

NBER WORKING PAPER SERIES

MEASURING EX-ANTE WELFARE IN INSURANCE MARKETS

Nathaniel Hendren

Working Paper 24470

<http://www.nber.org/papers/w24470>

NATIONAL BUREAU OF ECONOMIC RESEARCH

1050 Massachusetts Avenue

Cambridge, MA 02138

March 2018

I am very grateful to Raj Chetty, David Cutler, Liran Einav, Amy Finkelstein, Mark Shepard, Mike Whinston, along with seminar participants at the NBER Summer Institute and University of Texas for helpful comments and discussions. Support from the National Science Foundation CAREER Grant is gratefully acknowledged. The views expressed herein are those of the author 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 They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2018 by Nathaniel Hendren. 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.

Measuring Ex-Ante Welfare in Insurance Markets

Nathaniel Hendren

NBER Working Paper No. 24470

March 2018

JEL No. H0,I11,I3

ABSTRACT

Revealed-preference measures of willingness to pay (WTP) capture the value of insurance only against the risk that remains when choosing insurance. This paper provides a method to translate observed market WTP and cost curves into an ex-ante value of insurance that can analyze the impact of insurance market policies on ex-ante expected utility. The key additional statistic required is the difference in marginal utilities between insured and uninsured, which generally requires an estimate of risk aversion. Applying the approach to previous literature, I estimate higher values of subsidies and mandates relative to methods based market surplus or "deadweight loss".

Nathaniel Hendren

Harvard University

Department of Economics

Littauer Center Room 235

Cambridge, MA 02138

and NBER

nhendren@gmail.com

Measuring Ex-Ante Welfare in Insurance Markets

Nathaniel Hendren*

March, 2018

Abstract

Revealed-preference measures of willingness to pay (WTP) capture the value of insurance only against the risk that remains when choosing insurance. This paper provides a method to translate observed market WTP and cost curves into an ex-ante value of insurance that can analyze the impact of insurance market policies on ex-ante expected utility. The key additional statistic required is the difference in marginal utilities between insured and uninsured, which generally requires an estimate of risk aversion. Applying the approach to previous literature, I estimate higher values of subsidies and mandates relative to methods based market surplus or "deadweight loss".

1 Introduction

There is a large and growing literature using price variation and other reduced-form methods to estimate individuals' willingness to pay for insurance and the costs they impose on the insurer. These are then used to assess optimal insurance subsidies and mandates by comparing individuals' willingness to pay

*Harvard University, nhendren@fas.harvard.edu. I am very grateful to Raj Chetty, David Cutler, Liran Einav, Amy Finkelstein, Mark Shepard, Mike Whinston, along with seminar participants at the NBER Summer Institute and University of Texas for helpful comments and discussions. Support from the National Science Foundation CAREER Grant is gratefully acknowledged.

to their costs, often termed market surplus or deadweight loss (Einav et al. (2010); Hackmann et al. (2015)). If the marginal individual is not willing to pay the cost s/he imposes on the insurer, then additional subsidies lower market surplus and are presumed to be socially undesirable.

In most economic settings, measures of willingness to pay that are directly revealed through observed choices are the gold standard input into welfare analysis. However, in insurance settings they can be misleading. Insurance obtains its value by insuring the realization of risk. Often individuals choose insurance products after some information about their risk has already been revealed. It is widely appreciated that this can lead to adverse selection. What is arguably less well-appreciated is that their willingness to pay will not capture their value insurance against the portion of the risk that has already been revealed when they make their choice.¹ Thus, in settings with adverse selection, policies that maximize market surplus will not generally maximize ex-ante expected utility.

To see this, consider the decision to purchase health insurance. Suppose some people making this choice have learned they have cancer. They will potentially be willing to pay more for health insurance, both because they have a higher expected cost (which will lead to adverse selection) and also because they want to insure against remaining uncertainty in the cost of chemotherapy, radiation treatments, etc. Their market surplus from insurance (the difference between their willingness to pay and costs) will capture the value of insuring their remaining risk (e.g. uncertainty in treatment costs). But, it will miss the value that health insurance provides against the risk of getting cancer, which leads to higher expected medical costs. In this sense, market surplus tends to understate the value of subsidies and mandates from an ex-ante expected utility perspective when markets suffer from adverse selection.

This goal of this paper is to provide a transparent empirical method to measure the impact of insurance market policies, such as subsidies and mandates, on individuals' ex-ante expected utility. To do so, I build on the framework

¹This point is arguably first shown in the classic work of (Hirshleifer (1971)). When individuals have some knowledge about their risk when choices are made, the average willingness to pay for insurance will generally be less than the ex-ante willingness to pay for insurance

of Einav et al. (2010) that estimates market willingness to pay (or “demand”) and cost curves for insurance from variation in prices or choice sets. I provide additional sufficient statistics that enable the researcher to translate observed willingness to pay curves into the ex-ante willingness to pay individuals would have had if the researcher observed willingness to pay prior to when those in the market learn about their own risk.

To measure this ex-ante willingness to pay, I show that one can combine information in the market willingness to pay and cost curves with one additional sufficient statistic: the difference in marginal utilities of income for those who do versus do not choose to purchase insurance. The ex-ante value of insurance is higher when individuals from behind the veil of ignorance wish to move money from the state of the world in which they forego insurance to the state of the world in which they choose to purchase insurance. In the cancer example above, to the extent to which the presence of cancer is a determinant of insurance purchase, this difference in marginal utilities would reflect the extent to which health insurance provides insurance against the risk of getting cancer.

To implement the approach, one needs to know the difference in marginal utilities between the insured and uninsured. I provide assumptions under which one can estimate this difference using the market willingness to pay and cost curves combined with a measure of risk aversion.² The measure of risk aversion can be assumed, or it can potentially be inferred from the willingness to pay and cost curves.

I apply the approach to the study of health insurance subsidies for low-income adults, using willingness to pay and cost curve estimates from Finkelstein et al. (2017), and to study employer-provided top-up insurance using estimates from Einav et al. (2010). In the case of low-income health insurance, the insurance prices that maximize expected utility are 30% lower than those that maximize market surplus. For plausible specifications, imposing a mandate can lower market surplus but increase ex-ante expected utility. In this sense, market surplus can be misleading as a guide to optimal public

²This method for approximating differences in marginal utilities is analogous to using consumption changes combined with risk aversion to infer willingness to pay for unemployment insurance (Baily (1978); Chetty (2006)).

health insurance policies for low-income adults.

In the top-up insurance setting of employer-provided health insurance in Einav et al. (2010), the distinction between policies that ex-ante expected utility and those that maximize observed market surplus are less pronounced. This reflects the fact that the difference between market surplus and ex-ante welfare is larger when the risk reflects a larger portion of individuals' budgets. Because Einav et al. (2010) considers a top-up insurance policy of more versus less generous insurance, the size of the insurable risk is smaller. Therefore, there is a smaller divergence between ex-ante expected utility measures of welfare and market surplus.

The rest of this paper proceeds as follows. To illustrate the distinction between ex-ante welfare and observed willingness to pay, I start with a stylized example in Section 2. Section 3 then provides a general characterization of the main result that the ex-ante willingness to pay for insurance requires the difference in marginal utilities between insured and uninsured. Section 4 provides a method to estimate this difference in marginal utilities using market demand and cost curves combined with a measure of risk aversion. Section 5 provides the application to the study of health insurance subsidies for low-income adults in Finkelstein et al. (2017) and employer-provided top-up insurance in Einav et al. (2010). Section 6 concludes.

2 Stylized Example

I begin with an example that illustrates the problem with using market surplus as a normative guide and outlines the proposed solution. Suppose individuals have \$30 dollars but face a risk of losing \$ m dollars, where m is uniformly distributed between 0 and 10. Consider individuals' willingnesses to pay for insurance against the realization of m that is measured before they have any particular knowledge about their risk of loss. Let $D^{Ex-ante}$ denote this willingness to pay or "demand". This solves

$$u(30 - D^{Ex-ante}) = E[u(30 - m)] \tag{1}$$

where $E[u(30 - m)] = \frac{1}{10} \int_0^{10} u(30 - m) dm$ is the expected utility if uninsured. Suppose individuals have a utility function with a constant coefficient of relative risk aversion of 3 (i.e. $u(c) = \frac{1}{1-\sigma} c^{1-\sigma}$ and $\sigma = 3$). In this case, it is straightforward to compute that they are willing to pay $D^{Ex-ante} = 5.50$ for insurance against m . This insurance policy would cost the insurer $E[m] = 5$, so that the individuals are willing to pay a markup of 0.50 over actuarially fair insurance. Full insurance generates a market surplus of \$0.50.

Figure 1: Example Willingness to Pay and Cost Curves

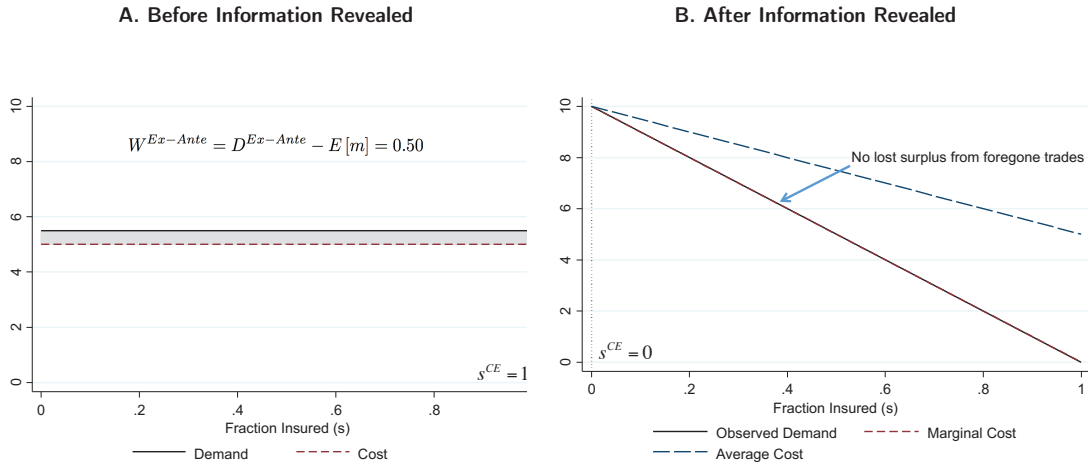


Figure 1, Panel A, translates this scenario using the demand and cost curve framework formalized in Einav et al. (2010). The horizontal axis enumerates the population in descending order of their willingness to pay for insurance (using an index $s \in [0, 1]$), and the vertical axis reflects prices, costs, and willingness to pay in the market. Each individual is willing to pay \$5.50 for insurance, generating a flat willingness to pay, or demand, curve of $D(s) = \$5.50$. Because no one knows anything about their particular cost, each individual imposes a cost of \$5 on the insurance company, generating a flat cost curve of $C(s) = \$5$. If a competitive market were to open up in this setting, one would expect everyone ($s^{CE} = 100\%$) to purchase insurance at a price of \$5. This allocation would generate $W^{Ex-Ante} = \$0.50$ of welfare, as reflected by the market surplus defined as the integral between demand and cost curve.

What happens if individuals learn some information about their costs before they choose whether to purchase insurance? For simplicity, consider the extreme case that individuals have fully learned their cost, m . Willingness to pay will equal individuals' known costs, $D(s) = m(s)$. Those who learn they will lose \$10 will be willing to pay \$10 for "insurance" against their loss; individuals who learn they will lose \$0 will be willing to pay nothing. The uniform distribution of risks generates a linear demand curve falling from \$10 at $s = 0$ to \$0 at $s = 1$. The cost imposed on the insurer by the marginal type s , $C(s)$, will equal their willingness to pay of $D(s)$. Therefore, the demand curve lies everywhere on top of the cost curve of the marginal types, as illustrated in Panel B.

If an insurer were to try to sell insurance, they would need to set prices to cover the average cost of those who purchase insurance. Let $AC(\tilde{s}) = E[C(s) | s \leq \tilde{s}]$ denote the average cost of those with willingness to pay above $D(\tilde{s})$. This average cost lies everywhere above the demand curve. Since no one is willing to pay the pooled cost of those with higher willingness to pay, the market would fully unravel. The unique competitive equilibrium would involve no one obtaining any insurance, $s^{CE} = 0\%$.

What is the welfare cost of this market unraveling? From a market surplus perspective, there is no welfare loss. Because the demand curve equals the cost curve, there are no valuable foregone trades that can take place at the time insurance choices are made. This reflects an extreme case of a more general phenomenon identified in Hirshleifer (1971). The market demand curve does not capture the value of insurance against the portion of risk that has already been realized at the time insurance choices are made. This means that policies that maximize market surplus may not maximize ex-ante expected utility.

How can one recover the ex-ante expected utility measure of welfare, $D^{Ex-Ante}$, in equation (1)? The traditional approach would require the econometrician to specify economic primitives including a utility function and information set. It would then also require measuring outcomes such as consumption (or assume proxies for consumption) to infer the ex-ante value of insurance from the model. Intuitively, if one knows the utility function, u , and the cross-sectional distribution of consumption ($30 - m$ in the example above), then one can use this information to compute $D^{Ex-Ante}$ in equation (1). For recent implemen-

tations of this approach, see Handel et al. (2015), Section IV of Einav et al. (2016), or Finkelstein et al. (2016).

