Measuring Ex-Ante Welfare in Insurance Markets

Nathaniel Hendren*

March, 2019

Abstract

The willingness to pay for insurance captures the value of insurance against only the risk that remains when choices are observed. This paper develops tools to measure the ex-ante expected utility impact of insurance subsidies and mandates when choices are observed after some insurable information is revealed. The approach retains the transparency of using reduced-form willingness to pay and cost curves, in contrast to structural approaches that require fully specifying the choice environment and information sets of individuals. To do so, the approach utilizes an additional sufficient statistic: the difference in marginal utilities between insured and uninsured. I provide an approach to estimate this statistic that uses only reduced-form willingness to pay and cost curves, combined with either (i) a measure of risk aversion or (ii) the reduction in variance of out of pocket expenditures generated by insurance. I illustrate the approach using existing willingness to pay and cost curve estimates from the low-income health insurance exchange in Massachusetts. Ex-ante optimal insurance prices are roughly 30% lower than prices that maximize market surplus. Mandates can increase expected utility despite increasing deadweight loss.

1 Introduction

Revealed preference theory is often used as a tool for measuring the welfare impact of government policies. Many recent applications use price variation to estimate the willingness to pay for insurance (Einav et al. (2010); Hackmann et al. (2015); Finkelstein et al. (2019); Panhans (2018)). Comparing willingness to pay to the costs individuals impose on insurers

*Harvard University, nhendren@fas.harvard.edu. I am very grateful to Raj Chetty, David Cutler, Liran Einav, Amy Finkelstein, Ben Handel, Pat Kline, Tim Layton, 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.

provides a measure of market surplus. This surplus potentially provides guidance on optimal insurance subsidies and mandates (Feldman and Dowd (1982)). If individuals are not willing to pay the costs they impose on the insurer, then greater subsidies or mandates will lower market surplus. From this perspective, subsidies and mandates would reduce welfare and be socially undesirable.

Measures of willingness to pay are generally a gold standard input into welfare analysis. But, in insurance settings they can be misleading. Insurance obtains its value by insuring the realization of risk. Often, individuals make insurance choices after learning some information about their risk. It is well-known that this can lead to adverse selection. What is less appreciated is that observed willingness to pay will not capture value of insuring against this learned information.[1] As a result, welfare conclusions based on market surplus can vary with the information that individuals have when the economist happens to observes choices. Policies that maximize observed market surplus will not generally maximize canonical measures of expected utility.

To see this, consider the decision to buy health insurance coverage for next year. Suppose some people have learned they need to undergo a costly medical procedure next year. Their willingness to pay will include the value of covering this known cost plus the value of insuring other future unknown costs. Market surplus - defined as the difference between willingness to pay and costs - will equal the value of insuring their unknown costs. But, it will not include any insurance value from covering the known costly medical procedure. This risk has already been realized when willingness to pay is observed.

Now, consider an economist seeking to measure the welfare impact of extending health insurance coverage next year to everyone through a mandate or large subsidy. The market surplus or deadweight loss generated from the policy will depend on how much people have learned about their health costs at the time the economist happened to measure willingness to pay. Existing literature (and introspection) suggests that individuals know more about expected costs and events in the near future (e.g. Finkelstein et al. (2005); Hendren (2013, 2017); Cabral (2017)). If willingness to pay had been measured earlier, market surplus could be larger because it would include the value of insuring against the costly medical procedure. This occurs even though the economic allocation generated by a mandate does not vary depending on when the economist measures willingness to pay. In contrast, traditional notions of the expected utility impact of a mandate would not depend on when the economist happens to measure willingness to pay. Expected utility provides a consistent framework for identifying optimal insurance policies that depends on economic allocations.

[1] This idea is related to Hirshleifer (1971), who shows that individuals may wish to insure against the realization of information.

