio.

The expansions in insurance coverage resulting from Medicare Part D and the ACA would only exacerbate the incentive for price increases resulting from a fragmented pharmaceutical market. This is both because of the increase in the size of the insurance population and the increased bundling of drugs under a single insurance product further encourages higher pricing. For example, Medicare Part D requires coverage of all new cancer drugs as well as drugs in several other classes. Private insurers often follow Medicare's lead, for reasons that are not clear. The combination of a fragmented market and bundled coverage may have created a perfect storm of incentives for sky high prices. While further empirical work would be required to establish causality and magnitudes of these effects, they are quite suggestive of an unexplored channel to explain the recent rise in drug prices.

8 Conclusions

While health insurance bears many similarities to more traditional financial insurance products, it also conveys the important benefit of breaking the liquidity constraint that many consumers face when attempting to purchase costly but valuable medical goods and services, including high priced pharmaceutical treatments. Previous work has asserted that this increase in access represents a benefit for consumers. However, our results, which endogenize the prices charged by the monopoly manufacturers of high value products, demonstrate that consumers may actually enjoy less consumer surplus, even in the absence of moral hazard. This occurs because insurance allows firms to more fully price out the expected benefits that their drugs generate for consumers, charging prices that may be far higher than what they would select if patients had to pay the

³⁹Note that this economic externality effect is distinct from, but has the same implications as, the oft heard claim that drug makers keep prices low simply to avoid the prospects of price regulation.

entire price of their drugs out of pocket. This not only reduces consumer surplus, it may also reduce total surplus if some consumers are liquidity constrained from purchasing insurance. Thus, regulations requiring insurers to provide coverage for all new high-value products, such as the minimum insurance definition of the ACA, can decrease consumer surplus and, potentially, decrease total surplus.

We also explore the traditional bundling of pharmaceutical insurance with insurance for other medical services, and find that bundling 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 products and services. This may explain why some the prices of some drugs seem to exceed their expected value without relying on consumer inattention, asymmetric information, or moral hazard. While regulations establishing minimum quality standards for insurance and stipulations about what pharmaceutical products must be included on a formulary can plausibly create consumer surplus gains, our results show that these must be weighed against losses that come from these pricing dynamics.

It is important to note that drug makers only benefit from the access motive while their drug is under patent. While our model considers only this time period, it is clear that following the expiration of a patent the decrease in prices resulting from the entry of generic competitors would transfer surplus from producers to consumers. This reminds us 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. Previous research has shown that pharmaceutical research and development efforts are correlated with expected profits (Acemoglu and Linn, 2004; Finkelstein, 2004; Blume-Kohut and Sood, 2012; Dranove, Garthwaite, and Hermsillo, 2014). To fully understand the welfare dynamics of the access motive of insurance, future work should consider whether increased profits (and lost consumer surplus) while the drug is manufactured by a monopoly producer provide the appropriate incentives for products that could have large welfare benefits after patent expiration. This would involve understanding both the innovators response to profits and the longevity of the welfare benefits of new products in comparison the effective patent period. 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.

9 Appendix

Derivation of expected utilities from purchasing insurance:

To determine the expected utility from purchasing and not purchasing insurance, we work backwards. Let’s first consider someone who does not purchase health insurance (of either kind, basic or expanded).

- If this individual *does not* become ill and *does not* need core medical services, the consumer problem is trivial: the individual sets $x_i = y_i = 0$ and $z_i = \widehat{W}_i$, resulting in utility $u_i^{NI}(s = 0) = \widehat{W}_i$.
- If the individual *does not* become ill but *needs core medical services*, the individual’s problem is:

$$\begin{aligned} & \max_{x_i, y_i, z_i} By_i + z_i \\ \text{subject to: } & Px_i + Cy_i + z_i \leq \widehat{W}_i. \end{aligned}$$

Clearly, $x_i = 0$. Moreover, by Assumption 2, $\widehat{W}_i \leq \widehat{W} + 1 < C$, so the individual is liquidity constrained from buying core medical services, and the individual must set $y_i = 0$, resulting in utility level $u_i^{NI}(s = 1) = \widehat{W}_i$.

- If this individual *does* become ill, the individual's utility maximization problem is:

$$\begin{aligned} & \max_{x_i, y_i, z_i} V_i x_i + z_i \\ \text{subject to: } & P x_i + C y_i + z_i \leq \widehat{W}_i. \end{aligned}$$

Clearly, $y_i = 0$. Now, if $\widehat{W}_i < P$, the individual is liquidity constrained from purchasing the treatment. In this case, $x_i = 0$, $y_i = 0$, and $z_i = \widehat{W}_i$, and the individual's utility is \widehat{W}_i . If $\widehat{W}_i \geq P$, the individual is not liquidity constrained from purchasing the treatment, and thus the individual could potentially purchase the treatment. By Assumption 1, $V_i > \widehat{W}_i$, so $\widehat{W}_i \geq P$ implies that $V_i - P > 0$, so the individual will, in fact, purchase the treatment, resulting in $x_i = 1$, $z_i = \widehat{W}_i - P$, and utility $\widehat{W}_i + V_i - P$. Thus, when the individual does become ill, his utility is

$$u_i^{NI}(s = 2) = \widehat{W}_i + (V_i - P) 1(\widehat{W}_i \geq P).$$

- Thus expected utility from not purchasing insurance is

$$\begin{aligned} EU^{NI} &= (1 - \rho_C - \rho_I) u_i^{NI}(s = 0) + \rho_C u_i^{NI}(s = 1) + \rho_I u_i^{NI}(s = 2) \\ &= \widehat{W}_i + \rho_I (V_i - P) 1(\widehat{W}_i \geq P). \end{aligned}$$

which is the expression in (1).

Next, consider the problem of someone who purchased the basic health insurance plan at premium $\rho_C C$. We note that given Assumption 3, this option is available to all consumers.

- If the individual *does not* become ill and *does not need* core medical services, the individual's utility maximization problem is:

$$\begin{aligned} & \max_{x_i, y_i, z_i} z_i \\ \text{subject to: } & P x_i + z_i \leq \widehat{W}_i - \rho_C C. \end{aligned}$$

Clearly, $x_i = 0$, $y_i = 0$, $z_i = \widehat{W}_i - \rho_C C$, so $u_i^{BI}(s = 0) = \widehat{W}_i - \rho_C C$.

- If the individual *does not become ill but needs core medical services*, the individual's utility maximization problem is:

$$\begin{aligned} & \max_{x_i, y_i, z_i} B y_i + z_i \\ \text{subject to: } & P x_i + z_i \leq \widehat{W}_i - \rho_C C. \end{aligned}$$

In this case, $x_i = 0$, $y_i = 1$, $z_i = \widehat{W}_i - \rho_C C$, so $u_i^{BI}(s = 0) = \widehat{W}_i + B - \rho_C C$.

- If the individual *does become ill*, the individual's problem is:

$$\begin{aligned} & \max_{x_i, y_i, z_i} V_i x_i + z_i \\ \text{subject to: } & P x_i + z_i \leq \widehat{W}_i - \rho_C C. \end{aligned}$$

In this contingency, we must distinguish between two cases: If $\widehat{W}_i - \rho_C C < P$, the individual will be liquidity constrained and will not be able to purchase the treatment. For such an individual, $x_i = 0$ and $z_i = \widehat{W}_i - \rho_C C$, and that individual gets utility $\widehat{W}_i - \rho_C C$. On the other hand, if $\widehat{W}_i - \rho_C C \geq P$, the individual can afford to purchase the treatment. The individual will purchase the treatment provided $V_i - P > 0$. Now, note that $\widehat{W}_i - \rho_C C \geq P$ implies $\widehat{W}_i \geq \rho_C C + P > P$, and since by Assumption 1, $V_i > \widehat{W}_i$, it follows that $\widehat{W}_i \geq P$ implies that $V_i - P > 0$, so the individual will, in fact, purchase

the treatment. This results in $z_i = \widehat{W}_i - \rho_C C - P$, and utility $\widehat{W}_i - \rho_C C + V_i - P$. Thus, when the individual does become ill, his utility is

$$u_i^{BI}(s=2) = \widehat{W}_i - \rho_C C + (V_i - P) 1(\widehat{W}_i - \rho_C C \geq P).$$

Thus, expected utility from purchasing the basic health insurance plan is

$$\begin{aligned} EU^{BI} &= (1 - \rho_C - \rho_I)u_i^{BI}(s=0) + \rho_C u_i^{BI}(s=1) + \rho_I u_i^{BI}(s=2) \\ &= \widehat{W}_i + \rho_C(B - C) + (V_i - P) 1(\widehat{W}_i - \rho_C C \geq P). \end{aligned}$$

Finally, let's consider the utility maximization problem of someone who has purchased the expanded insurance plan that covers the treatment and whose premium is $\rho_C C + \rho_I P$. We note that for this to be feasible given the consumer's budget constraint, it must be the case that that $\widehat{W}_i \geq \rho_C C + \rho_I P$.