In contrast, the goal of this paper is to estimate $D^{Ex-Ante}$ without knowledge of the full distribution of primitives (e.g. u and m). I also want to evaluate the ex-ante willingness to pay for subsidies and mandates that lead to market outcomes with only a fraction of the market choosing to purchase insurance, $s < 1$, in addition to the value of full insurance ($s = 1$) given by $D^{Ex-Ante}$ in equation (1).

To do so, let p_I denote the price of insurance and p_U denote the price of being uninsured (so that $p_I - p_U$ is the marginal price of obtaining insurance). Consider the willingness to pay for a larger insurance market using a budget-neutral shift in insurance prices that requires the total amount of money collected to equal the total cost of the insured, $sp_I + (1 - s)p_U = sAC(s)$. Budget neutrality is not essential for the approach. For non-budget neutral policies that use government funds to subsidize insurance, Section 3 shows how to construct an ex-ante measure of the marginal value of public funds (MVPF), as in Hendren (2016), for spending government resources on health insurance subsidies.

Suppose that prices are set such that a fraction $s = 0.5$ of the population chooses to purchase insurance, as illustrated in Figure 2, Panel A. It is straightforward to show that this corresponds to $p_I = 6.25$ and $p_U = 1.25$, so that the marginal price of insurance is \$5. Now, consider expanding the size of the insurance market from $s = 0.5$ to $0.5 + ds$ by decreasing p_I financed by an increase in p_U . This lowers the marginal price of insurance, $p_I - p_U$, by $D'(s)ds$. The resource constraint implies that the price faced by the uninsured increases by $dp_U = -sD'(s)ds$, and the price of insurance must decrease by $dp_I = (1 - s)D'(s)ds$.³

This change in insurance prices generates a transfer from the uninsured to the insured, as indicated by the blue arrow in Figure 2, Panel B. From a market surplus perspective, this transfer has no welfare impact. But, from an ex-ante expected utility perspective, these transfers have value to the extent to which the marginal utilities of income differ for the insured and uninsured. If the

³Appendix A provides this calculation.

marginal utility of income is higher (lower) for the insured than uninsured, then lowering (raising) the price of insurance increases welfare. Accounting for these difference in marginal utilities of income between the insured and uninsured is the key to constructing ex-ante measures of welfare.

Prior to learning one's willingness to pay, there is a chance s of being insured. The impact of lower insurance prices on ex-ante expected utility is given by

$$s \frac{dp_I}{ds} E[u_c | Insured] ds = s(1-s) D'(s) E[u_c | Insured] ds$$

where $E[u_c | Insured]$ is the average marginal utility of income for the fraction s of the market that is insured. Conversely, the cost of having a higher price on ex-ante expected utility is given by

$$(1-s) \frac{dp_U}{ds} E[u_c | Uninsured] ds = -s(1-s) D'(s) E[u_c | Insured] ds$$

where $E[u_c | Uninsured]$ is the average marginal utility of income for the fraction $1-s$ of the market that is uninsured (for notational simplicity, I suppress the dependence of these marginal utilities on s , p_I , and p_U). Summing these two effects yields the ex-ante value of expanding the size of the insurance market from s to $s + ds$:

$$EA(s) = \underbrace{s(1-s)(-D'(s))}_{\text{Transfer}} \underbrace{\frac{E[u_c | Insured] - E[u_c | Uninsured]}{E[u_c]}}_{\text{Difference in Marginal Utilities}} \quad (2)$$

where I normalize by the average marginal utility of income, $E[u_c]$, to generate a willingness to pay out of consumption averaged over all states of the world. The first term, $s(1-s)(-D'(s))$, can loosely be interpreted as the size of the blue arrow in Figure 2, Panel B. Steeper slopes of demand imply greater price changes (and thus larger transfers) one moves from s to $s + ds$ of the market being insured. The second term, $\frac{E[u_c | Insured] - E[u_c | Uninsured]}{E[u_c]}$, is the percentage difference in marginal utilities between the insured and uninsured population. Weighting by the difference in marginal utilities recovers the ex-ante value

of insurance.⁴ To the extent to which those choosing to buy insurance have a higher marginal utility of income, the transfer from the uninsured to the insured increases ex-ante expected utility.⁵

Given $EA(s)$ in equation (2), I define the ex-ante demand curve, $D^{Ex-Ante}(s)$, as the sum of the willingness to pay revealed by choices, $D(s)$, and the additional ex-ante value of insurance, $EA(s)$:

$$D^{Ex-Ante}(s) = D(s) + EA(s) \quad (3)$$

Prior to learning their willingness to pay for insurance, individuals are willing to pay $D^{Ex-Ante}(s)$ to have prices set such that a fraction s of the market is insured. In particular, the value of having everyone insured, $D^{Ex-Ante}(1)$, is equal to $D^{Ex-ante}$ in equation (1).⁶ Equations (2) and (3) are an illustration of the first main result of the paper, formalized in Section 3. The key additional component to measure ex-ante willingness to pay is the percentage difference in marginal utilities between the insured and uninsured.

A key barrier to estimating $D^{Ex-Ante}(s)$ is that one does not readily observe the differences in marginal utilities between the insured and uninsured. The second main result of the paper borrows tools from the literature on optimal unemployment insurance (e.g. Baily (1978); Chetty (2006)) to approximate the difference in marginal utilities for the insured versus uninsured using Taylor expansions of the marginal utility function. Under conditions outlined below, this difference in marginal utilities between insured and uninsured is a function of (i) the willingness to pay curve, $D(s)$, and (ii) an estimate of risk aversion.

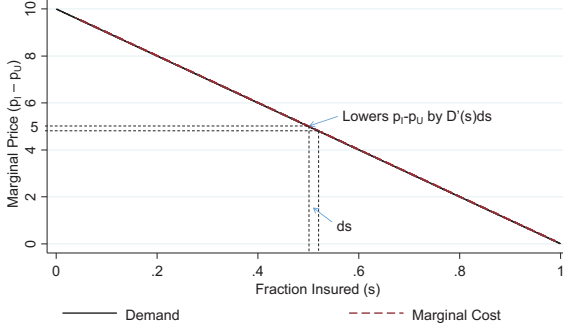
⁴This result is akin to the Baily-Chetty condition in optimal unemployment insurance that measures the value of more generous social insurance using the marginal utility of the beneficiaries (e.g. unemployed) relative to non-beneficiaries (e.g. employed) (Baily (1978); Chetty (2006)). Here, the beneficiaries of lower insurance prices are those who choose to purchase insurance.

⁵It is also possible that those who are uninsured have a higher marginal utility of income than the insured. This could be the case if the reason for not obtaining coverage is liquidity constraints, so that those choosing to forego insurance have a higher return to other forms of spending. This is ruled out in the simple example presented in this introduction, but will be possible in the more general model in the next Section.

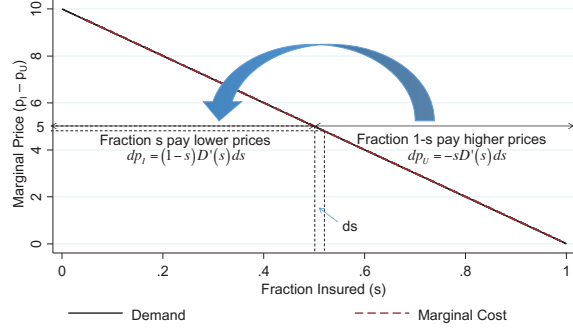
⁶More precisely, this is true up to an approximation error resulting from the fact that the average marginal utility, $E[u_c]$, varies with market size s ,

Figure 2: Recovering Ex-Ante Willingness to Pay

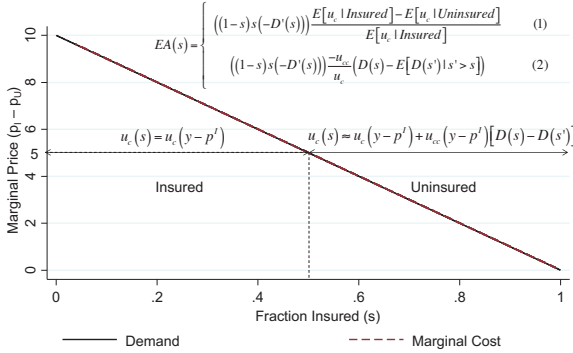
A. Marginal Increase in Fraction Insured



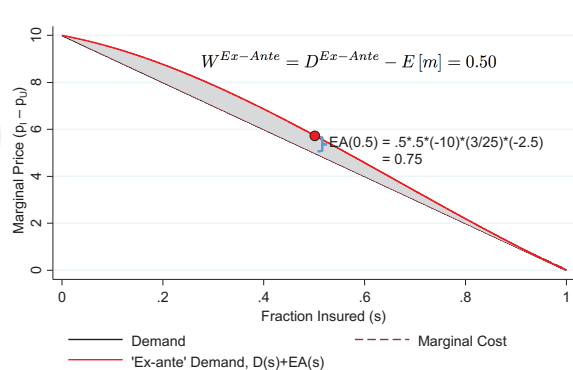
B. Transfer from Uninsured to Insured



C. Valuation of Transfer using Marginal Utilities



D. Recovering Ex-Ante Willingness to Pay



To illustrate how this is possible, return to the example above. The insured have consumption of $30 - p_I$. So, their marginal utility is given by $u_c(30 - p_I)$, where u_c is the marginal utility function (e.g. $u_c(c) = c^{-\sigma}$ if $u(c)$ is constant relative risk aversion). The consumption of the uninsured facing known loss $m(s)$ is given by $30 - p_U - m(s)$, so that their marginal utility is $u_c(30 - p_U - m(s))$. Averaging across the uninsured with different loss sizes and using the identity $D(s) = m(s)$, the average marginal utility of the uninsured is given by $E[u_c(30 - p_U - D(s')) | s' \geq s]$.

Now, consider a first order Taylor expansion to the marginal utility function

of the insured around a consumption level c^* . This yields

$$\begin{aligned} u_c(30 - p_U - D(s')) &\approx u_c(c^*) + u_{cc}(c^*) [(30 - p_U - D(s')) - c^*] \\ &\approx u_c(c^*) + u_{cc}(c^*) [p_I - p_U - D(s')] \end{aligned}$$

Similarly, the marginal utility of the insured is given by

$$u_c(30 - p_I) \approx u_c(c^*) + u_{cc}(c^*) [30 - p_I - c^*]$$

So, the difference between insured and uninsured is given by

$$\begin{aligned} E[u_c|Insured] - E[u_c|Uninsured] &\approx u_{cc}(c^*) [(30 - p_I - c^*) - (30 - p_U - D(s') - c^*)] \\ &\approx u_{cc}(c^*) [D(s') - D(s)] \end{aligned}$$

where $p_I - p_U = D(s)$ is the equilibrium price of insurance when a fraction s purchases insurance. Now, take expectations over the uninsured types, s' , and normalize by $E[u_c] \approx u_c(c^*)$, where c^* is the average consumption in the population. This yields an expression for the percentage difference between the marginal utility of insured and uninsured:

$$\frac{E[u_c|Insured] - E[u_c|Uninsured]}{E[u_c]} \approx \frac{-u_{cc}}{u_c} (D(s) - E[D(s') | s < s']) \quad (4)$$

where $\frac{-u_{cc}}{u_c}$ is the coefficient of absolute risk aversion (evaluated at c^*) and $D(s) - E[D(s') | s < s']$ is the difference between the willingness to pay of the average uninsured person and the price, $D(s) = p_I(s) - p_U(s)$, when a fraction s of the market is insured.

Equation (4) provides a method to estimate the ex-ante measures of welfare using the market demand curve and a measure of risk aversion. Risk aversion can either be imported from another setting, or one can infer it by comparing the markup individuals are willing to pay for insurance to the variance reduction offered by the insurance product, as discussed in Section 4 and shown in Appendix B.⁷

⁷For example, in a CARA-Normal model the coefficient of absolute risk aversion is equal to twice the ratio of the markup individuals are willing to pay for insurance relative to the variance reduction in out of pocket expenses it provides. Appendix B provides a more gen-

In the stylized example, the coefficient of relative risk aversion is 3 and the average consumption in the population is 25. So, the coefficient of absolute risk aversion is approximately $3/25$. Using equation (2), the ex-ante value of insurance from expanding the market when exactly 50% have insurance is $EA(0.5) = 0.5 * 0.5 * (-10) * (3/25) * (5 - 2.5) = 0.75$. From behind the veil of ignorance, individuals are willing to pay \$0.75 to expand the size of the insurance market from 50% to 51% insured relative to what would be indicated by their demand curve (which equals $D(0.5) = 5$). This is illustrated in Figure 2, Panel D.

Panel D of Figure 2 uses equations (2) and (4) to calculate $EA(s)$ for all values of $s \in [0, 1]$. Adding this ex-ante value to the market demand curve yields the ex-ante demand curve, $D^{Ex-Ante}(s) = D(s) + EA(s)$, depicted by the solid red line. At each value of s , $D^{Ex-Ante}(s)$ measures the impact on ex-ante expected utility of expanding the size of the insurance market from s to $s+ds$. Integrating from $s = 0$ to $s = 1$ yields the value of insuring everyone,

$$\int_0^1 D^{Ex-Ante}(s) = 5.50 = D^{Ex-Ante}$$

Integrating under the ex-ante demand curve in Figure 2, Panel D, yields \$5.50. Not coincidentally, this equals the integral under the demand curve in Figure 1, Panel A. In this sense, the approach ex-ante demand curve recovers the willingness to pay individuals would have for everyone to be insured ($s = 1$) if they were asked this willingness to pay prior to learning m . Moreover, the ex-ante demand curve can be used to evaluate the impact of insurance taxes and subsidies that expand the size of the market from, e.g., 50% to 51% on ex-ante expected utility.