The goal of this paper is to enable researchers to evaluate the impact of insurance market policies on expected utility, where the expectation is taken prior to when insurance choices are made. Traditional methods to estimating ex-ante expected utility would estimate a structural model. Among other things, the model would specify what individuals know when choosing whether to buy an insurance plan. It would then be estimated using observed insurance choices along with data on the realized utility-relevant outcomes, such as health and consumption.[2] Intuitively, if one has a structural model and knows what information has been realized when individuals choose their insurance policies, one can infer the value of insuring the risk that has been revealed before making those choices. But, in practice it is especially difficult to observe individuals' information sets when they make choices. This is especially true in insurance markets that suffer from adverse selection driven by asymmetric information.

To that aim, this paper develops a new approach to measure the expected utility impact of insurance market policies. The approach does not impose structural assumptions about individuals' information sets at the time of choice, nor does it require specifying a utility function or observing the distribution of utility-relevant outcomes in the economy. Instead, I consider a setting where the economist observes price variation in the insurance market of interest, which can be used to estimated reduced-form willingness to pay and cost curves as outlined by Einav et al. (2010). As noted above, these measures of willingness to pay do not capture the value of insurance against risk that has been realized at the time of choice. To deal with this, I characterize the additional sufficient statistics required to measure expected utility of subsidies and mandates in this environment. The first main result of the paper shows that one can measure ex-ante expected utility using one additional sufficient statistic: the difference in marginal utilities of income for those who do versus do not buy insurance. This measures how much individuals wish to move money to the state of the world in which they buy insurance. In the example above, it reflects the desire to insure against the need to undergo the costly medical procedure.

The second result of the paper shows how one can estimate this difference in marginal utilities using a benchmark approach that makes assumptions common in previous literature. The method uses only the reduced-form willingness to pay and cost curves combined with a measure of risk aversion. This additional risk aversion parameter can be assumed,

[2]For example, Handel et al. (2015) estimate the welfare value of insuring dynamic risk realizations over time. They specify individuals' information sets as the empirical cross-sectional distribution of future out-of-pocket medical costs given observable claims and health history in their data (Section 3.2, Handel et al. (2015)). This imposes that (1) individuals know how their observable characteristics map into potential medical costs, and (2) individuals have no additional information beyond what is potentially observable to the insurance company (i.e. no asymmetric information conditional on observables). See also Section IV of Einav et al. (2016) for another structural approach to measuring ex-ante expected utility.

or it can be inferred from the observed markup individuals are willing to pay for insurance, combined with the extent to which insurance reduces the variance in out of pocket expenditures. The method builds on the optimal unemployment insurance literature (Baily (1978); Chetty (2006)) that uses measures of consumption to estimate differences in marginal utilities. But, because consumption is seldom observed, I provide conditions under which one can exploit the information in the reduced-form willingness to pay curve for insurance instead of consumption.

The approach has several advantages over a traditional structural approach. Most importantly, researchers do not need to know individuals' information sets when they make insurance choices. But moreover, they also do not need to observe the ultimate economic allocations in the environment (e.g. distribution of consumption or health); nor do they need to specify a utility function. The difference in marginal utilities between insured and uninsured is sufficient when combined with the reduced-form willingness to pay and cost curves to measure expected utility. Yet, the implementation assumptions are not without loss of generality and can be restrictive in some applications. Therefore, I discuss how they can be relaxed with additional data elements.

I apply the framework to study the optimal subsidies and mandates for low-income health insurance in Massachusetts. Finkelstein et al. (2019) use price discontinuities as a function of income to estimate willingness to pay and cost curves for those with incomes near 150% of the federal poverty level (FPL). Their results show that an unsubsidized private insurance market would unravel.[3] Without subsidies, the market would not exist.

I use the framework to provide guidance on both budget neutral and non-budget neutral policies. I first consider a budget neutral mandate that requires everyone with incomes near 150% FPL to purchase insurance. The results suggest that increasing the fraction insured from a competitive allocation of 0% to 100% would generate a deadweight loss. The willingness to pay for insurance for those with incomes near 150% of the federal poverty level (FPL) is $45 per person lower than the cost of insuring everyone. Mandates would lower market surplus, and would therefore be inefficient from a market surplus perspective, as in Einav et al. (2010). But, applying the ex-ante approach shows that a mandate can increase ex-ante expected utility. If individuals had been asked their willingness to pay prior to learning their risk, the estimates suggest they would be willing to pay $70 per person for Massachusetts to impose a health insurance mandate for those with incomes near 150% FPL. Mandates increase expected utility, despite increasing deadweight loss.[4]

[3]This unraveling is due to a combination of adverse selection and uncompensated care externalities.