- If the individual *does not* become ill and *does not need* core medical services, the individual's utility maximization problem is

$$\begin{aligned} &\max_{x_i, y_i, z_i} z_i \\ \text{subject to:} & z_i \leq \widehat{W}_i - \rho_C C - \rho_I P. \end{aligned}$$

Clearly, $x_i = 0$, $y_i = 0$, $z_i = \widehat{W}_i - \rho_C C - \rho_I P$, and this gives the individual a utility level $u_i^I(s=0) = \widehat{W}_i - \rho_C C - \rho_I P$.

- If the individual *does not become ill but needs* core medical services, the individual's utility maximization problem is

$$\begin{aligned} &\max_{x_i, y_i, z_i} B y_i + z_i \\ \text{subject to:} & z_i \leq \widehat{W}_i - \rho_C C - \rho_I P. \end{aligned}$$

Clearly, $x_i = 0$, $y_i = 1$, $z_i = \widehat{W}_i - \rho_C C - \rho_I P$: the individual consumes the core medical services he is insured for and does not purchase the (unneeded) treatment. This gives the individual a utility level $u_i^I(s=1) = \widehat{W}_i + B - \rho_C C - \rho_I P$.

- If the individual *does become ill*, the individual's problem is

$$\begin{aligned} &\max_{x_i, y_i, z_i} V_i x_i + z_i \\ \text{subject to:} & z_i \leq \widehat{W}_i - \rho_C C - \rho_I P. \end{aligned}$$

Clearly, $x_i = 1$, $y_i = 0$, $z_i = \widehat{W}_i - \rho_C C - \rho_I P$: the individual obtains the treatment for free and does not purchase the (unneeded) core medical services. This gives the individual a utility level $u_i^I(s=2) = \widehat{W}_i + V_i - \rho_C C - \rho_I P$.

Thus, if the purchases the health insurance plan at premium $\rho_C C + \rho_I P$, expected utility EU_i^I is given by

$$\begin{aligned} EU_i^I &= (1 - \rho_C - \rho_I)u_i^I(s=0) + \rho_C u_i^I(s=1) + \rho_I u_i^I(s=2) \\ &= \widehat{W}_i + \rho_C(B - C) + \rho_I(V_i - P). \end{aligned}$$

Note that all individuals who purchase expanded insurance purchase the treatment if they become ill. ■

Proof of Proposition 1:

There are three relevant cases to consider:

- $\widehat{W}_i - \rho_C C \geq P$, or equivalently, $\widehat{W}_i \geq P + \rho_C C > P$. For a consumer in this situation, $EU_i^{BI} = \widehat{W}_i + \rho_C(B - C) + \rho_I(V_i - P) > \widehat{W}_i + \rho_I(V_i - P) = EU_i^{NI}$. The consumer would buy basic health

insurance. (Note: by Assumption 3, the basic plan is affordable to this set of consumers and indeed all consumers). Moreover, the consumer can afford to purchase the treatment if he becomes ill (because $\widehat{W}_i - \rho_C C \geq P$) and will do so (because $\widehat{W}_i > P$ implies $V_i > P$). Note that the consumer is strictly better off purchasing basic health insurance than not. This is, ultimately, due to the fact that basic insurance breaks the liquidity constraint on purchases of core medical services, a consequence of Assumptions 2 and 3.

- $\widehat{W}_i < P$. For a consumer in this situation $EU_i^{BI} = \widehat{W}_i + \rho_C(B - C) > \widehat{W}_i = EU_i^{NI}$. The consumer would buy basic health insurance. However, because $\widehat{W}_i < P < P + \rho_C C$, and thus $\widehat{W}_i - \rho_C C < P$, the consumer would be unable to afford to purchase the treatment if he becomes ill and thus will not do. (Note: this would also have been the case if the consumer had not purchased basic health insurance.)
- $P \leq \widehat{W}_i < P + \rho_C C$. For a consumer in this situation, $EU_i^{BI} = \widehat{W}_i + \rho_C(B - C)$ and $EU_i^{NI} = \widehat{W}_i + \rho_I(V_i - P)$. A consumer in this situation who purchases basic insurance can afford core medical services if he needs them, but he cannot afford the treatment if he becomes ill. On the other hand, if the consumers does not purchase basic insurance, he cannot afford core medical services if he needs them, but he can afford to purchase the treatment if he becomes ill. To determine consumer choice in this case note that

$$\begin{aligned} EU_i^{BI} - EU_i^{NI} &= \rho_C(B - C) - \rho_I(V_i - P) \\ &> \rho_C(B - C) - \rho V_i \\ &> \rho_C(B - C) - \rho v(\widehat{W} + 1) \\ &> 0, \end{aligned}$$

where the last inequality follows from Assumption 4. Thus, a consumer in this situation purchases basic health insurance, but if he becomes ill, he will be unable to afford to purchase the treatment. ■

Derivation of the demand curve for the treatment under expanded health insurance:

From (3) a consumer's expected utility if he purchases expanded health insurance that covers the treatment is

$$EU_i^I = \widehat{W}_i + \rho_C(B - C) + \rho_I(V_i - P), \quad (17)$$

while his expected utility if he does not purchase health insurance is:

$$EU_i^{NI} = \widehat{W}_i + \rho_I(V_i - P) 1(\widehat{W}_i \geq P). \quad (18)$$

We first show that individuals who *can afford* to purchase expanded insurance— $\widehat{W}_i \geq \rho_C C + \rho_I P$, or equivalently, $V_i \geq \rho_I P$ —do indeed purchase if and only if $V_i \geq P - \frac{\rho_C}{\rho_I}(B - C)$.

Proof that if $V_i \geq P - \frac{\rho_C}{\rho_I}(B - C)$, then individuals who can afford expanded insurance purchase it:

Individuals who can afford to purchase expanded insurance make their decision by evaluating $EU_i^I - EU_i^{NI}$. If $V_i \geq P - \frac{\rho_C}{\rho_I}(B - C)$, there are two possibilities: either $V_i \geq P$ or $P > V_i \geq P - \frac{\rho_C}{\rho_I}(B - C)$. When $V_i \geq P$,

$$EU_i^I - EU_i^{NI} = \rho_C(B - C) + \rho_I(V_i - P) [1 - 1(\widehat{W}_i \geq P)] > 0,$$

because $\rho_C(B - C) > 0$, $[1 - 1(\widehat{W}_i \geq P)] \geq 0$, and by assumption $V_i - P \geq 0$. If $P > V_i \geq P - \frac{\rho_C}{\rho_I}(B - C)$, then we also have $P > V_i > \widehat{W}_i$, and thus $EU_i^{NI} = \widehat{W}_i$ and

$$EU_i^I - EU_i^{NI} = \rho_C(B - C) + \rho_I(V_i - P) \geq 0,$$

where the inequality follows because $V_i \geq P - \frac{\rho_C}{\rho_I}(B - C)$.