Equation (4) illustrates the second main result of the paper outlined in Section 4: under certain conditions, one can recover the ex-ante willingness to pay for insurance using the observed market demand and cost curves combined with a measure of risk aversion. This provides a benchmark method to

eral characterization for more general utility functions and risk distributions. In this simple example here, there is no remaining risk that drives insurance demand. As a result, willingness to pay does not reveal anything about risk aversion; but in more realistic empirical applications one can potentially estimate this risk aversion coefficient internally.

empirically implement the ex-ante welfare framework.

The model in this section is highly stylized. There is no moral hazard, no preference heterogeneity, and the model assumed all information about costs, m , was revealed at the time of making the insurance decision. The next section extends these derivations to capture more realistic features of insurance markets encountered in common empirical applications, such as the one considered in Einav et al. (2010) or Finkelstein et al. (2017).

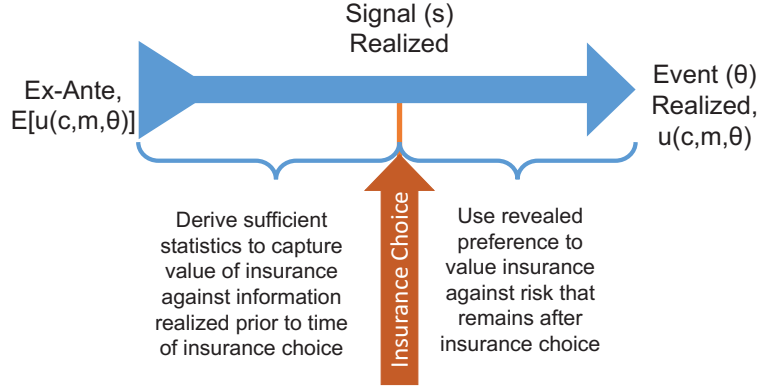
3 General Model

This section provides a general method for recovering the ex-ante willingness to pay for insurance. The language of the model will refer to a health insurance context. But, it is straightforward to amend the model to capture other insurance settings, such as unemployment insurance. Proposition 1 will provide a generalization of equation (2) for measuring the impact of mandates and subsidies on ex-ante expected utility. As in the stylized example, the key additional piece of information that is required to measure this ex-ante value is the difference in marginal utilities of income between the insured and uninsured. Section 4 will provide conditions under which one can approximate this difference using the demand and cost curves combined with a measure of risk aversion, as in equation (4) in the stylized example.

3.1 Setup

Individuals face uncertainty over a future event, captured by a random variable θ . After learning θ , individuals choose their non-medical consumption, c , and medical expenditures, m . Therefore, one should think of θ as capturing all information that goes into particular treatment and medical expenditure decisions. Individuals have a utility function over these choices, $u(c, m; \theta)$. For generality, I allow θ to affect both preferences (e.g. the value of medical services) and the budget constraint through effects on income, $y(\theta)$.

Figure 3: Timeline of Information Revelation and Insurance Purchase



The budget constraint depends on whether individuals have purchased an insurance policy to help cover some of their medical expenditures, m . I assume there exists a single insurance contract at price p_I that allows individuals to pay $x(m; \theta)$ for medical services m . To nest settings beyond standard health insurance products, I allow this cost, $x(m; \theta)$ to vary with θ . This captures indemnity insurance payments made independent of the individual's choice of m . This yields a budget constraint for the insured:

$$c^I(\theta) + x(m^I(\theta); \theta) + p_I \leq y(\theta) \quad (5)$$

Conversely, uninsured individuals pay the full price of m . This yields a budget constraint

$$c^U(\theta) + m^U(\theta) + p_U \leq y(\theta) \quad (6)$$

where p_U is a penalty or tax paid by individuals that are uninsured. For simplicity, I consider only a binary insurance choice, leaving future work to extend the approach to multiple insurance contracts. Let $\{c^I(\theta), m^I(\theta)\}$ denote the choice of consumption and medical spending of an insured type θ , and let $\{c^U(\theta), m^U(\theta)\}$ denote the choices of an uninsured type θ .⁸

⁸I adopt the common assumption (e.g. Einav et al. (2010)) that $m^I(\theta)$ does not depend on p_I . In principle, the choice of $m^I(\theta)$ could depend on p_I ; for example, if insurance is

Figure 3 presents a timeline of information revelation and outlines the empirical approach. At the time individuals make the decision to be insured or uninsured, I allow individuals to potentially know something about their particular type θ , which I denote by a signal $\tilde{s} \in [0, 1]$. After learning s , individuals choose to be insured and face the budget constraint in (5) or uninsured and face the budget constraint in (6).

Given \tilde{s} , let $D(\tilde{s})$ denote the marginal price that a type \tilde{s} is willing to pay for insurance. This solves

$$E[u(y(\theta) - x(m^I(\theta); \theta) - D(\tilde{s}) - p_U, m^I(\theta); \theta) | \tilde{s}] = E[u(y(\theta) - m^U(\theta) - p_U, m^U(\theta); \theta) | \tilde{s}] \quad (7)$$

All \tilde{s} such that $p_I - p_U \leq D(\tilde{s})$ will choose to purchase insurance, whereas types \tilde{s} for which $D(\tilde{s}) > p_I - p_U$ will choose to remain uninsured and pay p_U . For simplicity, I follow Einav et al. (2010) and assume that only the relative price of insurance, $p_I - p_U$ affects demand.⁹ Without loss of generality, assume that \tilde{s} is ordered so that demand, $D(\tilde{s})$, is decreasing in \tilde{s} . This means that if insurance prices are p_U and p_I , a fraction s will purchase insurance where s solves $D(s) = p_I - p_U$.

I will use the information revealed through insurance choices to calculate the value of insurance against θ given s , and will derive sufficient statistics that augment them to measure the ex-ante expected utility impact of subsidies and mandates. Following Einav et al. (2010), define the average cost imposed on the insurer when a fraction s of the market owns insurance by

$$AC(s) = E[m^I(\theta) - x(m^I(\theta); \theta) | \tilde{s} \leq s] \quad (8)$$

so that $sAC(s)$ is the total cost of insuring a fraction s of the market. Define $C(s)$ to characterize how the total cost to the insurer changes as the size of the market expands, $C(s) = \frac{d}{ds} [sAC(s)]$. This cost is the net difference between

cheaper, individuals may make riskier choices that increase health costs later on. In this case, one would need to account for the impact of price changes on the costs of the insured pool.

Similarly, I make the simplifying assumption that $m^U(\theta)$ does not depend on p_U . However, in contrast to the assumption that $m^I(\theta)$ does not depend on p_I , this assumption is without loss of generality because of the envelope theorem: $m^U(\theta)$ is fully paid by the individual so that behavioral responses of m^U do not affect welfare measures.

⁹Appendix C provides a generalized Proposition 1 to the case when demand is affected differentially by increases of p_U as opposed to decreases in p_I .

expenditures and out-of-pocket spending for those with signal¹⁰ s :

$$C(s) = E [m^I(\theta) - x(m^I(\theta); \theta) | \tilde{s} = s] \quad (9)$$

Finally, let $p_I(s)$ and $p_U(s)$ denote the prices of insurance and remaining uninsured when a fraction s of the market owns insurance. By definition, these prices must be consistent with the definition of willingness to pay,

$$D(s) = p_I(s) - p_U(s) \quad (10)$$

Lastly, let $G(s)$ denote the total cost (net of premiums collected) to the insurer of insuring a fraction s of the market by setting prices $p_I(s)$ and $p_U(s)$:

$$G(s) = \underbrace{sAC(s)}_{\text{Cost of Insured}} - \underbrace{[sp_I(s) + (1-s)p_U(s)]}_{\text{Premiums Collected}} \quad (11)$$

In the case in which insurers earn zero profits, or in which the government breaks even, one can set $G(s) = 0$ so that prices $p_I(s)$ and $p_U(s)$ are then defined implicitly as solutions to equations (11) and (10). More generally, $G(s)$ captures the net resource expenditures (e.g. government subsidies) for this health insurance market. Below, I illustrate how to conduct welfare analysis for both budget neutral ($G(s) = 0$) and non-budget neutral settings in which the government subsidizes the market ($G(s) \neq 0$).

3.2 Ex-Ante Welfare

The goal is to use an ex-ante expected utility perspective to evaluate the desirability of insurance market policies (e.g. subsidies and mandates) that change the fraction who obtain insurance. Let $W(s)$ denote the ex-ante expected utility when prices, $p_I(s)$ and $p_U(s)$, are such that a fraction s of the market

¹⁰This relies on the assumption noted above that individuals' choices are not affected by prices p_U and p_I . If prices do affect the cost to the insurer, this marginal cost function contains an additional term reflecting the net cost of those behavioral responses on the insurance company.

owns insurance. This is given by

$$\begin{aligned}
 W(s) = & \int_0^s E[u(y(\theta) - p_I(s), m^I(\theta)) | \tilde{s}] d\tilde{s} \\
 & + \int_s^1 E[u(y(\theta) - m^U(\theta) - p_U(s), m^U(\theta); \theta) | \tilde{s}] d\tilde{s}
 \end{aligned} \tag{12}$$

The first term integrates over those who choose to be insured, $\tilde{s} \leq s$. The second term integrates over those who choose to be uninsured, $\tilde{s} > s$. If one observed or estimated the utility function and its arguments, one could directly measure $W(s)$. This would be analogous to the approach to measuring welfare taken by Handel et al. (2015) and Finkelstein et al. (2016). Here, I instead follow the “sufficient statistics” approach of Einav et al. (2010) and build a measure of $W(s)$ from the willingness to pay and cost curves.

Before proceeding further, it is important to be clear about the definition of ex-ante expected utility in equation (12). This expectation integrates over the entire distribution of s . This means that the conceptual experiment involves holding fixed the definition of the “market” but measuring utility prior to when individuals learn their particular willingness to pay for insurance in this market. For example, if one estimated $D(s)$ and $C(s)$ curves for those employed at a large firm, then $W(s)$ would recover the expected utility impact of firm policies that lead to a fraction s of the insurance-eligible population in the firm purchasing insurance. But, it does not measure the willingness to pay for insurance prior to when individuals learn they are employed at the firm. Similarly, if $D(s)$ and $C(s)$ are estimates from a low-income health insurance program for those at 150% of the federal poverty line (FPL) as in Finkelstein et al. (2017), equation (12) will measure the expected utility of those at 150% FPL. It will not capture any insurance value against the risk of earning only 150% FPL. Extending the analysis to consider the ex-ante value of insurance against being in the eligible market at all amounts to asking whether the government should increase subsidies to the market, and will be addressed by considering the marginal value of public funds of additional expenditures in Subsection 3.2.2 below.

Given this apparatus, the ex-ante welfare impact of expanding the insurance market by a small amount starting with a fraction s insured is $W'(s)$.

Dividing $W'(s)$ by $E[u_c]$ to form $\frac{W'(s)}{E[u_c]}$ yields the ex-ante willingness to pay out of their own income to expand the insurance market.¹¹

To characterize $\frac{W'(s)}{E[u_c]}$, I proceed as follows. I use the willingness to pay function, $D(s)$, to capture the impact on the utility of the uninsured.¹² This yields an expression for $W(s)$ that does not require keeping track of the uninsured utility:

$$W(s) = \int_0^s E[u(y(\theta) - p_I(s), m^I(\theta)) | \tilde{s}] d\tilde{s} + \int_s^1 E[u(y(\theta) - D(\tilde{s}) - p_U, m^I(\theta); \theta) | \tilde{s}] d\tilde{s}$$

The marginal welfare impact of expanding the size of the insurance market is given by taking the derivative with respect to s ,

$$W'(s) = -sp'_I(s) E[u_c(y(\theta) - p_I(s), m^I(\theta)) | \tilde{s} \leq s] - (1-s)p'_U(s) E[u_c(y(\theta) - D(\tilde{s}) - p_U, m^I(\theta); \theta) | \tilde{s} \geq s] \quad (13)$$

The first term captures the welfare increase from lower prices for the insured ($p'_I < 0$). From behind the veil of ignorance, this price reduction of p'_I occurs with chance s and is valued using the marginal utility of income of the insured, $E[u_c(y(\theta) - p_I(s), m^I(\theta)) | \tilde{s} \leq s]$. The second term captures the welfare cost of having higher prices faced by the uninsured ($p'_U > 0$). This price

¹¹To see this, let $\tilde{W}(s, \delta)$ denote the ex-ante expected utility if fraction s are insured and have income $y(\theta) - \delta$, so that they pay δ out of their ex-ante income for insurance. Let $\Delta(s, s')$ denote the willingness to pay to move from a world with a fraction s insured to a world with a fraction s' insured. This is given by the solution to

$$\tilde{W}(s, \Delta(s, s')) = \tilde{W}(s', 0) = W(s')$$

where $W(s')$ is given by equation (12). Differentiating $\Delta(s, s')$ with respect to s' and evaluating at $s' = s$ yields

$$\frac{d}{ds} \Big|_{s'=s} \Delta(s, s') = \frac{W'(s)}{-\frac{\partial \tilde{W}}{\partial \delta}} = \frac{W'(s)}{E[u_c]}$$

where the second equality follows from the fact that the ex-ante utility impact of additional δ is the average marginal utility of income, $-\frac{\partial \tilde{W}}{\partial \delta} = E[u_c]$.