[4]In Appendix I, I also apply the approach to the estimates of the willingness to pay and costs in the top-up insurance market setting considered by Einav et al. (2010). In that setting, the distinction between policies that ex-ante expected utility and those that maximize observed market surplus are less pronounced.

The approach also guides optimal insurance subsidies that lead to partial insurance coverage. Market surplus is maximized when insurance premiums are \$1,581. In contrast, a price of \$1,089 maximizes ex-ante expected utility. Subsidizing prices below \$1,581 generates a marginal deadweight loss. But, lower insurance prices provide ex-ante risk protection against having to buy expensive insurance. At the ex-ante optimum, the value of ex-ante risk protection equals the marginal deadweight loss. This results in prices that are 30% lower than those that maximize market surplus.

Lastly, I evaluate non-budget neutral subsidies by estimating their marginal value of public funds (MVPF). Following Hendren (2016), the MVPF for an additional insurance subsidy is the individual's willingness to pay for it divided by its net cost to the government. The results show that an ex-ante perspective can lead to different conclusions. If insurance prices are \$2,000, the MVPF of additional subsidies (i.e. premiums of \$1,999 instead of \$2,000) would be roughly 1.2 if one used the average observed willingness to pay. At the time of purchase, individuals would be willing to pay \$1.20 for each government dollar spent on the subsidy. But, the MVPF would be 1.8 if one measured their ex-ante willingness to pay before revealing the risk known to individuals at the time of making insurance choices. As a result, insurance subsidies can be a more efficient method of redistribution than suggested by observed willingness to pay.

The benchmark implementation measures the ex-ante expected utility (before learning income or health risk) of insurance subsidies for those with incomes near 150% FPL. In this sense, welfare is measured behind a veil of ignorance. It is both ex-ante expected utility and ex-post utilitarian social welfare. This utilitarian perspective provides a natural benchmark as it does not make a normative distinction between redistribution and insurance (e.g. Harsanyi (1978)).

But, the approach does not require a utilitarian perspective, nor does it require measuring expected utility behind a complete veil of ignorance. One can conduct the analysis conditional on any observable subgroup, $X = x$ (e.g. old vs. young, male vs. female, with vs. without chronic health conditions, etc.). Doing so requires willingness to pay and cost curves for each subgroup. Applying the approach measures the expected utility impact of the policy for each $X = x$. For any policy, one could construct its MVPF for each $X = x$ and aggregate according to the social marginal utilities of income at each x (Saez and Stantcheva (2016)). One can then adopt any social welfare criteria by choosing appropriate weights for those with different values of X.

The paper considers a health insurance setting with random price variation. Yet the

This reflects a general tendency for the difference between market surplus and ex-ante welfare to be larger when the risk reflects a larger portion of individuals' budgets.

ideas extend to other settings, such as valuing social insurance. Often behavioral responses such as labor supply changes are used to measure the value of social insurance. The more individuals are willing to adjust their labor supply to become eligible for insurance, the more they value the insurance (e.g. Keane and Moffitt (1998); Gallen (2014); Dague (2014)). Yet, this captures the value of insurance against only the risk that is revealed after they adjust their behavior.[5] Similarly, other papers infer willingness to pay for social insurance from changes in consumption around a shock (e.g. Gruber (1997); Meyer and Mok (2013)). When information is revealed over time, the consumption change may vary depending on the time horizon used (Hendren (2017)). In the extreme, there may be no change around the event (e.g. smooth consumption around onset of disability or retirement). Consumption should change when information about the event is revealed, not when the event occurs. The methods in this paper can be applied to conceptualize which consumption difference is most appropriate for the desired notion of expected utility.