Proof that if individuals who can afford expanded insurance purchase it, then $V_i \geq P - \frac{\rho_C}{\rho_I}(B - C)$:

Suppose, to the contrary, that individuals who can afford to insurance do purchase it, but $V_i < P - \frac{\rho_C}{\rho_I}(B - C)$. This implies $V_i < P$, which in turn implies $\widehat{W}_i < P$, since $\widehat{W}_i < V_i$. Using (17) and (18) we then have

$$EU_i^I - EU_i^{NI} = \rho_C(B - C) + \rho_I(V_i - P) < 0,$$

where the inequality follows from the contrapositive assumption $V_i < P - \frac{\rho_C}{\rho_I}(B - C)$. But this contradicts our premise that individuals who can afford insurance purchase it.

Now consider consumers for whom expanded insurance is not affordable: $\widehat{W}_i \leq \rho_C C + \rho_I P$, or equivalently, $W_i \leq \rho_I P$. However, if the individual becomes ill, the individual might still purchase the treatment. The individual would be able to afford to do so if $\widehat{W}_i > P$ (or equivalently $W_i > P - \rho_C C$), and if that is the case, then $V_i \geq P$ (since $V_i > W_i$) and thus the person would obtain the treatment.

Thus, summing up, the demand for the treatment comes from two groups of consumers:

(1) Consumers for whom $\widehat{W}_i \geq \rho_C C + \rho_I P$ (or equivalently, $W_i \geq \rho_I P$) and $V_i \geq P - \frac{\rho_C}{\rho_I}(B - C)$. These are consumers can afford to purchase insurance and they do so, thus generating demand for the treatment when they fall ill;

(2) Consumers for whom $\widehat{W}_i < \rho_C C + \rho_I P$ (or equivalently, $W_i < \rho_I P$) cannot afford to purchase insurance and will not do so. However, a subset of these consumers for whom $\widehat{W}_i \geq P$, or equivalently $W_i > P - \rho_C C$, can afford to purchase insurance if they fall ill, and they will do so since $V_i > \widehat{W}_i > P$. These are consumers who cannot afford to purchase insurance but would purchase the treatment if the price P was sufficiently low.

In light of this, the demand for the treatment is by the measure of consumers such that

$$\left[\begin{array}{c} \left\{ \frac{W_i}{\rho_I} \geq P \text{ and } V_i \geq P - \frac{\rho_C}{\rho_I}(B - C) \right\} \\ \text{and } \left\{ \frac{W_i}{\rho_I} < P \text{ and } W_i \geq P - \rho_C C \right\} \end{array} \right]$$

Using Assumption 5 that $V_i = \underline{V} + \alpha [W_i - \underline{W}]$, the demand $D(P)$ for the treatment is given by

$$D(P) = D^*(P) + D_u(P)$$

where

$$\begin{aligned} D^*(P) &= \int_{\{W_i \in [\underline{W}, \underline{W}+1] | W_i > \max\{\rho_I P, \frac{P - \frac{\rho_C}{\rho_I}(B-C) - \underline{V}}{\alpha} + \underline{W}\}\}} dW_i \\ &= \int_{\max\{\rho_I P, \frac{P - \frac{\rho_C}{\rho_I}(B-C) - \underline{V}}{\alpha} + \underline{W}\}}^{\underline{W}+1} dW_i \\ &= \underline{W} + 1 - \max\left\{\rho_I P, \frac{P - \frac{\rho_C}{\rho_I}(B-C) - \underline{V}}{\alpha} + \underline{W}\right\} \\ &= \begin{cases} 1 & \max\{\rho_I P, \frac{P - \frac{\rho_C}{\rho_I}(B-C) - \underline{V}}{\alpha} + \underline{W}\} \leq \underline{W} \\ \underline{W} + 1 - \min\left\{\rho_I P, \frac{P - \frac{\rho_C}{\rho_I}(B-C) - \underline{V}}{\alpha} + \underline{W}\right\} & \underline{W} \leq \max\{\rho_I P, \frac{P - \frac{\rho_C}{\rho_I}(B-C) - \underline{V}}{\alpha} + \underline{W}\} \leq \underline{W} + 1 \\ 0 & \max\{\rho_I P, \frac{P - \frac{\rho_C}{\rho_I}(B-C) - \underline{V}}{\alpha} + \underline{W}\} \geq \underline{W} + 1 \end{cases} \end{aligned}$$

$$\begin{aligned} D_u(P) &= \int_{\{W_i \in [\underline{W}, \underline{W}+1] | P - \rho_C C < W_i < \rho_I P\}} dW_i \\ &= 1[P < \frac{\rho_C}{1 - \rho_I} C] \times \begin{cases} 1 & P > \frac{\underline{W}+1}{\rho_I}, P < \underline{W} + \rho_C C \\ \rho_C C - (1 - \rho_I)P & P \in \left[\frac{\underline{W}}{\rho_I} + \rho_C C, \frac{\underline{W}+1}{\rho_I}\right] \\ \underline{W} + 1 + \rho_C C - P & \max\left\{\frac{\underline{W}+1}{\rho_I}, \underline{W} + \rho_C C\right\} \leq P \leq \underline{W} + \rho_C C \\ \rho_I P - \underline{W} & \frac{\underline{W}}{\rho_I} \leq P \leq \min\left\{\frac{\underline{W}+1}{\rho_I}, \underline{W} + \rho_C C\right\} \\ 0 & P < \frac{\underline{W}}{\rho_I} \end{cases} \end{aligned}$$

We now prove that $D_u(P) = 0$ when $\underline{W} \geq \frac{\rho_C \rho_I}{1 - \rho_I} C$. 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 $D^*(P)$. This can be rewritten as $P' < \frac{\rho_C}{1 - \rho_I} C$, where $\frac{\rho_C}{1 - \rho_I} < 1$ (because $1 - \rho_I - \rho_C > 0$). 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., $\underline{W} < \rho_I P'$. Otherwise, all consumers would be “located” on $D^*(P)$. This can be rewritten as $P' > \frac{\underline{W}}{\rho_I}$. These two necessary conditions imply $\frac{\underline{W}}{\rho_I} < \frac{\rho_C}{1-\rho_I} C$, or equivalently $\underline{W} < \frac{\rho_C \rho_I}{1-\rho_I} C$. Without this condition, i.e., $\underline{W} \geq \frac{\rho_C \rho_I}{1-\rho_I} C$, there can be no prices that at which consumers would be unable to afford insurance but could afford the treatment if they became ill.

To illustrate $D^*(P)$ graphically and at the same time, to derive the expression for its inverse $P^*(Q)$, pick an arbitrary price P' and substitute it into

$$Q = \underline{W} + 1 - \rho_I P'$$

and

$$Q = \underline{W} + 1 - \left[\frac{P' - \frac{\rho_C}{\rho_I}(B - C) - \underline{V}}{\alpha} + \underline{W} \right].$$

This yields

$$\begin{aligned} Q'_1 &= \underline{W} + 1 - \rho_I P' \text{ which implies } P' = \frac{P_0(Q'_1)}{\rho_I} \\ Q'_2 &= \underline{W} + 1 - \left[\frac{P' - \frac{\rho_C}{\rho_I}(B - C) - \underline{V}}{\alpha} + \underline{W} \right] \text{ which implies } P' = P_1(Q'_2) + \frac{\rho_C}{\rho_I}(B - C), \end{aligned}$$

where recall $P_1(Q) = \underline{V} + \alpha - \alpha Q$. There are two possibilities. First, suppose that

$$\max \left\{ \rho_I P', \frac{P' - \frac{\rho_C}{\rho_I}(B - C) - \underline{V}}{\alpha} + \underline{W} \right\} = \rho_I P',$$

which implies $Q'_1 < Q'_2$. Because $\rho_I P' = \max\{\rho_I P', \frac{P' - \underline{V}}{\alpha} + \underline{W}\}$, (Q'_1, P') is a point along $D^*(P)$. Moreover, $P' = \frac{\underline{W} + 1 - Q'_1}{\rho_I} = \frac{P_0(Q'_1)}{\rho_I}$ and so (Q'_1, P') is a point along $\frac{P_0(Q)}{\rho_I}$ as well. Moreover, since $Q'_1 < Q'_2$ it follows that $P_1(Q'_1) > P_1(Q'_2) = P' = \frac{P_0(Q'_1)}{\rho_I}$. Thus, in the range where the inverse of $D^*(P) = \frac{P_0(Q)}{\rho_I}$, we have $P_1(Q) + \frac{\rho_C}{\rho_I}(B - C) > \frac{P_0(Q)}{\rho_I}$, i.e., the inverse demand curve with no liquidity constraints lies above the “grossed up” liquidity constrained inverse demand curve.