¹²Recall equation (7) implies $E[u(y(\theta) - m^U(\theta) - p_U, m^U(\theta); \theta) | \tilde{s}] = E[u(y(\theta) - D(\tilde{s}) - p_U, m^I(\theta); \theta) | \tilde{s}]$.

increase occurs with a chance $1 - s$ and is valued using the average marginal utility of income, $E[u_c(y(\theta) - D(\tilde{s}) - p_U(s), m^I(\theta); \theta) | \tilde{s} \geq s]$.

The value of $W'(s)$ depends on how prices are affected by the expansion of the insurance market, $p'_I(s)$ and $p'_U(s)$. This in turn depends on whether the policy is budget neutral.

3.2.1 Budget Neutral Policies

I first consider the case when the premiums collected cover the cost of the insured, as in Einav et al. (2010). Combining equation (13) with the resource constraint in equation (11) when $G'(s) = 0$ yields the following result.

Proposition 1. *For budget neutral policies satisfying $G'(s) = 0$, the marginal welfare impact of expanding the size of the insurance market from s^* to $s^* + ds$ is given by*

$$\frac{W'(s^*)}{E[u_c]} \approx \underbrace{D(s^*) + EA(s^*)}_{DEx-Ante(s)} - C(s^*) \quad (14)$$

where $EA(s^*)$ is the additional ex-ante value of expanding the size of the insurance market,

$$EA(s^*) = \underbrace{(1 - s^*)(C(s^*) - D(s^*) - s^*D'(s^*))}_{\text{Transfer from Uninsured to Insured}} \beta(s^*) \quad (15)$$

and $\beta(s)$ is the percentage difference in marginal utilities of income for the insured relative to the uninsured,

$$\beta(s) = \frac{E[u_c(y(\theta) - p_I(s), m^I(\theta); \theta) | \tilde{s} \leq s] - E[u_c(y(\theta) - D(\tilde{s}) - p_U, m^I(\theta); \theta) | \tilde{s} \geq s]}{E[u_c]} \quad (16)$$

Proof. See Appendix D. □

Equation (14) shows that the marginal ex-ante willingness to pay for a larger insurance market is given by the sum of $D(s) + EA(s) - C(s)$. The term $D(s) - C(s)$ is market surplus: expanding the size of the insurance market increases ex-ante welfare to the extent to which individuals are willing to pay more than their costs for insurance. $EA(s)$ captures the additional ex-ante value of expanding the size of the market through its impact on insurance

prices. Expanding the insurance market induces a transfer from uninsured to insured of size $(1 - s^*) (C(s^*) - D(s^*) - s^* D'(s^*))$. This term reduces to the transfer in equation (2) when demand equals marginal cost, $D(s) = C(s)$, as in the stylized example. Moving financial resources from the uninsured to the insured increases ex-ante welfare to the extent to which the marginal utility of income is higher for the insured than the uninsured. This difference is captured by the term $\beta(s^*)$.

The optimal size of the insurance market Ex-ante expected utility is maximized at the value of $s = s_{ea}$ such that $W'(s_{ea}) = 0$. Equation (15) shows that this occurs when $D(s_{ea}) + EA(s_{ea}) - C(s_{ea}) = 0$. This contrasts with the size of the market that maximizes market surplus, $D(s_{ms}) - C(s_{ms}) = 0$. Instead of setting market demand equal to costs, the size of the market that maximizes ex-ante expected utility includes the ex-ante value from having lower insurance prices. At the optimum, marginal lost surplus, or deadweight loss, is equated to this ex-ante value of insurance,

$$EA(s_{ea}) = C(s_{ea}) - D(s_{ea})$$

So, as long as $EA(s) > 0$, the size of the market that maximizes expected utility is larger than the size of the market that maximizes market surplus. Optimally set insurance subsidies involve deadweight loss to the extent to which it provides ex-ante risk protection.

The sign of $\beta(s^*)$ The ex-ante term, $EA(s)$, is positive whenever the marginal utility of income is higher for the insured than uninsured, $\beta(s) > 0$. In canonical models of insurance, one would expect $\beta(s^*) > 0$. For example, in the stylized example in Section 2, those who choose to purchase insurance expect to face a higher financial loss than those who remain uninsured. This means that the consumption levels of the insured are lower than those of the uninsured. Concavity of the utility function then implies that the marginal utilities of the insured are higher than the uninsured, so that $\beta(s^*) > 0$.

But, it is also possible to have $\beta(s^*) < 0$. For example, θ could reflect a liquidity or income shock to $y(\theta)$ so that the primary driver of the decision to

purchase insurance is not a higher expected cost, but rather a liquidity shock that makes the value of medical care less than the value of additional other consumption. If the uninsured are foregoing insurance purchase because of this liquidity shock, then it is feasible that those who forego insurance have a higher marginal utility of income than those who purchased, $\beta(s^*) < 0$. In this case, expanding the size of the insurance market will transfer resources from the liquidity constrained to those who are less constrained, which would suggest that $EA(s^*) < 0$. Going forward, most of the discussion will consider the benchmark case where $\beta(s^*) > 0$. But, this highlights the generality of the sufficient statistic approach for capturing many potential underlying models. And, it suggests a value of future work estimating $\beta(s)$ in a wide class of settings.

3.2.2 The MVPF for Non-Budget Neutral Policies

Proposition 1 applied to budget-neutral policy changes where the cost of the insured was fully covered by premiums and mandate penalties. One can also consider the case in which the government subsidizes the price of insurance so that funding for the insurance market comes from those outside of the insurance market (e.g. via taxation of others in the economy). To assess this case, I consider the marginal value of public funds (MVPF) of lower insurance prices. Following Hendren (2016), the MVPF is defined as the individual's willingness to pay for a larger insurance market divided by the cost to the government of using subsidies to expand the insurance market,

$$MVPF(s) = \frac{\frac{W'(s)}{E[u_c]}}{G'(s)} = \frac{\text{Marginal WTP}}{\text{Marginal Cost}}$$

Hendren (2016, 2014) shows how one can compare this MVPF to the MVPF of other policies affecting similar populations to study the optimality of government policies: if the MVPF of lowering health insurance prices is higher than the MPVF of a tax cut to a similar population (e.g. EITC in the case of low-income health insurance subsidies), then welfare can be increased by lowering health insurance prices financed by a reduction in tax subsidies.

Proposition 2 provides a characterization of the MVPF for non-budget

neutral policies when uninsured individuals pay no penalty, $p_U(s) = 0$.

Proposition 2. *Suppose $p_U(s) = 0$. The MVPF of additional insurance market subsidies is given by*

$$MVPF(s) = \frac{1 + (1 - s)\beta(s)}{1 + \frac{C(s) - D(s)}{s(-D'(s))}} \quad (17)$$

where $\beta(s)$ is the percentage difference in marginal utilities of income for the insured relative to the uninsured given by equation (16).

Proof. See Appendix E □

The denominator in equation (17), $1 + \frac{C(s) - D(s)}{s(-D'(s))}$ is the marginal cost of lowering insurance prices. This has two components. First, lowering premiums by \$1 increases the cost by \$1 for each of the s enrollees. Second, there is an additional cost from those induced to purchase insurance by the lower prices. These enrollees pay $D(s) = p_I(s)$ but cost the insurer $C(s)$. So, they impose a net cost of $C(s) - D(s)$. A \$1 price reduction increases the size of the market by $\frac{1}{-D'(s)}$. Hence, the total cost normalized by the size of the market s , of lowering premiums by \$1 is $1 + \frac{C(s) - D(s)}{s(-D'(s))}$.¹³

The numerator in equation (17) reflects the willingness to pay for lower insurance premiums. An individual who has already learned their signal s and decided to purchase insurance is willing to pay \$1 to have premiums that are \$1 lower. So, if welfare were not being calculated from behind the veil of ignorance, the numerator would simply be 1 and the welfare impact would be $\frac{1}{1 + \frac{C(s) - D(s)}{s(-D'(s))}}$. This corresponds to the MVPF reported in Finkelstein et al. (2017). But, from behind the veil of ignorance, individuals are willing to pay an additional $(1 - s)\beta(s)$ to have premiums that are \$1 lower.

Propositions 1 and 2 characterize how to measure the ex-ante expected utility impact of insurance market policies. In addition to the market demand and cost curves, the key additional component required is an estimate of the

¹³More generally, if there are additional behavioral responses that affect the government budget (e.g. if insurance improves health and increases taxable income, or if the subsidies distort labor supply, etc.), these would also need to be incorporated into the marginal cost of lowering premiums.

difference in marginal utilities between insured and uninsured, $\beta(s)$. The next section discusses a benchmark path to providing an estimate of $\beta(s)$.

4 Implementation Using Market Demand and Cost Curves

This section provides conditions under which one can write $\beta(s)$ as a function of market level demand curves combined with a measure of risk aversion analogous to equation 4 in the stylized example of Section 2. To be specific, Proposition 3 will provide conditions under which

$$\beta(s^*) = \gamma [D(s^*) - E_s [D(s) | s \geq s^*]] \quad (18)$$

where γ is the coefficient of absolute risk aversion and $D(s^*) - E_s [D(s) | s \geq s^*]$ is the difference in willingness to pay between the marginal type, s^* , and the average uninsured type, $s \geq s^*$. Loosely, $D(s^*) - E_s [D(s) | s \geq s^*]$ captures the amount of information that is revealed about demand at the time of choosing insurance; γ then translates this into a difference in marginal utilities. The estimate of risk aversion can either be imported from external settings, or it can be estimated internally using the relationship between the markup individuals are willing to pay and the reduction in consumption variance provided by the insurance, as discussed in Appendix B.

The assumptions required to generate equation (18) are strict but common in the literature on optimal unemployment insurance (Baily (1978); Chetty (2006)). Moreover, many of these assumptions are satisfied in the structural models used to estimate the WTP for insurance. Thus, it provides a benchmark method for researchers having estimated demand and cost curves to infer whether an ex-ante welfare perspective can lead to welfare conclusions that differ from a focus on market surplus. Section 4.2 shows how relaxing these assumptions is possible if one observes additional data elements.

4.1 Implementation Assumptions

The implementation relies on a Taylor expansion of the marginal utility utility function. Let $\bar{y} = E[y(\theta)]$ denote the average income of the population. Let $\bar{c} = \bar{y} - p_I$ denote average consumption. To help illustrate the role of preference heterogeneity, assume θ is a uni-dimensional index, $\theta \in \mathbb{R}$, and assume that the utility function, $u(c, m; \theta)$, is continuously differentiable with respect to θ . Let $\bar{\theta} = E[\theta]$ denote the average θ in the population. To a first order Taylor approximation, the average marginal utility in the population, $E[u_c(y(\theta) - p_I(s), m^I(\theta); \theta)]$, is given by the marginal utility of the average type, $\bar{u}_c = u_c(\bar{c}, \bar{m}, \bar{\theta})$.¹⁴

The marginal utility of other uninsured types, θ , with demand $D(s)$ is approximately¹⁵

$$\begin{aligned} & u_c(y(\theta) - D(s) - p_U, m^I(\theta); \theta) - u_c(\bar{c}, \bar{m}, \bar{\theta}) \\ & \approx \underbrace{u_{cc}(y(\theta) - D(s) - p_U - (\bar{y} - p_I))}_{\text{Consumption}} + \underbrace{u_{cm}(m^I(\theta) - \bar{m})}_{\text{Medical Spending}} + \underbrace{u_{c\theta}(\theta - \bar{\theta})}_{\text{Preferences}} \end{aligned} \quad (19)$$

Using equation (19), one can aggregate the insured and uninsured marginal

¹⁴To see this, note that one can write $u_c(y(\theta) - p_I(s), m^I(\theta); \theta)$ as

$$u_c(y(\theta) - p_I(s), m^I(\theta); \theta) - u_c(\bar{c}, \bar{m}, \bar{\theta}) \approx u_{cc}(y(\theta) - p_I - \bar{c}) + u_{cm}(m^I(\theta) - \bar{m}) + u_{c\theta}(\theta - \bar{\theta})$$

where subscripts denote derivatives and u_{cc} , u_{cm} , and $u_{c\theta}$ are evaluated at $(\bar{c}, \bar{m}, \bar{\theta})$. Hence,

$$\begin{aligned} & E[u_c(y(\theta) - p_I(s), m^I(\theta); \theta)] - u_c(\bar{c}, \bar{m}, \bar{\theta}) \\ & \approx E[u_{cc}(y(\theta) - p_I - \bar{c}) + u_{cm}(m^I(\theta) - \bar{m}) + u_{c\theta}(\theta - \bar{\theta})] \\ & = u_{cc}(E[y(\theta)] - p_I - \bar{c}) + u_{cm}(E[m^I(\theta)] - \bar{m}) + u_{c\theta}(E[\theta] - \bar{\theta}) \end{aligned}$$

which equals zero by the definition of \bar{c} , \bar{m} , and $\bar{\theta}$.