The rest of this paper proceeds as follows. Section 2 provides a stylized example that develops the intuition for the approach. Section 3 derives the result in a general model that shows 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 willingness to pay and cost curves combined with a measure of risk aversion. Section 5 applies the approach to health insurance subsidies for low-income adults in Massachusetts, using estimates from Finkelstein et al. (2019). Section 6 concludes.

2 Stylized Example

I begin with a stylized example that illustrates the distinction between market surplus and ex-ante expected utility, and outlines the proposed method to recover ex-ante expected utility. Suppose individuals have \$30 dollars but face a risk of losing \$$m$ dollars, where m is uniformly distributed between 0 and 10. This uniform distribution of m can be the unconditional risk faced by individuals behind a complete "veil of ignorance", or the risk that remains after conditioning on a particular observable characteristic of individuals.[6]

Consider individuals' willingnesses to pay for insurance against the realization of m. Let $D^{Ex-ante}$ denote the willingness to pay or "demand" that is measured prior to individuals

[5] Indeed, Gallen (2014) shows many individuals who respond have particularly costly health conditions. The measure of willingness to pay in Gallen (2014) will miss the value of insurance against those health conditions.

[6] For example, in the Massachusetts health insurance example in Section 5, the empirical design will condition on having income near 150% of the Federal Poverty Level, living in Massachusetts, and not having employer-provided insurance. As discussed in Section 3.2, calculating the MVPF of these policies allows the researcher to then conduct a cost and benefit of redistributing to/from this particular subpopulation.

learning anything about their particular realization of m from the uniform distribution. This solves

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

where $E\left[u\left(30-m\right)\right] = \frac{1}{10}\int_0^{10} u\left(30-m\right)dm$ is the expected utility if uninsured. Suppose individuals have a utility function with a constant coefficient of relative risk aversion of 3 (i.e. $u\left(c\right) = \frac{1}{1-\sigma}c^{1-\sigma}$ 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 $E\left[m\right] = 5$, 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

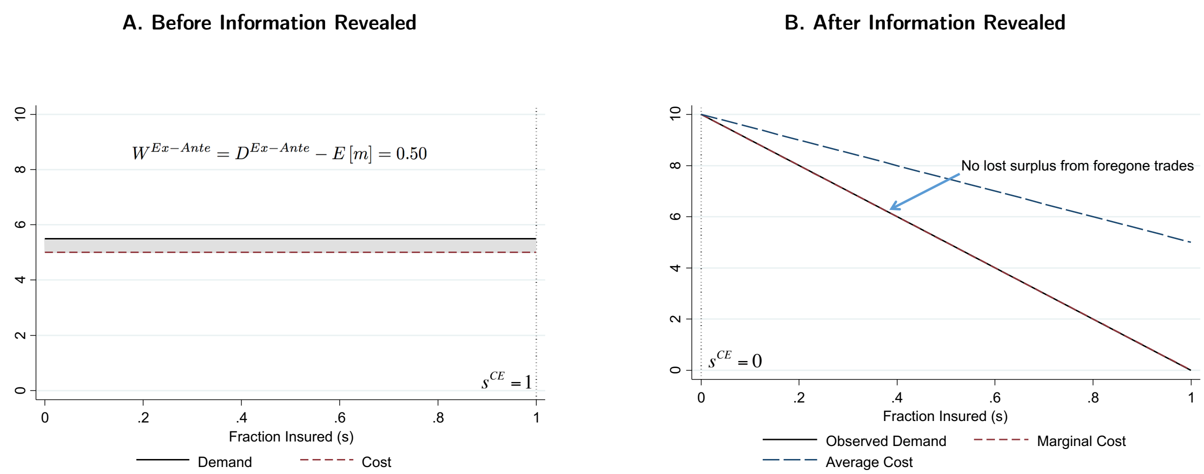


Figure 1, Panel A, draws the reduced-form demand and cost curves that would be revealed through random variation in prices in this environment, as formalized in Einav et al. (2010). The horizontal axis enumerates the population in descending order of their willingness to pay for insurance (using an index $s \in [0,1]$), 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 $D\left(s\right) = \$5.50$. 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 $C\left(s\right) = \$5$. 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. Prior to learning about m, individuals would be willing to be \$0.50 poorer and consume \$24.50 if they lived in a world with full insurance instead of being exposed to the uniform distribution of risk but consume an average of \$25.