Next, suppose that

$$\max \left\{ \rho_I P', \frac{P' - \frac{\rho_C}{\rho_I}(B - C) - \underline{V}}{\alpha} + \underline{W} \right\} = \frac{P' - \frac{\rho_C}{\rho_I}(B - C) - \underline{V}}{\alpha} + \underline{W},$$

which implies that $Q'_1 > Q'_2$. Because $\frac{P' - \frac{\rho_C}{\rho_I}(B - C) - \underline{V}}{\alpha} + \underline{W} = \max\{\rho_I P', \frac{P' - \frac{\rho_C}{\rho_I}(B - C) - \underline{V}}{\alpha} + \underline{W}\}$, (Q'_2, P') is a point along $D^*(P)$. Moreover, $P' = \underline{V} + \alpha - \alpha Q'_2 + \frac{\rho_C}{\rho_I}(B - C) = P_1(Q'_2) + \frac{\rho_C}{\rho_I}(B - C)$, so (Q'_2, P') is a point along $P_1(Q) + \frac{\rho_C}{\rho_I}(B - C)$ as well. Moreover, since $Q'_1 > Q'_2$ it follows that $\frac{P_0(Q'_2)}{\rho_I} > \frac{P_0(Q'_1)}{\rho_I} = P' = P_1(Q'_2) + \frac{\rho_C}{\rho_I}(B - C)$. Thus, in the range where the inverse of $D^*(P) = P_1(Q) + \frac{\rho_C}{\rho_I}(B - C)$, we have $\frac{P_0(Q)}{\rho_I} > P_1(Q) + \frac{\rho_C}{\rho_I}(B - C)$, i.e., the “grossed up” liquidity constrained inverse demand curve lies above the inverse demand curve with no liquidity constraints plus $\frac{\rho_C}{\rho_I}(B - C)$. From all of this it follows that when $\underline{W} \geq C$, the demand curve for the treatment is the lower envelope of $\frac{P_0(Q)}{\rho_I}$ and $P_1(Q) + \frac{\rho_C}{\rho_I}(B - C)$. We can thus write that demand curve in inverse form as

$$\begin{aligned} P^*(Q) &= \min \left\{ \frac{P_0(Q)}{\rho_I}, P_1(Q) + \frac{\rho_C(B - C)}{\rho_I} \right\} \\ &= \min \left\{ \frac{\underline{W} + 1 - Q}{\rho_I}, \underline{V}' + \alpha - \alpha Q \right\}, Q \in [0, 1], \end{aligned}$$

where

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

Finally, we illustrate that this demand curve can have kinks, and we show where the kinks are located. First, let Q_K be such that

$$\frac{P_0(Q_K)}{\rho_I} = P_1(Q_K) + \frac{\rho_C(B-C)}{\rho_I},$$

which implies

$$Q_K = \frac{W + 1 - \rho_I(\underline{V}' + \alpha)}{1 - \alpha\rho_C}.$$

If

$$\frac{W + 1 - \rho_I(\underline{V}' + \alpha)}{1 - \alpha\rho_C} \in (0, 1)$$

the demand curve for the treatment has a kink at Q_K , which corresponds to a price of P_K

$$P_K = \frac{\underline{V}' - \alpha W}{1 - \alpha\rho_C}.$$

The point (Q_K, P_K) corresponds to point B in Panel C of Figure 5: the point at which we switch from treatment demand being determined by affordability of insurance to it being determined by the value of insurance. There is also a kink at point A that corresponds to the price

$$P^*(1) = \min \left\{ \frac{W}{\rho_I}, \underline{V}' \right\}.$$

■

Proof that innovator's profit-maximization problem under expanded insurance has a corner solution:

The general solution to the optimization problem in (12) is

$$Q^* = \begin{cases} \frac{1}{2} & \underline{V}' \geq \alpha \\ \frac{\underline{V}' + \alpha}{2\alpha} & \underline{V}' \leq \alpha \end{cases}. \quad (19)$$

$$P^* = \begin{cases} \underline{V}' & \underline{V}' \geq \alpha \\ \frac{\underline{V}' + \alpha}{2} & \underline{V}' \leq \alpha \end{cases}. \quad (20)$$

We now show that Assumptions 1, 4, and 5 imply that $\underline{V}' > 1$, which in turn (from (19) and (20)) imply that $Q^* = 1$ and $P^* = \underline{V}'$.

By definition

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

By Assumption 1, $\underline{V} + \alpha > \widehat{W} + 1$, so

$$\underline{V}' > \widehat{W} + 1 - \alpha + \frac{\rho_C}{\rho_I}(B-C). \quad (21)$$

Now, by Assumption 4, $\frac{\rho_C}{\rho_I}(B-C) > \underline{V} + \alpha$, or equivalently, $\frac{\rho_C}{\rho_I}(B-C) - \alpha > \underline{V}$, which from (21) implies

$$\underline{V}' > \widehat{W} + \underline{V} + 1 > 1,$$

since $\underline{V} > \widehat{W} \geq 0$. ■

Proof of Proposition 2:

Note that

$$P^* = \underline{V}' = \underline{V} + \frac{\rho_C}{\rho_I}(B-C). \quad (22)$$

Given Assumption 4, we have

$$\frac{\rho_C}{\rho_I}(B-C) > \underline{V} + \alpha > \alpha,$$

which in conjunction with (22) implies

$$\begin{aligned} P^* &> \underline{V} + \alpha \\ &> V_i = v(W_i) \text{ for all } W_i \in [\underline{W}, \underline{W} + 1], \end{aligned}$$

where the second inequality follows because $\alpha > 0$, and $\underline{V} + \alpha$ is the highest possible consumer valuation. ■

Proof of Proposition 3:

We have two possibilities:

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

$$P^* - P^0 = \underline{V}' - \left(\frac{\underline{W} + 1}{2} \right). \quad (23)$$

Because $\underline{V}' > \underline{V}$ and (from Assumption 5), $\underline{V} + \alpha > \underline{W} + 1$, it follows that

$$\underline{V}' > \underline{W} + 1 - \alpha. \quad (24)$$

Now, as part of the proof that the innovator's profit-maximization problem under expanded insurance has a corner solution, we established that $\underline{V}' > 1$, and since $\alpha < 1$, it follows that

$$\underline{V}' > \alpha. \quad (25)$$

Together (24) and (25) imply $2\underline{V}' > \underline{W} + 1$, which given (23) implies that $P^* - P^0 > 0$. ■

Proof of Proposition 4:

First, consider the case in which $\underline{W} \geq 1$, so $Q^0 = Q^* = 1$. In this case, coverage of the treatment leaves the quantity of the treatment purchased unchanged. Moreover, it also does not reduce health insurance coverage, so consumption of core medical services is unchanged. However, recall from Proposition 3 that the price of the treatment is higher when it is covered by insurance. Note that the expression for the consumer surplus function $CS(Q, Y, P)$ in (10) implies that $\frac{\partial CS}{\partial P} = -Q < 0$, so it follows that when $\underline{W} \geq 1$,

$$CS^0 = CS(1, 1, P^0) > CS(1, 1, P^*) = CS^*.$$

Next, consider the case in which $\underline{W} < 1$, so $Q^0 = \frac{\underline{W}+1}{2} < Q^* = 1$. Using (13) and (14), we can write the difference in consumer surplus as

$$CS^0 - CS^* = \frac{\rho_C(B - C)}{\rho_I} + \frac{\underline{W} + 1}{2} \left(\underline{V} + \frac{3\alpha}{4} - \frac{\alpha\underline{W}}{4} - \frac{\underline{W}}{2} - \frac{1}{2} \right) - \frac{\alpha}{2}. \quad (26)$$

We now show that this expression is positive. To do so, we note that

$$\begin{aligned} \frac{\partial [CS^0 - CS^*]}{\partial \alpha} &= \left(\frac{\underline{W} + 1}{2} \right) \left(\frac{3}{4} - \frac{\underline{W}}{4} \right) - \frac{1}{2} \\ &= -\frac{1}{8} (\underline{W} - 1)^2 < 0. \end{aligned}$$

Thus to establish that $CS^0 - CS^* > 0$ it suffices to show that $CS^0 - CS^*|_{\alpha=\underline{V}'} > 0$. This is because $\alpha < 1 < \underline{V}'$ —recall that we established the latter inequality as part of the proof that $Q^* = 1$ —and thus if $CS^0 - CS^*|_{\alpha=\underline{V}'} \geq 0$, it follows that $CS^0 - CS^* > 0$ for all $\alpha < 1$.