¹⁵The Taylor expansion relies on there being only three utility arguments: consumption, medical expenditure, and preference heterogeneity. If there were additional arguments of the utility function, this would need to be incorporated into the calculation. One example of this is dynamics. If there were multiple arguments to consumption and individuals who are insured can spread their payment of insurance across multiple periods, then consumption will not be given by $y(\theta) - p_I$, but rather $y(\theta) - \frac{1}{T}p_I$ where T is the number of periods one can smooth the payments. It is straightforward to show that this implies $\beta(s^*) = \frac{1}{T}\gamma E[D(s^*) - D(s) | s \geq s^*]$. Intuitively, only $1/T$ of the willingness to pay actually comes from consumption in any given period.

utilities to write the difference as

$$\begin{aligned}
\beta(s^*) &\approx \frac{\bar{u}_{cc}}{\bar{u}_c} (E[(y(\theta) - p_I - (\bar{y} - p_I)) | s \leq s^*] - E[(y(\theta) - D(s) - p_U - (\bar{y} - p_I)) | s > s^*]) \\
&\quad + \frac{\bar{u}_{cm}}{\bar{u}_c} (E[m^I(\theta) | s \leq s^*] - E[m^I(\theta) | s > s^*]) \\
&\quad + \frac{\bar{u}_{c\theta}}{\bar{u}_c} (E[\theta | s \leq s^*] - E[\theta | s > s^*])
\end{aligned} \tag{20}$$

where \bar{u}_{cX} is the derivative of u_c with respect to $X \in \{c, m, \theta\}$, evaluated at the average type, $(\bar{c}, \bar{m}, \bar{\theta})$. The difference in marginal utilities between the insured and uninsured depends on the average differences in consumption, medical spending, and preferences for the insured versus uninsured.

Three assumptions facilitate an estimate of $\beta(s^*)$ as in equation (18). The first assumption is that the marginal utility of consumption does not depend on the level of medical spending.

Assumption 1. *(No Complementarities/Substitutabilities between c and m)*
The marginal utility function, $u_c(c, m; \theta)$ does not depend on m .

Assumption 1 is satisfied in the broad class of models that assume a single consumption argument in the utility function, such as in the example in Section 2 (and the more general model of Handel et al. (2015)) where $c = y - m$ and utility is only an argument of consumption, c .

Next, I assume away preference heterogeneity in the marginal utility function over consumption, u_{cc} .

Assumption 2. *The marginal utility function, $u_c(c, m; \theta)$, does not depend on θ*

This assumption does not prevent preference heterogeneity in general, but it implies that $u_{c\theta}(\theta - \bar{\theta}) = 0$ so that there is no covariance between types and the marginal utility of consumption.

Under Assumptions 1-2, the difference in marginal utilities depends only on the difference in consumption between insured and uninsured, $y(\theta) - D(\bar{s}) - p_U - (\bar{y} - p_I)$, and the curvature of the utility function, u_{cc} . The third assumption rules out differences in marginal utilities driven by systematic differences in $y(\theta)$.

Assumption 3. (*No Liquidity / Income Differences*) Income does not systematically vary between insured and uninsured, $\bar{y} = E[y(\theta) | s^* \leq s] = E[y(\theta) | s^* > s]$.

Assumption 3 rules out liquidity effects as a primary source of variation in demand for insurance. As discussed in Section 4.2, one can incorporate liquidity effects if one is able to observe the average income levels of the insured and uninsured. Assumption 3 implies that the difference in demand between the marginal insured type, $D(s^*) = p_I - p_U$, and the average uninsured type, $E[D(s) | s \geq s^*]$ drives differences in consumption between the insured and uninsured. Combining Assumptions 1-3 yields the result.

Proposition 3. *Suppose Assumptions 1-3 hold. Then,*

$$\beta(s^*) \approx \gamma(D(s^*) - E_s[D(s) | s \geq s^*]) \quad (21)$$

where the \approx denotes a first-order Taylor approximation and $\gamma = \frac{-\bar{u}_{cc}}{\bar{u}_c}$ is the coefficient of absolute risk aversion evaluated at the average level of consumption, \bar{c} , medical spending, \bar{m} , and health status, $\bar{\theta}$. The ex-ante component of willingness to pay is given by

$$EA(s^*) \approx (1 - s^*) (C(s^*) - D(s^*) - s^* D'(s^*)) \gamma (D(s^*) - E[D(s) | s \geq s^*]) \quad (22)$$

so that it is identified from the demand and cost curves, combined with a coefficient of absolute risk aversion, γ .

Proof. Imposing Assumptions 1-3 to equation (20) yields the result. \square

When $C(s) = D(s)$ as in the stylized example, equation (22) reduces to equation (4). In this more general setup, equation (21) provides a benchmark method estimate $\beta(s)$ and implement the ex-ante welfare approach.

4.2 Violations of Assumptions 1-3

Assumptions 1-3 provide a benchmark method to estimate $\beta(s)$. But, they are strong assumptions. Here, I discuss these limitations and illustrate how to relax them with suitable additional empirical estimates.

Assumption 1 Assumption 1 is violated if consumption of medical spending is a substitute (or complement) to consumption. In the more general case with $u_{cm} \neq 0$, $\beta(s^*)$ can be written as:

$$\beta(s^*) \approx \gamma(D(s^*) - E[D(s) | s \geq s^*]) + \frac{\bar{u}_{cm}}{\bar{u}_c} (E[m(\theta) | s \geq s^*] - E[m(\theta) | s < s^*])$$

where $\frac{\bar{u}_{cm}}{\bar{u}_c} = \frac{u_{cm}(\bar{c}, \bar{m}; \bar{\theta})}{u_c(\bar{c}, \bar{m}; \bar{\theta})}$ measures how the marginal utility of consumption varies with the level of medical spending (holding c and θ constant) and $E[m(\theta) | s \geq s^*] - E[m(\theta) | s < s^*]$ is the difference in spending between the uninsured and insured. This complementarity/substitutability of the utility function determines how individuals' budget allocation between c and m varies if one faces higher prices for m but is compensated with an equivalent increase in income. Thus, it could be estimated with exogenous variation in both income and prices of medical spending, m . With such an estimate of $\frac{\bar{u}_{cm}}{\bar{u}_c}$ and the difference in medical spending for insured and uninsured, one could relax Assumption 1.

Assumption 2 Assumption 2 would be violated if two types, θ , with the same level of consumption have different marginal utilities of consumption so that $u_{c\theta} \neq 0$. One potential reason for $u_{c\theta} \neq 0$ would be if the marginal utility of consumption depended on health status. If sicker people have lower marginal utilities of income (as in Finkelstein et al. (2013)), and the sick are more likely to purchase insurance, then those who purchase insurance may have lower marginal utilities of income than those who choose not to purchase insurance. In the more general case,

$$\beta(s^*) = \gamma(D(s^*) - E[D(s) | s \geq s^*]) + \frac{\bar{u}_{c\theta}}{\bar{u}_c} (E[\theta | s \geq s^*] - E[\theta | s < s^*])$$

The term $\frac{\bar{u}_{c\theta}}{\bar{u}_c} (E[\theta | s \geq s^*] - E[\theta | s < s^*])$ measures how much insured and uninsured would value a transfer from uninsured to insured even if they had the same level of consumption. Given measures of this systematic state-dependence of the utility function, one could relax Assumption 2.

Heterogeneous Risk Aversion. One might have also thought that heterogeneous risk aversion could generate a violation of Assumption 2. For example,

one would expect that the insured might have higher risk aversion than the uninsured. But, this does not necessarily lead to a violation of Assumption 2. This is because it is not necessary that those with greater curvature in the utility function (second derivatives of u) also have greater (or lower) marginal utilities (first derivatives of u). For example, if individuals have CRRA preferences, $u(c) = \frac{1}{1-\theta}c^{1-\theta}$ with heterogeneous θ , then those with higher θ would have a higher preference for insurance but would have a lower marginal utility of consumption, $c^{-\theta}$, so that $u_{c\theta} < 0$. But, if individuals have CRRA preferences, $u(c) = k^\theta c^{1-\theta}$, this utility function exhibits the same willingness to pay for insurance for a type θ but will have $u_c = k^\theta (1-\theta) c^{-\theta}$, which will be increasing in θ for sufficiently large k . In this sense, the marginal utility of consumption need not be correlated with the curvature of the utility function. As a result, heterogeneous risk aversion in and of itself does not provide a clear source of bias in the benchmark implementation in equation (22)

Assumption 3 Heterogeneous income or liquidity shocks, $y(\theta)$, could be a driver of insurance demand, as noted in the end of Section 3.2. If incomes differ between the insured and uninsured, then one can estimate a modified formula for $\beta(s^*)$ as

$$\beta(s^*) = \gamma(D(s^*) - E[D(s) | s \geq s^*] + E[y(\theta) | \tilde{s} \leq s] - E[y(\theta) | \tilde{s} > s])$$

where $E[y(\theta) | \tilde{s} \leq s] - E[y(\theta) | \tilde{s} > s]$ is the difference in incomes between the insured and uninsured. If the insured have higher incomes than the uninsured, then the benchmark formula for $\beta(s^*)$ in equation (21) will understate the ex-ante willingness to pay for insurance. However, if one can estimate this difference in average incomes, one can modify the ex-ante demand curve to account for this heterogeneity by simply adding this difference to $D(s^*) - E_s[D(s) | s \geq s^*]$.

5 Applications

I illustrate how to apply the approach to study the ex-ante welfare impact of health insurance subsidies for low-income adults in Massachusetts using

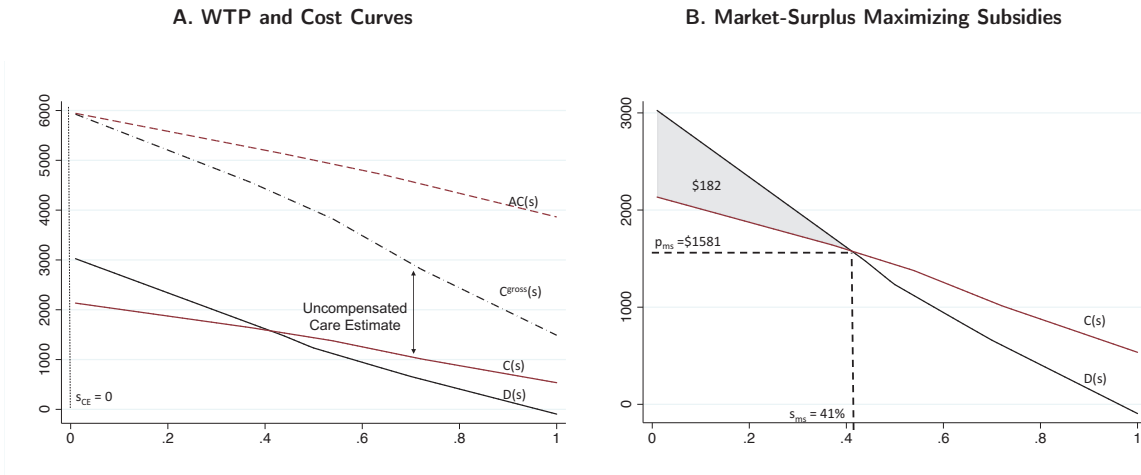
demand and cost estimates from Finkelstein et al. (2017). And, I explore the ex-ante welfare impact of different prices for more versus less generous insurance at a large employer using estimates from Einav et al. (2010).

5.1 Health Insurance Subsidies for Low-Income Adults

Using administrative data from Massachusetts' subsidized insurance exchange, Commonwealth Care, Finkelstein et al. (2017) exploit discontinuities in the subsidy schedule to estimate willingness to pay and costs of insurance among low-income adults. As subsidies decline and prices rise, insurance take-up falls. Figure 4 (Panel A) depicts the resulting willingness to pay and cost curves for those at with incomes at 150% of the Federal Poverty Line (FPL)¹⁶. Scaling the monthly premiums and costs in Finkelstein et al. (2017) by a factor of 12 to translate the monthly premiums and costs into annual figures.

Throughout the entire eligible population, willingness to pay falls below average costs of the insured, $D(s) < AC(s)$ for all s . In the absence of subsidies, a private market for low-income health insurance in MA would fully unravel ($s = 0$).

Figure 4: Willingness to Pay and Cost of Health Insurance for Low-Income Adults



¹⁶150% FPL corresponds to roughly \$16K in income for an individual with no children.

Figure 4 also presents the cost of these marginal enrollees that is paid by the insurer, $C^{gross}(s)$. For reasons described below, I refer to these claims paid directly by the insurer as the “gross” cost. The figure presents clear evidence of adverse selection: the marginal enrollees tend to be lower-cost (i.e. $C^{gross}(s)$ slopes downward). But, Finkelstein et al. (2017) also find that enrollee willingness to pay is far below individuals’ own expected costs paid by the insurer, $D(s) < C^{gross}(s)$. They suggest the rationale for this is the presence of uncompensated care. Because low-income uninsured can either obtain charity care from hospitals or default on medical debt, $C^{gross}(s)$ does not reflect the net cost of insurance. Some of the costs paid by the insurer compensate those who are otherwise providing uncompensated care.

To abstract from the complexities of these potential externalities on uncompensated care providers, I focus on the case in which the insurer/government is the payer of uncompensated care. This means that the cost to the insurer is the net resource cost that subtracts the cost of displaced uncompensated care from the insurer’s cost. I denote this net cost as $C(s)$.