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 equals 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 $AC(s) = E[C(S)|S \leq s]$ denote the average cost of those with willingness to pay above $D(s)$.[7] 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 expected utility if one measures expected utility prior to when all information about m is revealed to the individuals.

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, such as a utility function and an assumption about individuals' information sets at the time of choice. It would then also involve measuring the distribution of outcomes that enter the utility function, such as consumption, and use this information 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 implementations of this approach, see Handel et al. (2015), Section IV of Einav et al. (2016), or Finkelstein et al.

[7] Throughout, we will let S denote the random variable corresponding to different types and s denote a number corresponding to the size of the insurance market.

(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). Moreover, while the value of $D^{Ex-Ante}$ measures the value of full insurance ($s = 1$), the approach developed here will also allow the researcher to evaluate the expected utility impact of subsidies and mandates that lead to market outcomes with only a fraction of the market choosing to purchase insurance, $s < 1$.

To illustrate the approach proposed in this paper, 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, $sp_I + (1-s)\,p_U = sAC\,(s)$.[8]

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 $dp_I = (1-s)\,D'\,(s)\,ds$.[9]

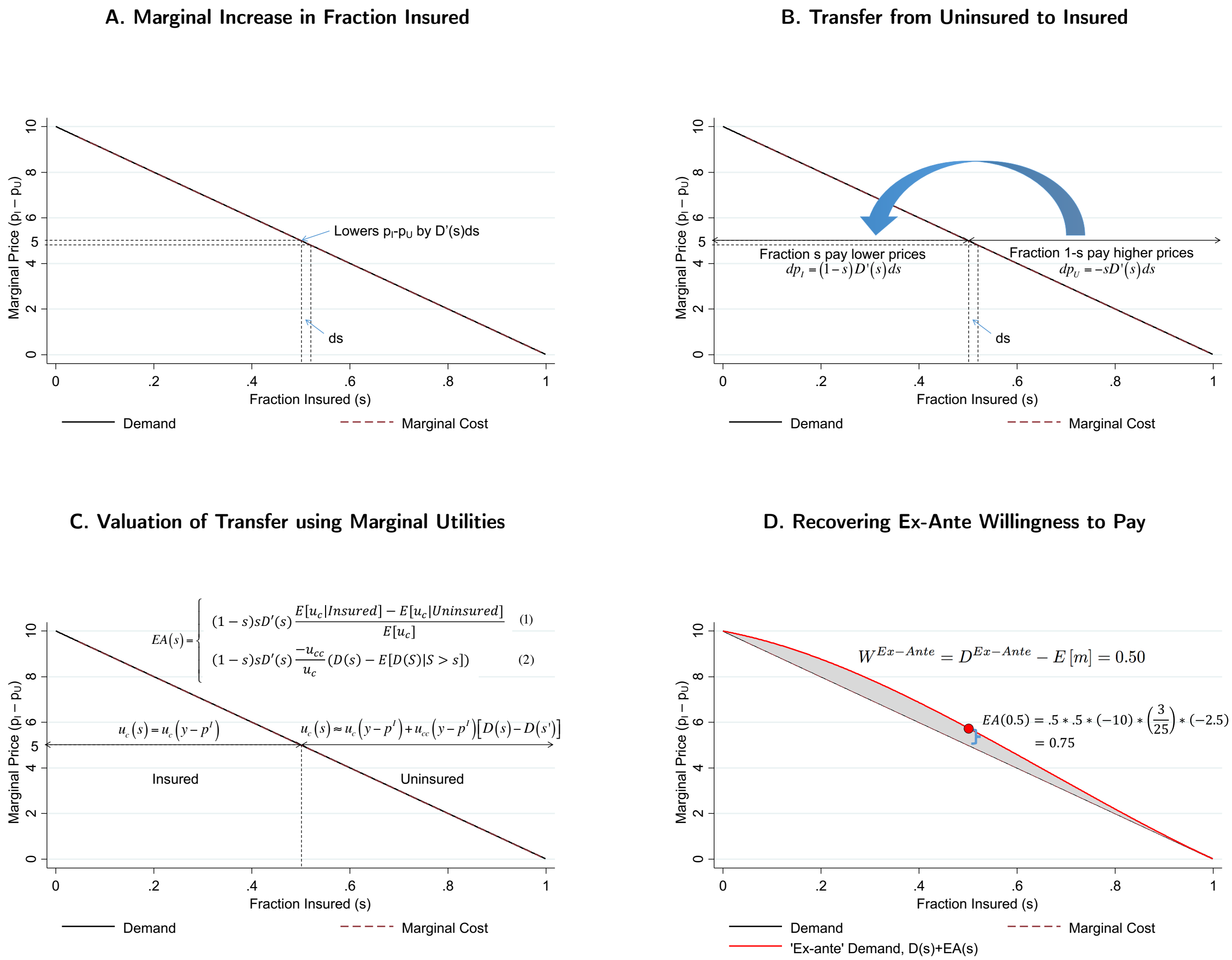
This change in insurance prices generates a transfer from the uninsured to the insured, as indicated by the blue arrow in Figure 2, Panel B. From a market surplus perspective, this transfer has no welfare impact. But, from an ex-ante expected utility perspective, these transfers have value to the extent that the marginal utilities of income differ for the insured and uninsured. If the 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 between the insured and uninsured is the key requirement for measuring ex-ante expected utility.[10]