Substituting \underline{V}' for α in (26) gives us

$$CS^0 - CS^*|_{\alpha=\underline{V}} = \frac{\rho_C(B-C)}{\rho_I} + \frac{W+1}{2} \left(\underline{V}' + \frac{3V'}{4} - \frac{WV'}{4} - \frac{W}{2} - \frac{1}{2} \right) - \frac{V'}{2} \quad (27)$$

Because

$$\underline{V} = \underline{V}' - \frac{\rho_C(B-C)}{\rho_I},$$

we can write the second term on the right-hand side of (27) as

$$\frac{W+1}{2} \left(\underline{V}' - \frac{\rho_C(B-C)}{\rho_I} + \frac{3V'}{4} - \frac{WV'}{4} - \frac{W}{2} - \frac{1}{2} \right) - \frac{1}{2} \left(\underline{V}' - \frac{\rho_C(B-C)}{\rho_I} \right).$$

Substituting this into (27) and rearranging terms gives us

$$CS^0 - CS^*|_{\alpha=\underline{V}} = \left(\frac{\rho_C(B-C)}{\rho_I} \right) \left(\frac{3}{2} - \frac{W+1}{2} \right) + \frac{W+1}{2} \left(\underline{V}' + \frac{3V'}{4} - \frac{WV'}{4} - \frac{W}{2} - \frac{1}{2} \right) - \frac{V'}{2}.$$

Now recall that we are considering the case in which we have an interior solution under basic insurance, i.e.,

$$Q^0 = \frac{W+1}{2} < 1.$$

Thus $\frac{3}{2} - \frac{W+1}{2} > 0$, and so

$$\begin{aligned} CS^0 - CS^*|_{\alpha=\underline{V}} &> \frac{W+1}{2} \left(\underline{V}' + \frac{3V'}{4} - \frac{WV'}{4} - \frac{W}{2} - \frac{1}{2} \right) - \frac{V'}{2} \\ &= \frac{3WV'}{4} - \frac{W^2V'}{8} - \frac{W^2}{4} - \frac{W}{2} + \frac{3V'}{8} - \frac{1}{4} \\ &\geq \frac{3WV'}{4} - \frac{WV'}{8} - \frac{W}{4} - \frac{W}{2} + \frac{3V'}{8} - \frac{1}{4} \\ &= \frac{5WV'}{8} - \frac{3W}{4} + \frac{3V'}{8} - \frac{1}{4} \\ &\geq \frac{5W}{8} - \frac{3W}{4} + \frac{3}{8} - \frac{1}{4} \\ &= -\frac{(W-1)}{8} > 0. \end{aligned}$$

The first and final inequalities follows because $\underline{W} < 1$ and thus $\underline{W}^2 < \underline{W}$. The second inequality follows because, as we showed as part of the proof that $Q^* = 1$, Assumptions 1, 4, and 5 imply $\underline{V}' > 1$. Thus, $CS^0 - CS^* > 0$. ■

Proof of Proposition 5:

When the demand curve for the treatment is the value constraint,

$$D^*(P) = \begin{cases} 0 & P \geq \underline{V}' + \alpha \\ \frac{\underline{V}'+\alpha}{\alpha} - \frac{P}{\alpha} & P \in [\underline{V}', \underline{V}' + \alpha] \\ 1 & P \leq \underline{V}' \end{cases} . \quad (28)$$

We have seen that with a single innovator we have a corner solution in which the innovator serves the entire market, i.e., $Q^* = D^*(P^*) = D^*(\underline{V}') = 1$. 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 $1 - \frac{\underline{V}'}{\alpha} \leq 0$, which we know holds because $\underline{V}' > 1 > \alpha$.

Now consider the profit $\pi(P_1, \dots, P_N)$ of one of $N > 1$ innovators:

$$\pi(P_1, \dots, P_N) = P_k D^*(P_A).$$

Thus,

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

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

$$\left. \frac{d\pi}{dP_k} \right|_{P_1 = \dots = P_N = \underline{V}'} = 1 - \frac{\underline{V}'}{N\alpha}.$$

Since $\underline{V}' > 1 > \alpha$, this expression will be non-positive for $N \leq N^* \equiv \frac{\underline{V}'}{\alpha}$. For such N , the symmetric Nash equilibrium price $P_N^* = \underline{V}'$. However, for $N > N^*$,

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

In this case, 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

$$\left. \frac{d\pi}{dP_k} \right|_{P_1 = \dots = P_N = P_N^*} = 0,$$

or

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

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

Proof of Proposition 6:

If $N < \frac{\underline{V}'}{\alpha}$, we have a corner solution in which $Q_N^* = 1$ and $P_N^* = \underline{V}'$, just as in the single-firm case. Thus, $CS_N^* = CS_1^* < CS^0$, as proven in Proposition 4.

If $N \geq \frac{\underline{V}'}{\alpha}$, we have an interior solution with N innovators, and

$$P_N^* = \frac{N}{N+1} (\underline{V}' + \alpha) > \underline{V}' = P_1^*.$$

Having already proved that $CS^0 > CS_1^*$, if we can prove that $CS_1^* > CS_N^*$, we will establish the result.

Given that (by assumption) there are no uninsured buyers of the treatment and given (10) then under expanded insurance

$$\begin{aligned} CS_N^* &= \frac{\rho_C}{\rho_I} (B - C) D^*(P_N^*) + (\underline{V} + \alpha - P_N^*) D^*(P_N^*) - \frac{\alpha}{2} D^*(P_N^*)^2 \\ &= (\underline{V}' + \alpha - P_N^*) D^*(P_N^*) - \frac{\alpha}{2} D^*(P_N^*)^2. \end{aligned}$$

Since we are in the case in which the demand curve for the treatment is the value constraint, $D^*(P)$ is given by (28) above. Thus for $P \in (\underline{V}', \underline{V}' + \alpha)$ consider the consumer surplus function

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

Differentiating with respect to P and noting that $D^*(P) = \frac{V'+\alpha}{\alpha} - \frac{P}{\alpha}$ for $P \in (\underline{V}', \underline{V}' + \alpha)$ implies that

$$\frac{dCS^*(P)}{dP} = -D^*(P) < 0, P \in (\underline{V}', \underline{V}' + \alpha). \quad (29)$$

Since $P_1^* < P_N^*$ when $N \geq \frac{V'}{\alpha}$, we have, light of (29),

$$CS_1^* = CS^*(P_1^*) > CS^*(P_N^*) = CS_N^*.$$

Thus,

$$CS^0 > CS_N^*.$$

■

Proof of Proposition 7:

The \hat{N} we refer to in the statement of the proposition is

$$\hat{N} = \max \left\{ N^*, \frac{2}{\alpha} \frac{V' + \alpha}{W + 1} - 1 \right\},$$

where, recall, $N^* = \frac{V'}{\alpha}$.

Using (9), let $TS^*(Q) = TS(Q, Q) = (V' + \alpha)Q - \frac{\alpha}{2}(Q)^2$. Thus

$$TS^0 = TS(Q^0, 1) \geq TS(Q^0, Q^0) = TS^*(Q^0) = (V' + \alpha)Q^0 - \frac{\alpha}{2}(Q^0)^2,$$

and

$$TS_N^* = TS^*(Q_N^*) = (V' + \alpha)Q_N^* - \frac{\alpha}{2}(Q_N^*)^2.$$

Now,

$$\frac{dTS^*}{dQ} = V' + \alpha - \alpha Q > 0,$$

i.e., $TS^*(Q)$ increases in Q over the relevant range $Q \in [0, 1]$.