Market Surplus Panel B of Figure 4 conducts a standard welfare analysis that focuses on market surplus, comparing observed willingness to pay to costs, $D(s) - C(s)$. Market surplus is maximized when 41% of the market owns insurance. The marginal price that leads to this allocation is \$1581.¹⁷ Relative to the competitive insurance market that fully unravels, $s = 0$, this allocation generates \$182 of market surplus, as indicated by the shaded region in Panel B.

Ex-ante Welfare How does this differ from a welfare perspective based on ex-ante expected utility? To move to ex-ante welfare, one requires an estimate of risk aversion. For the baseline case, I take a common estimate from the health insurance literature of $\gamma = 5 \times 10^{-4}$ (e.g. similar to estimates in Handel et al. (2015)).¹⁸

¹⁷This marginal price is less than the average cost of the 41% of the market who would purchase. The amount paid regardless of insurance purchase, p_U , would cover the remainder.

¹⁸Handel et al. (2015) estimates this for a relatively middle to high income population making choices over insurance plans. Under the natural assumption that absolute risk aversion decreases in consumption levels, this provides a lower bound on the size γ . But,

Figure 5: Ex-Ante Welfare of Health Insurance for Low-Income Adults

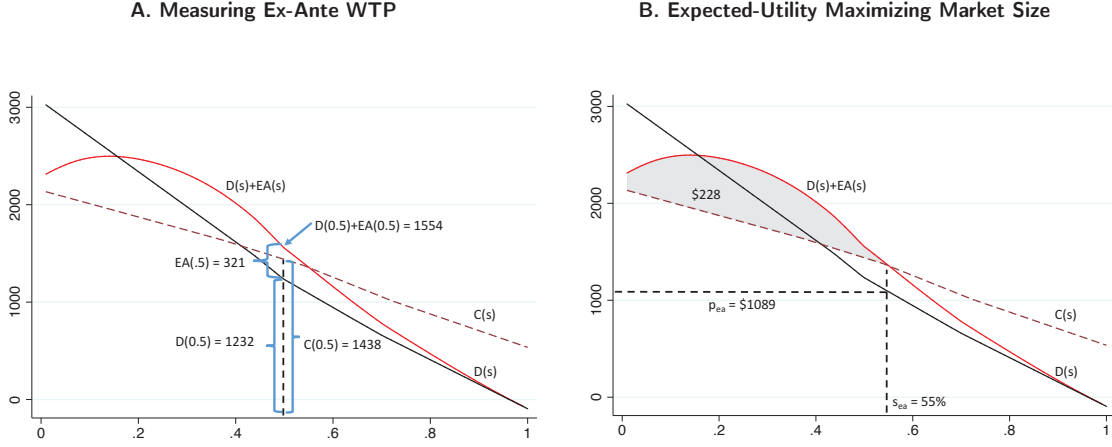


Figure 5 presents the ex-ante demand curve from Proposition 1 using $\gamma = 5 \times 10^{-4}$ and equation (22). Panel A illustrates the calculation of $EA(s)$ when 50% of the population owns insurance. The cost of the marginal enrollee is given by $C(0.5) = 1438$, willingness to pay is $D(0.5) = 1232$, and the slope of willingness to pay is $D'(0.5) = -3405$.¹⁹ The average willingness to pay with those whose demand is below $D(0.5) = 1232$ is 559. Combining using equation (22), the ex-ante willingness to pay for a larger insurance market is

$$\begin{aligned}
 EA(s) &= (1 - s^*) (C(s^*) - D(s^*) - s^* D'(s^*)) \gamma 2 (D(s^*) - E[D(s) | s \geq s^*]) \\
 &= .5 (1438 - 1232 + 0.5 * 3405) (5 \times 10^{-4}) (1232 - 559) \\
 &= 321
 \end{aligned}$$

this choice of risk aversion is that it implies a fairly high coefficient of relative risk aversion. For income levels of around 10,890 for those at 150% of FPL, it implies a coefficient of relative risk aversion of 8. Appendix F provides an alternative implementation that uses a coefficient of relative risk aversion of 3, which corresponds to an absolute risk aversion coefficient of 1.8×10^{-4} .

¹⁹Finkelstein et al. (2017) estimate a piece-wise linear demand curve. To obtain smooth estimates of the slope of demand, I regress the estimates of $D(s)$ from Finkelstein et al. (2017) on a 10th order polynomial in s . The results are similar for other smoothed functions.

While the median individual is willing to pay \$1,232 for insurance, prior to learning their demand individuals are willing to pay an additional \$321 for a larger insurance market. Dividing by 100, everyone would have been willing to pay \$3.21 from behind the veil of ignorance to have the opportunity to purchase insurance at the prices that lead to 51% of the market insured instead of 50% of the market insured.

Figure 5 presents the Ex-ante WTP curve for all values of s .²⁰ Expected utility is maximized when $W'(s) = 0$, or $D(s) + EA(s) = C(s)$. This occurs when 55% of the market owns insurance and the marginal price of insurance is \$1,089. This contrasts with the market surplus-maximizing size of the market of 41% and the optimal price is roughly 30% lower than surplus-maximizing price of \$1,581.

What is the welfare gain from pricing insurance optimally relative to the full unraveling of the competitive market ($s = 0$)? Everyone would be willing to contribute \$228 per person if they could live in a world in which insurance prices set at $p = \$1089$ so that the optimal 55% of the market obtains insurance as opposed to have a non-existence market with no one obtaining insurance. This contrasts with the loss of market surplus of \$182 shown in Panel B of Figure 4. After learning their willingness to pay for insurance, $D(s)$, individuals would only be willing to contribute an average of \$182 per person to set prices to maximize economic surplus. In this case, an ex-ante welfare perspective leads to fairly different conclusions about optimal insurance prices and the welfare cost of adverse selection.

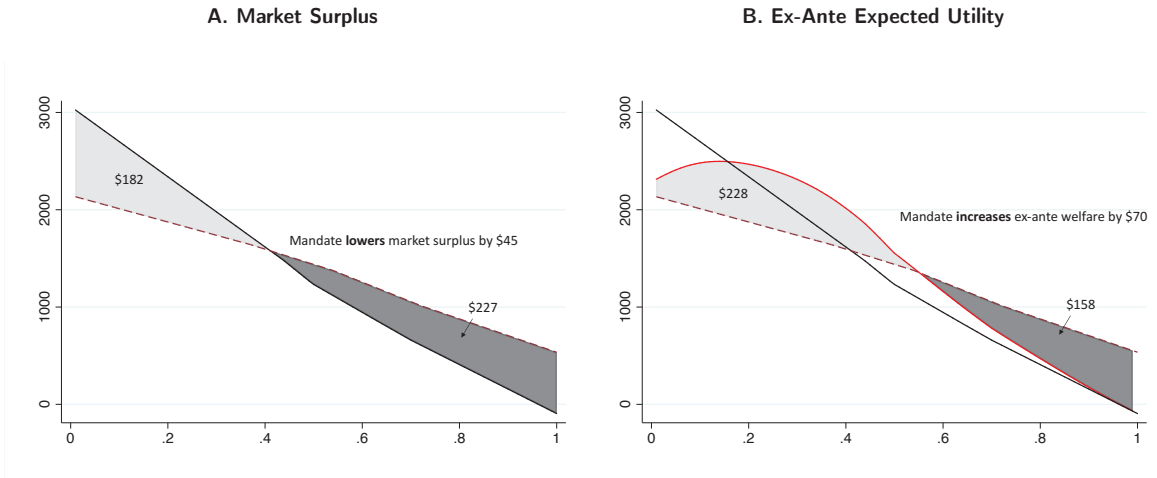
Mandates What is the welfare impact of requiring insurance coverage, $s = 1$? Figure 6 depicts the welfare impact of imposing a mandate from a market surplus perspective (Panel A) and an ex-ante expected utility perspective (Panel B). From a market surplus perspective, insuring the first 41% with the highest willingness to pay yields a surplus of \$182. In contrast, the lost sur-

²⁰It is perhaps surprising that $EA(s) < 0$ for low values of s . Mathematically, this is because for low values of s , $C(s) - D(s) < sD'(s)$. Economically, this means that expanding the size of the insurance market actually generates a Pareto improvement, as it can lower prices for both the insured and uninsured because the marginal cost of the new enrollees is sufficiently below their willingness to pay. As a result, market surplus actually over-states the welfare impact of expanding the insurance market for low values of s .

plus from insuring the remaining 59% of the market is \$227. Thus, a mandate imposes a net loss of market surplus of \$45. On aggregate, individuals in the market would be willing to pay \$45 per person to prevent a mandate.

In contrast, from an ex-ante expected utility perspective, the value of insuring the 55% of the market with the highest willingness to pay is \$228, and the cost of insuring the remainder of the market is \$158. Prior to learning their willingness to pay, individuals would pay an average of \$70 per person to have a mandate instead of having no insurance. Mandates increase ex-ante expected utility, but decrease observed market surplus. In this sense, an ex-ante welfare perspective can lead to different conclusions about the desirability of government mandates.

Figure 6: Welfare Impact of Mandating Insurance ($s = 1$) Relative to Full Unraveling ($s = 0$)



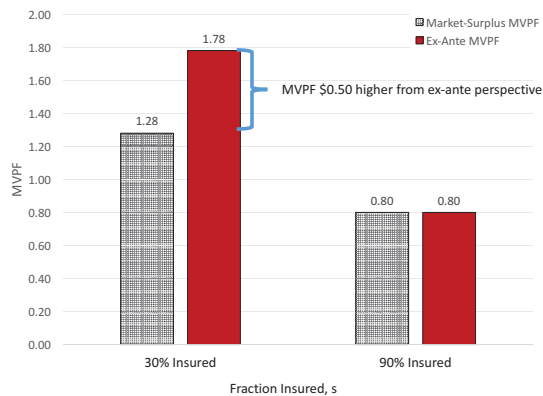
Non-budget neutral policies In practice, the insurance subsidies in Massachusetts are not paid by low-income individuals choosing to forego insurance. Rather, they are paid by other taxpayers out of government funds. Here, I show how to estimate the marginal value of public funds of higher/lower sub-

sidies. This is given by

$$MVPF(s) = \frac{1}{1 + \frac{C(s) - D(s)}{s(-D'(s))}} (1 + (1 - s) \gamma (D(s^*) - E[D(s) | s \geq s^*]))$$

where $\beta(s) = \gamma (D(s^*) - E[D(s) | s \geq s^*])$ is the difference in marginal utilities between the insured and uninsured. Figure 7 calculates this MVPF of greater insurance subsidies. I do so for two values of s that corresponds to the range of take-up estimates in Finkelstein et al. (2017). The price variation in Finkelstein et al. (2017) leads to between 30-90% of the market choosing to purchase insurance. I therefore consider the MVPF of greater subsidies when $s = 0.3$ and $s = 0.9$.

Figure 7: MVPF for Health Insurance Subsidies for Low-Income Adults



When 30% of the market is insured, $s = 0.3$, annual costs are given by $C(0.3) = 1738$, willingness to pay is given by $D(0.3) = 1978$, and the slope of willingness to pay is given by $D'(0.3) = -3610$. The average willingness to pay for those with $s \geq 0.3$ is 853. Therefore, the MVPF is given by

$$\begin{aligned} MVPF(0.3) &= \frac{1}{1 - \frac{1978 - 1738}{0.3 * 3610}} (1 + .3 * 5x10^{-4} * (1978 - 853)) \\ &= 1.28 * 1.39 \\ &= 1.78 \end{aligned}$$

From a market surplus perspective, the willingness to pay for additional subsi-

dies is \$1.28 per dollar of government spending. Every \$1 of subsidy generates \$1.28 lower prices for the insured. This is greater than \$1 because the marginal types that are induced to enroll from lower prices have a lower cost of being insured, $D(0.3) > C(0.3)$. But, behind the veil of ignorance, these lower prices to the insured have additional value because the insured have a 40% higher marginal utility of income relative to the average person in this setting. So, the ex-ante MVPF of larger insurance subsidies is 1.78, not 1.28.

When most of the market already has insurance, $s = 0.9$, the willingness to pay of the marginal type is below her cost, $D(s) < C(s)$, so that $\frac{1}{1 + \frac{C(s) - D(s)}{s(-D'(s))}} = 0.8$. Moreover, the distinction between market surplus and ex-ante expected utility is smaller. This is because the difference between the implied consumption of the insured relative to the average in the population population is smaller. In this case, the MVPF is 0.8 for both the ex-ante and market surplus perspective.