[8] Budget neutrality is not an essential assumption. For non-budget neutral policies that use government funds to subsidize insurance, Section 3.2 shows how to construct measures of the marginal value of public funds (MVPF), as defined in Hendren (2016), for spending government resources on health insurance subsidies.

[9] Appendix A provides this calculation.

[10] As discussed in Section 3.3 and in more detail in Appendix F, the value of this transfer can be measured conditional on any observables available to the econometrician. And, the researcher can apply any particular social welfare weights to those with different observable characteristics. In this sense, the value of the transfer need not be thought of as “redistribution”, but rather as a method for generating a consistent welfare measure that does not depend on the amount of information that happens to be revealed at the time the individuals make their insurance choices.

Figure 2: Recovering Ex-Ante Willingness to Pay



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

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

where $E\left[u_c|Insured\right]$ 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

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

where $E\left[u_c|Uninsured\right]$ 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 $s+ds$:

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

Normalizing the denominator by the average marginal utility of income, $E[u_c]$, provides an ex-ante willingness to pay out of consumption taken equally from all states of the world. The first term, $s(1-s)(-D'(s))$, is loosely 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.

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. It 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 here, but will be possible in the more general model in Section 3.

The reduced-form demand curve, $D(s)$, asks how much the marginal individual is willing to pay for insurance when a fraction s of the market is insured. From a market surplus perspective, this measures the marginal willingness to pay for a larger insurance market (i.e. increasing the fraction insured from s to $s+ds$). Given $EA(s)$ in equation (2), one can now ask: how much are individuals willing to pay for a larger insurance market prior to learning anything about m? The ex-ante demand curve $D^{Ex-Ante}(s)$ answers this question by summing the observed willingness to pay, $D(s)$, with the additional ex-ante value of being able to purchase insurance at a lower marginal price, $EA(s)$,

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

In particular, the ex-ante willingness to pay to have everyone insured is equal to the average willingness to pay across all values of s, $D^{Ex-Ante} = \int_0^1 D^{Ex-Ante}(s)\,ds$ in equation (1).[11] Equations (2) and (3) are an illustration of the first main result of the paper, formalized in

[11]More 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,

Section 3. If one knows the difference in marginal utilities between the insured and uninsured, one can recover the ex-ante willingness to pay for insurance.