If $\underline{W} \geq 1$, $Q^0 = 1$, and as we saw in the proof of Proposition 5, if $N \geq N^* = \frac{V'}{\alpha}$, then $Q_N^* < 1$. Since, by assumption, $N > \hat{N}$, then $N > N^*$, so indeed, $Q_N^* < Q^0 = 1$, and thus,

$$TS_N^* = TS^*(Q_N^*) < TS^*(Q^0) = TS^0.$$

If $\underline{W} < 1$, $Q^0 = \frac{W+1}{2}$. Since $N > \hat{N}$, then $N > N^*$ and $N > \frac{2}{\alpha} \frac{V'+\alpha}{W+1} - 1$. Together, these imply $Q_N^* = \frac{1}{\alpha} \frac{V'+\alpha}{N+1} < \frac{W+1}{2} = Q^0$, and thus

$$TS_N^* = TS^*(Q_N^*) < TS^*(Q^0) < TS^0.$$

■

Proof of Proposition 8:

Preliminaries: Consider the optimization problem

$$\begin{aligned} \max PQ \\ \text{s.t. } P &\leq \frac{W+1-Q}{\rho_I} \end{aligned} \quad (30)$$

$$P \leq \nu + \alpha - \alpha Q \quad (31)$$

$$Q \leq 1 \quad (32)$$

$$Q \geq 0$$

When $\nu = \underline{V}$, we have the innovator's profit-maximization problem when consumers have a choice of health plans, and when $\nu = \underline{V}' = \underline{V} + \frac{\rho_C(B-C)}{\rho_I}$, we have the innovator's profit-maximization problem when the only health insurance plan offered is the expanded plan. Thus, the two problems differ only in the level of

ν . Let $\widehat{Q}(\nu)$ and $\widehat{P}(\nu)$ be the solution to this problem. The approach of the proof is to show that $\widehat{P}(\nu)$ is non-decreasing in ν , and since $\underline{V}' \geq \underline{V}$, it would follow that the price of the treatment under a MCS is at least as high as without MCS.

To characterize that solution to the profit-maximization problem, define the quantity $Q_K(\nu)$ to be the solution for Q to

$$\frac{W+1-Q}{\rho_I} = \nu + \alpha - \alpha Q,$$

which implies

$$Q_K(\nu) = \frac{\frac{W+1}{\rho_I} - (\nu + \alpha)}{\frac{1}{\rho_I} - \alpha},$$

and thus (recalling $\rho_I < 1 < \frac{1}{\alpha}$), $Q_K(\nu)$ strictly decreases in ν . Let $P_K(\nu) = \frac{W+1-Q_K(\nu)}{\rho_I} = \nu + \alpha - \alpha Q_K(\nu)$ be the corresponding price:

$$P_K(\nu) = \frac{\nu - \alpha W}{1 - \rho_I \alpha}.$$

Characterizing the solution to the profit-maximization problem: In the profit-maximization problem, we can ignore the constraint $Q \geq 0$ as it will end up being satisfied for all $\nu > 0$. Letting λ_1 , λ_2 , and μ be the Lagrange multipliers for (30), (31), and (32), respectively and letting

$$\mathcal{L} = PQ - \lambda_1 \left[P - \left(\frac{W+1-Q}{\rho_I} \right) \right] - \lambda_2 [P - (\nu + \alpha - \alpha Q)] - \mu [Q - 1].$$

be the Lagrangian, the Kuhn-Tucker conditions are

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial P} &= 0 \Rightarrow Q = \lambda_1 + \lambda_2. \\ \frac{\partial \mathcal{L}}{\partial Q} &= 0 \Rightarrow P = \frac{\lambda_1}{\rho_I} + \alpha \lambda_2 + \mu. \\ \lambda_1 &\geq 0; P \leq \frac{W+1-Q}{\rho_I}; \lambda_1 \left[P - \left(\frac{W+1-Q}{\rho_I} \right) \right] = 0. \\ \lambda_2 &\geq 0; P \leq \nu + \alpha - \alpha Q; \lambda_2 = 0; [P - (\nu + \alpha - \alpha Q)] = 0. \\ \mu &\geq 0; Q \leq 1; \mu [Q - 1] = 0. \end{aligned}$$

Now, let Θ_1 and Θ_2 be the values ν that solve

$$Q_K(\nu) = \frac{\nu + \alpha}{2\alpha}$$

and

$$Q_K(\nu) = \frac{W+1}{2},$$

respectively. This gives us

$$\begin{aligned} \Theta_1 &= \frac{\frac{W+1}{\rho_I}}{\frac{1}{2\alpha} \left(\frac{1}{\rho_I} - \alpha \right) + 1} - \alpha. \\ \Theta_2 &= \frac{W+1}{2\rho_I} + \alpha \left(\frac{W-1}{2} \right). \end{aligned}$$

Straightforward analysis of the Kuhn-Tucker conditions implies that six possible configurations of the solution depending on whether the solution occurs along the liquidity constraint, the value constraint, or at the kink where they intersect, and whether we have an interior or corner solution for Q . These configurations depend on ν :

Case 1: $\nu \geq \Theta_2$ and $\underline{W} < 1$:

$$\begin{aligned}\widehat{Q}(\nu) &= \frac{W+1}{2} = Q^0 < 1. \\ \widehat{P}(\nu) &= \frac{W+1}{2\rho_I} = \frac{P^0}{\rho_I}.\end{aligned}$$

In this case, we have an interior solution to the profit-maximization problem that occurs on the liquidity constraint. The value constraint does not bind. We get the same quantity we got under basic insurance, with the price grossed up by $\frac{1}{\rho_I}$ relative to that case, assuming that that quantity is interior.

Case 2: $\nu \geq \frac{W}{\rho_I}$ and $\underline{W} \geq 1$:

$$\begin{aligned}\widehat{Q}(\nu) &= Q^0 = 1. \\ \widehat{P}(\nu) &= \frac{W}{\rho_I} = \frac{P^0}{\rho_I}.\end{aligned}$$

Here we have a corner solution to the profit-maximization problem that occurs on the liquidity constraint.

Case 3: $\nu \in [\Theta_1, \Theta_2]$ and $\nu \geq \frac{W}{\rho_I}$:

$$\begin{aligned}\widehat{Q}(\nu) &= \widehat{Q}_K(\nu) < 1. \\ \widehat{P}(\nu) &= \widehat{P}_K(\nu).\end{aligned}$$

In this case, profit-maximization occurs at the kink at which the liquidity constraint and value constraint cross, and the corresponding quantity is interior.

Case 4: $\nu = \frac{W}{\rho_I}$:

$$\begin{aligned}\widehat{Q}(\nu) &= \widehat{Q}_K(\nu) = 1 \\ \widehat{P}(\nu) &= \widehat{P}_K(\nu).\end{aligned}$$

This is the same as the previous case, but with the intersection of the value and liquidity constraints occurring at $Q = 1$. (In other words, the corner solution occurs at the kink.)

Case 5: $\nu \leq \Theta_1$ and $\nu < \alpha$:

$$\begin{aligned}\widehat{Q}(\nu) &= \frac{\nu + \alpha}{2\alpha}. \\ \widehat{P}(\nu) &= \frac{\nu + \alpha}{2}.\end{aligned}$$

In this case, we have an interior solution to the profit-maximization problem that occurs on the value constraint. The liquidity constraint does not bind.

Case 6: $\nu \geq \alpha$ and $\nu < \frac{W}{\rho_I}$:

$$\begin{aligned}\widehat{Q}(\nu) &= 1. \\ \widehat{P}(\nu) &= \nu.\end{aligned}$$

Here we have a corner solution to the profit-maximization problem that occurs on the value constraint.

Lemmas to help characterize the solution to the profit maximization problem: We now prove five results that, together, will help pin down how we “move” through these cases as ν increases or decreases.