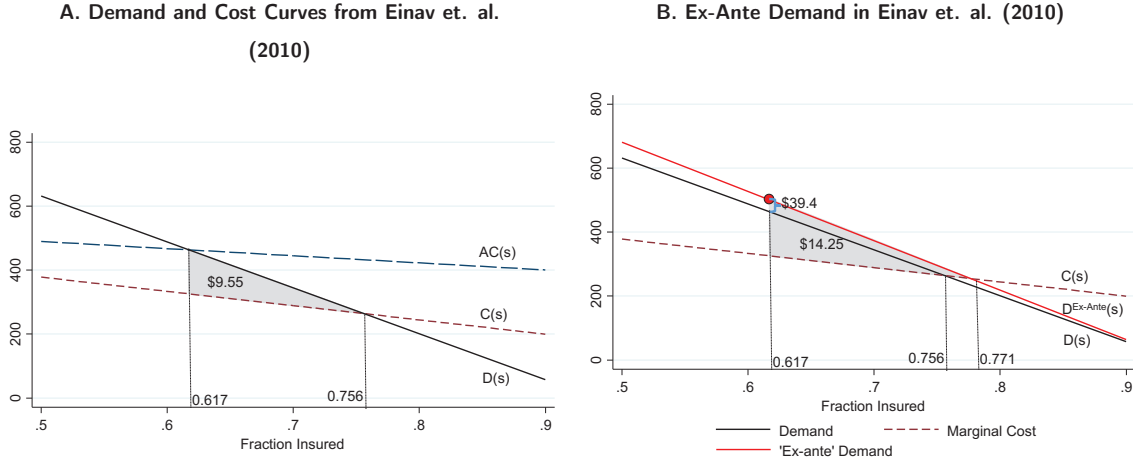
5.2 More vs. Less Generous Employer-Provided Health Insurance

Einav et al. (2010) use variation in prices across business units of Alcoa to estimate demand and cost curves for a more generous health insurance policy relative to a less generous policy. Figure 8, Panel A presents their demand and cost curve estimates. A competitive equilibrium in this environment would result in $s^{CE} = 61.7\%$ of the market purchasing the more generous policy, reflected by the intersection between the average cost curve and the demand curve in Panel A. This occurs with a price of $D(s^{CE}) = AC(s^{CE}) = \463.5 . But, those that are indifferent to purchasing insurance at a price of \$463.5 on average impose a cost on the insurance company, $C(s^{CE})$, that is less than their willingness to pay. Aggregating across these potential trades for which demand is above marginal cost, the lost surplus from adverse selection is \$9.57. This is given by the shaded region in Panel A of Figure 8.

How does this compare to an ex-ante measure of the welfare cost of adverse selection? Panel B of Figure 8 presents the estimated $D^{Ex-Ante}(s)$ assuming

$\gamma = 5 \times 10^{-4}$.²¹ The ex-ante demand curve intersects the cost curve when a fraction 77.7% of the market is insured. This is fairly similar to the size of the market that maximizes market surplus of 75.6%.

Figure 8: Ex-Ante WTP in Einav, Finkelstein, and Cullen (2010)



Integrating between the ex-ante demand curve and cost curve between 61.7% and 77.7% yields an ex-ante welfare cost of adverse selection of \$14.25. Put differently, from an ex-ante perspective prior to learning their demand

²¹To illustrate its calculation at $s = 0.617$, note that the linearity of demand in Einav et al. (2010) implies $E[D(s^*) - D(s) | s \geq s^*] = D' \frac{(1-s^*)}{2}$. So, equation (22) becomes

$$EA(s) = (1-s)(C(s) - D(s) - sD')\gamma D' \frac{(1-s)}{2}$$

At $s = 61.7\%$, the ex-ante component of demand is given by

$$EA(0.617) = (1-0.617)(MC(0.617) - D(0.617) - 0.617D')\gamma D' \frac{(1-0.617)}{2}$$

Plugging in $MC(0.617) = \$325.88$, $D(0.617) = \$463.5$, and $D' = -1435.97$, along with $\gamma = 5 \times 10^{-4}$, yields

$$EA(0.617) = \$39.4$$

This \$39.4 is reflected by the difference between $D(0.617)$ and $D^{Ex-Ante}(0.617)$ in Panel B of Figure 8.

for insurance in this market, individuals would be willing to pay \$14.25 to have an optimally priced insurance market in which 77.7% of the population is insured. This suggests that market surplus captures the two thirds (67%) of the ex-ante welfare cost of adverse selection.

Overall, the value of insurance is higher from an ex-ante expected utility perspective in the Einav et al. (2010) setting than is suggested by market surplus measures of welfare. But, as illustrated in Figure 8, the difference is empirically small. In contrast to the health insurance for low-income adults setting above, both the market surplus and ex-ante welfare perspective lead to similar conclusions about optimal policies towards the insurance market.

Size of Insurable Risk Why is this? This is an illustration of a general phenomenon that the ex-ante adjustment tends to be increasing in the size of the insurable loss. Because the Finkelstein et al. (2017) setting considers a willingness to pay for a full insurance (i.e. uninsured versus insured) product, the size of the insured risk is larger than the top-up (more versus less generous) insurance product considered in Einav et al. (2010).

To see why the size of the insurable risk matters, suppose that one scales willingness to pay and cost curves by a factor α so that willingness to pay goes from $D(s) \rightarrow \alpha D(s)$ and costs go from $C(s) \rightarrow \alpha C(s)$. Equation (22) shows that the size of the ex-ante adjustment is increasing in the square of the increase in willingness to pay and costs, $EA(s) \rightarrow \alpha^2 EA(s)$. When the insurable event comprises a large fraction of one's income, the difference between the marginal utility of the insured and uninsured is larger. Therefore, the distinction between ex-ante and observed willingness to pay is most important in settings where the risk comprises a larger fraction of one's consumption or income.²²

²²Similarly, the MVPF is increasing roughly linearly in $\beta(s)$, so that if one increases demand and cost by a factor of α , one would expect the difference in marginal utilities between insured and uninsured to scale by a factor of α , so that the MVPF of insurance subsidies is increasing in the size of the insurable risk.

6 Conclusion

Measuring willingness to pay is generally thought to be the gold standard for measuring welfare. And, there's a growing literature that either directly or indirectly measures willingness to pay for insurance and uses it to make welfare inferences. But, in the case of insurance it misses the value of insurance against the portion of risk that has been realized at the time of measuring willingness to pay. This paper develops methods to measure the ex-ante expected utility impact of policies that affect insurance markets.

The method retains the empirical transparency of the revealed-preference approach. But, it introduces additional sufficient statistics to capture the value of insurance against the portion of the risk that has already been realized at the time of purchasing insurance. In general, the key sufficient statistic required to make this calculation is the difference in marginal utilities of income for the insured and uninsured. I provide a benchmark method for estimating this difference using market demand and cost curves combined with a measure of risk aversion.

The ex-ante welfare distinction can matter in practice. For example, in the low-income health insurance setting of Finkelstein et al. (2017), the estimates suggest that mandates are optimal relative to a competitive market allocation, despite the fact that the mandates generate lower market surplus. Policies that maximize ex-ante utility generally involve deadweight loss.

Future work can extend the analysis in many directions. For example, one could consider choices over a greater number of insurance contracts instead of the binary choice considered here. From an empirical perspective, it would be valuable to understand the difference in marginal utilities between the insured and uninsured. Although most models of insurance imply higher marginal utilities of consumption for those with unobservables that lead them to purchase insurance, it is possible that this is incorrect in some settings. For example, if those who choose to forego insurance under the Affordable Care Act are doing so because they have a negative liquidity shock that prevents them from being able to afford insurance, it could be the case that the insured actually have a lower marginal utility of income than the uninsured ($\beta(s) < 0$). If $\beta(s) < 0$, then increasing the mandate penalty and increasing insurance sub-

sidies to expand the size of the insurance may actually deliver lower welfare than is suggested by market surplus.

Lastly, the approach can be extended to cases in which one does not directly observe willingness to pay. While the approach and notation here focused on the context of the Einav et al. (2010) in which the researcher has estimated demand and cost curves, one could readily extend the approach to settings where willingness to pay is inferred from other choices. For example, Gallen (2014) infers willingness to pay from income and labor supply choices around Medicaid eligibility notches. Such an approach measures willingness to pay for insurance against the risk that remains after choosing their income and labor supply. But, to the extent to which individuals make these choices after learning their health conditions (as evidenced in Gallen (2014)), the resulting willingness to pay measures will not include the value of Medicaid as insurance against those health conditions. The ex-ante methods developed here could perhaps be adapted to recover the ex-ante expected utility impact of Medicaid.

In the end, this paper provides a method to translate observed willingness to pay and cost curves into a framework that can help identify policies that maximize expected utility.

References

- Baily, M. N. (1978). Some aspects of optimal unemployment insurance. *Journal of Public Economics* 10(3), 379–402. 2, 2, 4, 4
- Chetty, R. (2006). A general formula for the optimal level of social insurance. *Journal of Public Economics* 90(10), 1879–1901. 2, 2, 4, 4
- Einav, L., A. Finkelstein, and M. R. Cullen (2010). Estimating welfare in insurance markets using variation in prices. *Quarterly Journal of Economics* 125(3). 1, 2, 2, 8, 3.1, 3.2, 3.2.1, 5, 5.2, 21, 5.2, 6, B
- Einav, L., A. Finkelstein, and H. Williams (2016). Paying on the margin for medical care: Evidence from breast cancer treatments. *American Economic Journal: Economic Policy* 8(1), 52–79. 2

- Finkelstein, A., N. Hendren, and E. F. Luttmer (2016). The value of medicaid: Interpreting results from the oregon health insurance experiment. *NBER Working Paper*,, <http://scholar.harvard.edu/hendren/publications/value-medicaid-interpreting-results-oregon-health-insurance-experiment>. 2, 3.2
- Finkelstein, A., N. Hendren, and M. Shepard (2017). Subsidizing health insurance for low-income adults: Evidence from massachusetts. *NBER Working Paper No. 23668*. 1, 2, 3.2, 3.2.2, 5, 5.1, 19, 5.1, 5.2, 6, D
- Finkelstein, A., E. F. Luttmer, and M. J. Notowidigdo (2013). What good is wealth without health: The effect of health on the marginal utility of consumption. *Journal of the European Economic Association* 11(s1), 221–258. 4.2
- Gallen, T. S. (2014). Measuring the value of Medicaid using TennCare. *Working Paper*. 6
- Hackmann, M. B., J. T. Kolstad, and A. E. Kowalski (2015). Adverse selection and an individual mandate: When theory meets practice. *The American Economic Review* 105(3), 1030–1066. 1
- Handel, B., I. Hendel, and M. D. Whinston (2015). Equilibria in health exchanges: Adverse selection versus reclassification risk. *Econometrica* 83(4), 1261–1313. 2, 3.2, 4.1, 5.1, 18
- Hendren, N. (2014). The inequality deflator: Interpersonal comparisons without a social welfare function. *NBER Working Paper*, <http://www.nber.org/papers/w20351>. 3.2.2
- Hendren, N. (2016). The policy elasticity. *Tax Policy and the Economy* 30. 2, 3.2.2
- Hirshleifer, J. (1971). The private and social value of information and the reward to inventive activity. *The American Economic Review* 61(4), 561–574. 1, 2

Online Appendix: Not For Publication

A Calculation of dp_U and dp_I in budget-neutral market expansion in Section 2

Note that differentiating the budget constraint of the insurer yields

$$\begin{aligned}\frac{d}{ds} [sp_I + (1-s)p_U] &= \frac{d}{ds} \int_0^s C(s) \\ p_I - p_U + sd p_I + (1-s) d p_U &= C(s)\end{aligned}$$

where $C(s) = m(s) = D(s) = p_I - p_U$. So,

$$d p_U = -\frac{s}{1-s} d p_I$$

Now, consider the demand identity:

$$p_I - p_U = D(s)$$

differentiating and re-arranging yields:

$$\begin{aligned}d p_I - d p_U &= D'(s) ds \\ d p_I &= D'(s) ds + d p_U \\ d p_I &= D'(s) ds - \frac{s}{1-s} d p_I \\ \frac{1}{1-s} d p_I &= D'(s) ds \\ d p_I &= (1-s) D'(s) ds\end{aligned}$$

and

$$\begin{aligned}d p_U &= d p_I - D'(s) ds \\ d p_U &= -s D'(s) ds\end{aligned}$$

B Measuring Risk Aversion

In addition to the demand and cost curves in the Einav et al. (2010) framework, measuring ex-ante willingness requires an estimate of risk aversion, $\gamma(s)$. This can be imported from another setting as in Section 5.2. Here, I illustrate how one can in principle infer risk aversion within the demand and cost curve setup. Risk aversion is revealed by comparing individual's willingness to pay for insurance to the reduction in variance of expenditures that is provided by the insurance product. For example, it is well-known that if preferences have a constant absolute risk aversion and the risk of medical expenditures is normally distributed (i.e. a "CARA-Normal" model), then the markup individual's are willing to pay for insurance is given by the variance reduction offered by the insurance multiplied by $\frac{\gamma(s)}{2}$.