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. This estimation problem is analogous to the problem faced in the large literature on on optimal unemployment insurance (e.g. Baily (1978); Chetty (2006)), which seeks to estimate the difference in marginal utilities between employed and unemployed individuals. The second main result of the paper builds on the tools developed in this literature 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 can be expressed as a function of (i) the willingness to pay curve, $D(s)$, and (ii) an estimate of risk aversion.[12]

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 uninsured 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

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

where $p_I - p_U = D(s)$ is the equilibrium price of insurance when a fraction s purchases

[12] In the more general formulation in Section 3 that allows for behavioral responses to insurance (i.e. moral hazard), the difference in marginal utilities will also be a function of the cost curve, in addition to the willingness to pay curve.

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:

$$\frac{E[u_c|Insured] - E[u_c|Uninsured]}{E[u_c]} \approx \frac{-u_{cc}}{u_c}\left(D(s) - E[D(S)|s < S]\right) \tag{4}$$

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.

The combination of equations (4) and (2) provide 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.[13]

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,

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

Numerically integrating under the ex-ante demand curve in Figure 2, Panel D, yields approximately \$5.50. Not coincidentally, this equals the integral under the demand curve in Figure 1, Panel A. In this sense, the ex-ante demand curve recovers the willingness to pay

[13] For 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 general 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.

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 that are satisfied in this example and more clearly spelled out in the next section, 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 measure ex-ante expected utility.

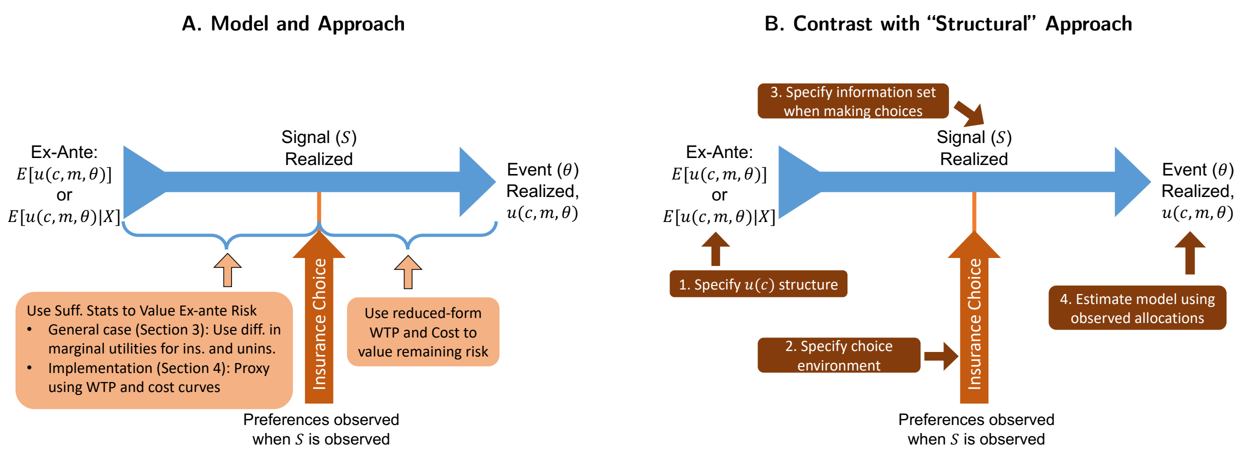
The model in this section is highly stylized. There is no moral hazard, no preference heterogeneity, and the model assumed all information about costs, m, was revealed at the time of making the insurance decision. The next two sections extend these derivations to capture more realistic features of insurance markets encountered in common empirical applications, such as the one considered in Einav et al. (2010) or Finkelstein et al. (2019). The main result of Section 3 will be to show that the difference in marginal utilities of income between the insured and uninsured continues to be the key additional sufficient statistic beyond the reduced-form demand and cost curves that is required to construct the ex-ante willingness to pay for insurance. However, the formula for $EA(s)$ is differs from above because the size of the transfer in equation (2) now depends not only on the demand curve but also on the cost curve. Section 4 will then consider the generality of the empirical estimation of the difference in marginal utilities between insured and uninsured. The key result will be to establish 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 General Model

Figure 3, Panel A, presents a graphical representation of the timeline for the general model. Individuals face uncertainty captured by the realization of a random variable θ. After learning θ, individuals realize income $y(\theta)$. They then choose their non-medical consumption, c, and medical expenditures, m, to maximize their utility function, $u(c, m; \theta)$, subject to a budget constraint that depends on whether they have an insurance policy to help cover some of their medical expenditures, m.