Lemma A: If $\underline{W} < 1$, then $\frac{W}{\rho_I} < \Theta_2$.

Proof:

Note that $\widehat{Q}_K(\Theta_2) = \frac{W+1}{2} < 1$, when $\underline{W} < 1$. Moreover,

$$\widehat{Q}_K\left(\frac{W}{\rho_I}\right) = \frac{\frac{W+1}{\rho_I} - \left(\frac{W}{\rho_I} + \alpha\right)}{\frac{1}{\rho_I} - \alpha} = 1.$$

Because $\widehat{Q}_K(\frac{W}{\rho_I}) > \widehat{Q}_K(\Theta_2)$ and $\widehat{Q}_K(\nu)$ strictly decreases in ν , it follows that $\frac{W}{\rho_I} < \Theta_2$. ■

Lemma B: $\Theta_1 < \Theta_2$.

Proof: To establish this, suppose to the contrary that $\Theta_1 \geq \Theta_2$. Then

$$\frac{\nu + \alpha}{2\alpha} \Big|_{\nu=\Theta_2} \leq \frac{\nu + \alpha}{2\alpha} \Big|_{\nu=\Theta_1} = Q_K(\Theta_1).$$

But also, since $Q_K(\nu)$ is strictly decreasing in ν , we have (given $\Theta_1 \geq \Theta_2$)

$$Q_K(\Theta_1) \leq Q_K(\Theta_2) = \frac{W+1}{2}.$$

Together, these inequalities imply $\frac{\Theta_2 + \alpha}{2\alpha} \leq \frac{W+1}{2}$, or

$$\Theta_2 \leq \alpha W.$$

But note that

$$\begin{aligned} \Theta_2 &= \frac{W}{2\rho_I} + \alpha \frac{W}{2} + \frac{1}{2} \left(\frac{1}{\rho_I} - \alpha \right) \\ &> \alpha \frac{W}{2} + \alpha \frac{W}{2} = \alpha W, \end{aligned}$$

where the inequality follows because $\frac{1}{\rho_I} > 1 > \alpha$. The contrapositive assumption that $\Theta_1 \geq \Theta_2$ yields a contradiction, which implies $\Theta_1 < \Theta_2$. ■

Lemma C: $\alpha > \Theta_1$ if and only if $\frac{W}{\rho_I} < \Theta_1$.

Proof: We first prove $\alpha > \Theta_1 \Rightarrow \frac{W}{\rho_I} < \Theta_1$. By definition of Θ_1 and $Q_K(\nu)$

$$\frac{\nu + \alpha}{2\alpha} \Big|_{\nu=\Theta_1} = Q_K(\Theta_1).$$

Moreover, we have

$$\frac{\nu + \alpha}{2\alpha} \Big|_{\nu=\alpha} = 1 = Q_K\left(\frac{W}{\rho_I}\right).$$

Since $\alpha > \Theta_1$ and $\frac{\nu + \alpha}{2\alpha}$ increases in ν , it follows that

$$Q_K\left(\frac{W}{\rho_I}\right) > Q_K(\Theta_1),$$

and since $Q_K(\nu)$ is strictly decreasing in ν , $\frac{W}{\rho_I} < \Theta_1$. The converse implication is proven through the same steps of logic. ■

Lemma D: If $\alpha = \Theta_1$ then $\frac{W}{\rho_I} = \Theta_1$.

Proof: If $\alpha = \Theta_1$ then

$$Q_K(\Theta_1) = \frac{\Theta_1 + \alpha}{2\alpha} = \frac{\alpha + \alpha}{2\alpha} \Big|_{\nu=\alpha} = 1 = Q_K\left(\frac{W}{\rho_I}\right).$$

■

Lemma E: If $W > 1$, then $\frac{W}{\rho_I} > \Theta_2$.

Proof: The result follows because

$$1 = Q_K\left(\frac{W}{\rho_I}\right) < \frac{W+1}{2} = Q_K(\Theta_2),$$

and $Q_K(\nu)$ is monotone decreasing. ■

Characterizing the profit-maximizing price: Taken together, these results restrict the ordering of Θ_1 , Θ_2 , $\frac{W}{\rho_I}$, and α that determine which of the above cases we “transition through” as we vary ν , which in turn can

allow us to characterize the profit-maximizing price. First, suppose that $\underline{W} < 1$. The following parameter orderings are possible:

- $\frac{W}{\rho_I} < \Theta_1 < \Theta_2$ and $\Theta_1 < \alpha$.
- $\frac{W}{\rho_I} = \Theta_1 < \Theta_2$ and $\Theta_1 = \alpha$.
- $\Theta_1 < \frac{W}{\rho_I} < \Theta_2$ and $\Theta_1 > \alpha$.

Consider increasing ν from 0. With the first and second orderings, we move from Case 5 (an interior solution along the value constraint) to Case 3 (an interior solution at the kink of the value and liquidity constraints) and (since $\underline{W} < 1$) to Case 1 (an interior solution along the liquidity constraint). In this case, the profit-maximizing price can be summarized as

$$\hat{P}(\nu) = \begin{cases} \frac{W+1}{2\rho_I} & v \geq \Theta_2 \\ \hat{P}_K(\nu) & \nu \in [\Theta_1, \Theta_2] \\ \frac{\nu+\alpha}{2} & \nu \leq \Theta_1 \end{cases} .$$

This function is continuous and non-decreasing in ν .

With the third ordering, as ν increases from 0 we move from Case 5 to Case 6 (a corner solution along the value constraint) to Case 4 (a corner solution coinciding with the kink of the value and liquidity constraints) to Case 3 to Case 1. In this case, we have

$$\hat{P}(\nu) = \begin{cases} \frac{W+1}{2\rho_I} & v \geq \Theta_2 \\ \hat{P}_K(\nu) & \nu \in [\frac{W}{\rho_I}, \Theta_2] \\ \nu & \nu \in [\alpha, \frac{W}{\rho_I}) \\ \frac{\nu+\alpha}{2} & \nu \leq \alpha \end{cases} .$$

This function is also continuous and non-decreasing in ν .

Now, let's turn to the case in which $\underline{W} \geq 1$. Given Lemmas E, B, and C, we have the following ordering:

$$\alpha \leq \Theta_1 < \Theta_2 < \frac{W}{\rho_I} .$$

In this case there is just one possible solution configuration, which can be most easily seen by decreasing ν from a high level, specifically $\nu \geq \frac{W}{\rho_I}$. We thus start (working backward) in Case 2, where the profit-maximization problem has a corner solution along the liquidity constraint. Since quantity Q cannot go above 1, we remain at a corner solution until reaching the range of v at which Case 6 obtains, where we have a corner solution along the value constraint. We continue in this case until ν becomes sufficiently small for Case 5 to arise, where we have an interior solution along the liquidity constraint. In this case, then, the profit-maximizing price is

$$\hat{P}(\nu) = \begin{cases} \frac{W}{\rho_I} & v \geq \frac{W}{\rho_I} \\ \nu & \nu \in [\frac{W}{\rho_I}, \alpha] \\ \frac{\nu+\alpha}{2} & \nu \leq \alpha \end{cases} .$$

This function is continuous and non-decreasing in ν . Thus, for all parameter combinations that are independent of ν , the profit-maximizing price $\hat{P}(\nu)$ is non-decreasing in ν . It then follows that

$$\hat{P} = \hat{P}(\underline{V}) \leq P^*(\underline{V}') = P^* .$$

Finally, we note that when $\underline{V}' \leq \Theta_1$ and $\underline{V}' \leq \alpha$ (and so $\underline{V} < \Theta_1$ and $\underline{V} < \alpha$), we have

$$\hat{P} = \frac{\underline{V} + \alpha}{2} \leq \frac{\underline{V}' + \alpha}{2} = P^* ,$$

establishing that the price of the treatment is strictly higher under a MCS when, in that case, price is determined by the value constraint. ■