More generally, one can consider a second-order Taylor approximation to equation (7) that characterizes willingness to pay, $D(\tilde{s})$. Let $p(\tilde{s}) = \frac{\partial x}{\partial m}$ denote the price of additional medical spending when insured. Under the additional assumption that $u_{mm} = 0$, then it is straightforward to show²³ that the

²³To see this, suppress notation w.r.t. θ and condition all expectations on \tilde{s} . Let $(\bar{c}, \bar{m}, \bar{\theta})$ denote the average bundle of an \tilde{s} type. Taking a Taylor expansion to the utility function around this bundle in equation (7) yields

$$\begin{aligned} & u_c \left(E \left[y - x^I - D(\tilde{s}) - p_U - \bar{c} \right] \right) + \frac{1}{2} u_{cc} \left(E \left[y - x^I - D(\tilde{s}) - p_U - \bar{c} \right]^2 \right) + u_m E \left[m^I - \bar{m} \right] \\ &= u_c \left(E \left[y - p_U \right] - \bar{c} \right) + \frac{1}{2} u_{cc} \left(E \left[y - p_U - \bar{c} \right]^2 \right) + u_m E \left[m^U - \bar{m} \right] \end{aligned}$$

or

$$\begin{aligned} & u_c \left(\left(E \left[y - x^I - D(\tilde{s}) - p_U - \bar{c} \right] \right) - \left(E \left[y - p_U \right] - \bar{c} \right) \right) \\ &= \frac{-1}{2} u_{cc} \left[E \left[y - x^I - D(\tilde{s}) - p_U - \bar{c} \right]^2 - E \left[y - p_U - \bar{c} \right]^2 \right] + u_m E \left[m^U - m^I \right] \end{aligned}$$

or

$$\begin{aligned} D(\tilde{s}) - \left(E \left[m^I - x^I \right] \right) &= \frac{\gamma(\tilde{s})}{2} V + \left(\frac{u_m}{u_c} - 1 \right) E \left[m^I - m^U \right] \\ \gamma(\tilde{s}) &= 2 \frac{D(\tilde{s}) - MC(\tilde{s}) + \left(1 - \frac{u_m}{u_c} \right) E \left[m^I - m^U \right]}{V} \end{aligned}$$

where $MC(\tilde{s}) = E \left[m^I - x^I | \tilde{s} \right]$ is the cost to the insurer of enrolling the type \tilde{s} .

coefficient of absolute risk aversion is given by:

$$\gamma(\tilde{s}) = 2 \frac{D(\tilde{s}) - C(\tilde{s}) + (1 - p(\tilde{s})) E[m^I - m^U | \tilde{s}]}{V} \quad (23)$$

where $D(\tilde{s}) - C(\tilde{s})$ is the markup individuals of type \tilde{s} are willing to pay above the cost they impose on the insurer, V is approximately the reduction in variance of consumption offered by the insurance:

$$V = E \left[(y - x^U - p_U - \bar{c})^2 | \tilde{s} \right] - E \left[(y - x^I - D(\tilde{s}) - p_U - \bar{c})^2 | \tilde{s} \right]$$

and $(1 - p(\tilde{s})) E[m^I - m^U | \tilde{s}]$ is a correction term to account for moral hazard. $E[m^I(\theta) - m^U(\theta) | \tilde{s}]$ is the causal effect of insurance on medical spending to a type θ . If $p(\tilde{s}) < 1$, some of this additional cost that is imposed on the insurer will not be fully valued by the individual.

In this sense, one needs to observe two additional pieces of information in order to generate an internal measure of risk aversion, $\gamma(\tilde{s})$: (1) the impact of insurance on medical spending for type \tilde{s} , $E[m^I(\theta) - m^U(\theta) | \tilde{s}]$ and (2) the impact of insurance on the variance of consumption, $V(\tilde{s})$. In this sense, one need not necessarily rely on an external measure of risk aversion, but can instead infer risk aversion from individuals revealed willingness to pay to reduce their variance in consumption.

C Case when $\frac{\partial D}{\partial p_U} \neq 1$

This appendix outlines the more general case when willingness to pay for the insurance policy depends not just on the relative price of insurance, $p_I - p_U$, but separately depends on p_I and p_U (e.g. because of income effects). To capture this, let $D(\tilde{s}, p_U)$ denote the price that a type \tilde{s} is willing to pay for insurance when facing a price p_U of being uninsured. This solves

$$E[u(y(\theta) - x(m^I(\theta); \theta) - D(\tilde{s}, p_U), m^I(\theta); \theta) | \tilde{s}] = E[u(y(\theta) - m^U(\theta) - p_U, m^U(\theta); \theta) | \tilde{s}]$$

Here, I re-state the main proposition for this general case. It is straightforward to see that the main

Proposition. *The marginal welfare impact of expanding the size of the insur-*

ance market from s^* to $s^* + ds$ is given by

$$\frac{V'(s^*)}{E[u_c(y(\theta) - p_I(s), m^I(\theta); \theta) | s \leq s^*]} \approx \underbrace{p_I(s^*) - p_U(s^*) + EA(s^*) - MC(s^*)}_{\text{Ex-Ante Demand}} \quad (24)$$

where $EA(s^*)$ is the ex-ante value of expanding the size of the insurance market,

$$EA(s^*) = \frac{1 - s^*}{1 + s^* \left(\frac{\partial D}{\partial p_U} - 1 \right)} \underbrace{\left(MC(s^*) - (p_I(s^*) - p_U(s^*)) - s^* \frac{\partial D}{\partial s} \right)}_{\text{Transfer from Uninsured to Insured}} \beta(s^*) \quad (25)$$

and $\beta(s)$ is the percentage difference in marginal utilities of income for the insured relative to the uninsured,

$$\beta(s) = \frac{E[u_c(y(\theta) - p_I(s), m^I(\theta); \theta) | \tilde{s} \leq s] - E\left[\frac{\partial D(\tilde{s}, p_U(s))}{\partial p_U} u_c(y(\theta) - D(\tilde{s}, p_U(s)), m^I(\theta); \theta) | \tilde{s} \geq s\right]}{E[u_c(y(\theta) - p_I(s), m^I(\theta); \theta) | \tilde{s} \leq s]} \quad (26)$$

It is straightforward to see that the main result in Proposition 1 is obtained by setting $\frac{\partial D}{\partial p_U} = 1$. For brevity, the proof of this proposition is provided in the proof of Proposition 1 below.

D Proof of Proposition 1

This Appendix walks through the proof of Proposition 1. I consider the general case in Appendix C that allows for take-up to depend separately on p_I and p_U , and use the results to consider the sub-case when the purchase decision only depends on the relative price, $p_I - p_U$.

Let $p_I(s)$ and $p_U(s)$ satisfy the resource constraint (11) and the constraint, $p_I(s) = D(s, p_U(s))$ when fraction s of the market purchasing insurance when facing those prices. Ex-ante expected utility, $W(s)$, is given by

$$\begin{aligned} W(s) &= \int_0^s E[u(y(\theta) - p_I(s), m^I(\theta); \theta) | \tilde{s}] d\tilde{s} \\ &\quad + \int_s^1 E[u(y(\theta) - m^U(\theta) - p_U(s), m^U(\theta); \theta) | \tilde{s}] d\tilde{s} \end{aligned}$$

Substituting the demand function yields:

$$W(s) = \int_0^s E[u(y(\theta) - p_I(s), m^I(\theta); \theta) | \tilde{s}] d\tilde{s} \\ + \int_s^1 E[u(y(\theta) - D(\tilde{s}, p_U(s)), m^I(\theta); \theta) | \tilde{s}] d\tilde{s}$$

so the marginal welfare impact is given by

$$W'(s) = -sp'_I(s) E[u'(y(\theta) - p_I(s), m^I(\theta); \theta) | \tilde{s} \leq s] \\ - (1-s)p'_U(s) E\left[\frac{\partial D(\tilde{s}, p_U(s))}{\partial p_U} u_c(y(\theta) - D(\tilde{s}, p_U(s)), m^I(\theta); \theta) | \tilde{s} \geq s\right]$$

Now,

$$sp_I(s) + (1-s)p_U(s) = sAC(s)$$

so that when $G'(s) = 0$

$$p_I(s) + s\frac{dp_I}{ds} + (1-s)\frac{dp_U}{ds} - p_U(s) = MC(s)$$

or

$$sp'_I(s) + (1-s)p'_U(s) = MC(s) - p_I(s) + p_U(s)$$

or

$$sp'_I(s) + (1-s)p'_U(s) = MDWL(s)$$

where $-MDWL(s) = p_I(s) - p_U(s) + MC(s)$. If there is sufficiently high DWL from expanding the insurance market, both p_I and p_U will go up (as was seen for low values of s in the example from Finkelstein et al. (2017)). But, if there is sufficiently high surplus, the resource constraint will imply that both prices must go down. For intermediate ranges of DWL, one expects the price of insurance to go down and the price of being uninsured to go up.

So, adding and subtracting $(1-s)p'_U(s) E[u_c(y(\theta) - p_I(s), m^I(\theta); \theta) | \tilde{s} \leq s]$ and then dividing by $E[u_c(y(\theta) - p_I(s), m^I(\theta); \theta) | \tilde{s} \leq s]$ yields

$$\frac{W'(s)}{E[u_c(y(\theta) - p_I(s), m^I(\theta); \theta) | \tilde{s} \leq s]} \\ = -[sp'_I(s) + (1-s)p'_U(s)] + (1-s)p'_U(s) \\ \times \left(\frac{E[u_c(y(\theta) - p_I(s), m^I(\theta); \theta) | \tilde{s} \leq s] - E\left[\frac{\partial D(\tilde{s}, p_U(s))}{\partial p_U} u_c(y(\theta) - D(\tilde{s}, p_U(s)), m^I(\theta); \theta) | \tilde{s} \geq s\right]}{E[u_c(y(\theta) - p_I(s), m^I(\theta); \theta) | \tilde{s} \leq s]} \right)$$

or

$$\begin{aligned} & \frac{W'(s)}{E[u_c(y(\theta) - p_I(s), m^I(\theta); \theta) | \bar{s} \leq s]} \\ &= -MDWL + (1-s)p'_U(s)\beta(s) \\ & \times \left(\frac{E[u_c(y(\theta) - p_I(s), m^I(\theta); \theta) | \bar{s} \leq s] - E\left[\frac{\partial D(\bar{s}, p_U(s))}{\partial p_U} u_c(y(\theta) - D(\bar{s}, p_U(s)), m^I(\theta); \theta) | \bar{s} \geq s\right]}{E[u_c(y(\theta) - p_I(s), m^I(\theta); \theta) | \bar{s} \leq s]} \right) \end{aligned}$$

where $p'_U(s) > 0$. Note that one can express $-p'_U(s)$ as follows. The derivative of the resource constraint with respect to s equals the difference between the marginal price of insurance and marginal cost of insuring the type s : $-sp'_I(s) - (1-s)p'_U(s) = (p_I(s) - p_U(s)) - C(s)$. Note that

$$sp_I(s) + (1-s)p_U(s) = sAC(s) \quad (27)$$

so that

$$p_I(s) + sp'_I(s) + (1-s)p'_U(s) - p_U(s) = C(s)$$

or

$$sp'_I(s) + (1-s)p'_U(s) = C(s) - p_I(s) + p_U(s)$$

Moreover, differentiating equation (27) yields $p'_I(s) = \frac{\partial D}{\partial s} + \frac{\partial D}{\partial p_U} p'_U(s)$. To see this, note that:

$$\begin{aligned} -sp'_I(s) - (1-s)p'_U(s) &= (p_I(s) - p_U(s)) - C(s) \\ -s \left[\frac{\partial D}{\partial s} + \frac{\partial D}{\partial p_U} p'_U(s) \right] - (1-s)p'_U(s) &= (p_I(s) - p_U(s)) - C(s) \\ -p'_U(s) \left[1 + s \left(\frac{\partial D}{\partial p_U} - 1 \right) \right] &= (p_I(s) - p_U(s)) - C(s) + s \frac{\partial D}{\partial s} \\ -p'_U(s) &= \frac{1}{1 + s \left(\frac{\partial D}{\partial p_U} - 1 \right)} \left[(p_I(s) - p_U(s)) - C(s) + s \frac{\partial D}{\partial s} \right] \end{aligned}$$

Under the additional approximation that $\frac{\partial D}{\partial p_U} = 1$, and replacing $D(s) = p_I(s) - p_U(s)$, this yields

$$\frac{W'(s)}{E[u_c(y(\theta) - p_I(s), m^I(\theta); \theta) | \bar{s} \leq s]} = D(s) - C(s) + (1-s) \left(D(s) - C(s) + s \frac{\partial D}{\partial s} \right) \beta(s)$$

which concludes the proof.

E Proof of Proposition 2

Every dollar the government spends on additional subsidies leads to $\frac{1}{1 - \frac{C(s) - D(s)}{sD'(s)}}$ dollars accruing to the insured. Hence, from behind the veil of ignorance, this generates a welfare impact of $E[u_c|Insured] \frac{1}{1 - \frac{C(s) - D(s)}{sD'(s)}}$. From behind the veil of ignorance, \$1 of additional resources leads to an increase in utility of $E[u_c]$. Hence, the MVPF is given by

$$MVPF(s) = \frac{E[u_c|Insured]}{E[u_c]} \frac{1}{1 - \frac{C(s) - D(s)}{sD'(s)}}$$

Now, note that

$$\begin{aligned} E[u_c] &= sE[u_c|Insured] + (1 - s)E[u_c|Uninsured] \\ &= E[u_c|Insured] + (1 - s)(E[u_c|Uninsured] - E[u_c|Insured]) \end{aligned}$$

so that

$$\frac{E[u_c|Insured]}{E[u_c]} = 1 - (1 - s)\beta(s)$$

Hence, the MVPF is given by

$$MVPF(s) = \frac{1 + (1 - s)\beta(s)}{1 + \frac{C(s) - D(s)}{s(-D'(s))}}$$

F Alternative Risk Aversion Estimates

Figure A1 Presents estimates of the ex-ante willingness to pay curve under the assumption that the coefficient of relative risk aversion is 3. The optimal size of the insurance market becomes 47% as opposed to 55% in the baseline specification. The welfare impact of going from $s = 0$ to $s = 47\%$ is \$193. The welfare impact of insuring the remaining 53% of the market is \$192. Hence, the welfare impact of a mandate is \$1. Prior to learning one's willingness to pay, individuals would be willing to pay \$1 to have a mandated insurance market with $s = 1$ relative to a world with $s = 0$.

Figure A1: Ex-Ante WTP Under Alternative Risk Aversion Calibrations