References

- [1] Acemoglu, Daron and Joshua Linn (2004). “Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry,” *Quarterly Journal of Economics*, Vol. 119, No. 3, pp. 1049-1090
- [2] Bagley, Nicholas, Amitabh Chandra and Austin Frakt (2015). “Correcting Signals for Innovation in Health Care,” Hamilton Project Discussion Paper #2015-09.
- [3] Becker, Gary (2007). “Health as Human Capital: Synthesis and Extensions,” *Oxford Economic Papers*, Vol. 59, No. 3, pp. 379-410.
- [4] Berkrot, Bill (2015). “Gilead Put Profit Ahead of Hepatitis C Patients: U.S. Senate report,” *Reuters*, (December 1).
- [5] Blume-Kohout, Margaret and Neeraj Sood, Neeraj (2013). “Market Size and Innovation: Effects of Medicare Part D on Pharmaceutical Research and Development,” *Journal of Public Economics*, Vol. 97, pp. 327-336.
- [6] Chhatwal, Jagpreet, Fasiha Kanwal, Mark S. Roberts, and Michael A. Dunn (2015). “Cost-Effectiveness and Budget Impact of Hepatitis C Virus Treatment With Sofosbuvir and Ledipasvir in the United States,” *Annals of Internal Medicine*, Vol. 162, No. 6, pp. 397-406.
- [7] Cournot, Augustin (1927) *Researches into the Mathematical Principles of the Theory of Wealth*, translated by Nathaniel T. Bacon (New York: Macmillan).
- [8] Dranove, David, Craig Garthwaite and Manuel Hermsillo (2014), “Pharmaceutical Profits and the Social Value of Innovation,” NBER Working Paper #20212.
- [9] Duggan, Mark and Fiona Scott-Morton (2010), “The Effect of Medicare Part D on Pharmaceutical Prices and Utilization,” *American Economic Review*, Vol. 100, No. 1, pp. 590-607.
- [10] Einav, Liran, Amy Finkelstein, Stephen P. Ryan, Paul Schrimpf, and Mark R. Cullen (2013). “Selection on Moral Hazard in Health Insurance,” *American Economic Review*, Vol. 103, No. 1, pp. 178-219.
- [11] Ellison, Sara and Catherine Wolfram (2006) “Coordinating on Lower Prices: Pharmaceutical Pricing under Political Pressure.” *The RAND Journal of Economics*, Vol. 37, No. 2, pp. 324-340.
- [12] Engelhardt, Gary V. and Jonathan Gruber (2011). “Medicare Part D and the Financial Protection of the Elderly,” *American Economic Journal: Economic Policy*, Vol. 3, No. 4, pp. 77-102.
- [13] Feldman, Roger and Bryan Dowd (1991). “A New Estimate of the Welfare Loss of Excess Health Insurance,” *American Economic Review*, Vol. 81, No. 1, pp. 531–537.
- [14] Feldstein, Martin (1973). “The Welfare Loss of Excess Health Insurance,” *Journal of Political Economy*, Vol 61, pp. 251–280.
- [15] Finkelstein, Amy (2004). “Static and Dynamic Effects of Health Policy: Evidence from the Vaccine Industry,” *Quarterly Journal of Economics*, Vol. 119, No. 2, pp. 527-564.

- [16] Finkelstein, Amy, Erzo Luttmer and Nathan Hendren (2015), "The Value of Medicaid: Interpreting Results from the Oregon Health Insurance Experiment," NBER Working Paper #21308.
- [17] Friedman, Bernad (1974) "Risk Aversion and the Consumer Choice of Health Insurance Option," *Review of Economics and Statistics*, Vol 56, No. 2, pp. 209-214.
- [18] Fuchs, Willi, Brett Green and David Levine. 2013. "Optimal Arrangements for Distribution in Developing Markets," working paper, http://faculty.haas.berkeley.edu/bgreen/files/FGL_13.pdf (accessed January 16, 2016).
- [19] Garthwaite, Craig, Tal Gross, and Matthew Notodiwidgo (2015), "Hospitals as Insurers of Last Resort," NBER Working Paper #21290.
- [20] Hirsch, Barry, Suresh Balu, and Kevin Shulman (2014). "The Impact Of Specialty Pharmaceuticals As Drivers Of Health Care Costs," *Health Affairs*, Vol 33. No. 10, pp. 1714-1720.
- [21] Howard, David H., Peter B. Bach, Ernst R. Berndt, and Rena M. Conti (2015). "Pricing in the Market for Anticancer Drugs," *Journal of Economic Perspectives*, Vol. 29, No. 1, pp. 139-162.
- [22] Kardish, Chris (2014). "The Risky Business of Limiting Medicaid Access to Sovaldi," *Governing.com*, <http://www.governing.com/topics/health-human-services/gov-hepatitis-coverage-solvaldi-lawsuits.html> (accessed December 22, 2015).
- [23] Kaczmarek, Stephen J. (2014) "The Impact of New Hepatitis C Drug Therapy on Individual Medicare Part D Spending," <http://www.pmanet.org/images/stories/uploads/2014/partdpremiumstudymilliman.pdf> (accessed July 9, 2015).
- [24] Liu, Shan, Daena Watcha, Mark Holodniy, and Jeremy Goldhaber-Fiebert (2015). "Sofosbuvir-Based Treatment Regimens for Chronic, Genotype 1 Hepatitis C Virus Infection in U.S. Incarcerated Populations: A Cost-Effectiveness Analysis," *Annals of Internal Medicine* Vol. 161, No 8, pp. 546-553.
- [25] Loftus, Peter (2015). "How Much Should Cancer Drugs Cost," *Wall Street Journal* (June 18).
- [26] Mahoney, Neale (2015). "Bankruptcy as Implicit Health Insurance," *American Economic Review*, Vol. 105, No. 2, pp. 710-746.
- [27] Managan, Dan (2015). "Pricey New Cholesterol Rx Covered by Big Drug Plan, but...," *CNBC.com*, <http://www.cnbc.com/2015/10/06/pricy-new-cholesterol-rx-covered-by-big-drug-plan-but.html> (accessed April 4, 2016).
- [28] Manning, Will and M. Susan Marquis, (1996). "Health Insurance: the Tradeoff Between Risk Pooling and Moral Hazard," *Journal of Health Economics*, Vol. 15, No. 5, pp. 609-640.
- [29] Najafzadeh, Mehdi, Karin Andersson, William H. Shrank, Alexis A. Krumme, Olga S. Matlin, Troyen Brennan, Jerry Avorn, MD, and Niteesh K. Choudhry (2015). "Cost-Effectiveness of Novel Regimens for the Treatment of Hepatitis C Virus," *Annals of Internal Medicine*, Vol. 162, No. 6, pp.407-419.
- [30] Newhouse, Joseph (1993). *Free for All? Lessons from the RAND Health Insurance Experiment*. (Cambridge, MA: Harvard University Press).

- [31] Nyman, John A. (1998). "Theory of Health Insurance," *Journal of Health Administration Education*, Vol. 16, No. 1, pp.41-66.
- [32] Nyman, John A. (1999). "Health Insurance: the Access Motive," *Journal of Health Economics*, Vol. 18, No. 2, pp.141-52.
- [33] Pauly, Mark (1968). "The Economics of Moral Hazard: Comment," *American Economic Review*, Vol. 58, No. 3, pp. 531-537.
- [34] Pfeiffer, Stuart (2015). "Woman Sues Anthem Blue Cross for Refusing to Cover Hepatitis C Drug," *Los Angeles Times* (May 18).
- [35] Pollack, Andrew (2014). "Gilead's Hepatitis C Drug, Sovaldi, Is on Pace to Become a Blockbuster," *New York Times* (July 23).
- [36] Pollack, Andrew (2015). "Drug Goes from \$13.50 a Table to \$750, Overnight," *New York Times* (September 20).
- [37] Rockoff, Jonathan and Ed Silverman (2015). "Pharmaceutical Companies Buy Rivals' Drugs, Then Jack Up the Prices," *Wall Street Journal* (April 26).
- [38] Viscusi, W. Kip and Joseph E. Aldy (2003) "The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World," *Journal of Risk and Uncertainty*, Vol. 27, No. 1, pp 5-76.
- [39] Walker, Joseph (2015). "High Prices for Drugs Attacked at Meeting," *Wall Street Journal* (June 1).