Prior to realizing θ but potentially after some information about θ is known to individuals, individuals choose whether or not to purchase an insurance product. Let S denote the signal known at the time of insurance purchase. The analysis will measure the ex-ante expected utility impact of subsidies and mandates towards an insurance market that exists when individuals observe S. The key distinction relative to measures of market surplus is that the expected utility will be evaluated at a point in time prior to when S is observed to the individuals making choices. This ex-ante expected utility can be unconditional, $E\left[u\left(c,m;\theta\right)\right]$, or it can involve conditioning on observable characteristics, X, $E\left[u\left(c,m;\theta\right)|X\right]$.[14] [15]

Figure 3: Model Timeline and Estimation Strategy



As outlined in Figure 3A, the strategy is to utilize the information contained in the willingness to pay and cost curves in the Einav et al. (2010) framework to measure the value of insurance against the information that remains at the time of insurance choice. Then, one steps back and characterizes the additional sufficient statistics to measure the value of insurance against the risk that has been revealed prior to making the insurance choices. The main result presented in Section 3 is that this is characterized by the difference in marginal utilities between the insured and uninsured. Section 4 then provides additional conditions

[14]To reiterate the difference relative to the market surplus perspective, note that market surplus would evaluate the welfare impact of policies from a normative perspective of after the signal S has been realized, thus not incorporating the impact of insurance policies if one were to measure welfare prior to when $\tilde{s}$ was realized to the individuals.

[15]The approach measures the expected utility of a single insurance contract period (e.g. 1 year of subsidized health insurance on the Massachusetts health exchange in a particular year). Yet, the results here are closely related to the work of Handel et al. (2015), study the impact of reclassification risk policies and the impact of exposure to changes in premium over time. Handel et al. (2015) measure the sum of expected utility from an ex-ante perspective, e.g. $W^{HHW} = E\left[\sum_{t>0} u_t\right]$, where u_t is utility in period t and the expectation, E, is taken from an ex-ante perspective. Here, one could apply the results of this paper to measure $E\left[u_t\right]$ in any period (where again E is taken from an ex-ante perspective) and sum to construct the full path of $E\left[u_t\right]$ for all t.

under which one can infer this difference in marginal utilities using again the information in the willingness to pay and cost curves, combined with a measure of risk aversion.

Panel B of Figure 3 shows how this approach differs from a more traditional "structural" approach. This approach would specify a utility function structure, a choice set, and an information set of the individuals at the time of insurance choice. It would then estimate the model using data on the ultimate outcomes in the economy (e.g. consumption and health). In contrast, the approach developed here does not require observing these ultimate allocations, nor does it require specifying an information set of the individuals. Rather, it builds upon the Einav et al. (2010) framework that does not require directly measuring the primitives of the economy.

WTP for Insurance 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:

$$c^I(\theta) + x\left(m^I(\theta);\theta\right) + p_I \leq y(\theta) \tag{5}$$

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

$$c^U(\theta) + m^U(\theta) + p_U \leq y(\theta) \tag{6}$$

where p_U is a penalty or tax paid by individuals that are uninsured. For simplicity, I consider only a binary insurance choice.[16] Let $\left\{c^I(\theta), m^I(\theta)\right\}$ denote the choice of consumption and medical spending of an insured type θ, and let $\left\{c^U(\theta), m^U(\theta)\right\}$ denote the choices of an uninsured type θ.[17]

Prior to when θ is realized, individuals have the opportunity to purchase insurance. I denote the information individuals have at the time of choosing insurance by a signal

[16] For multiple contracts, it seems natural to suppose that the marginal utilities of income in the event of making each choice are the key additional sufficient statistics, but I leave a detailed analysis of this to future work.

[17] I 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 cheaper, individuals may make riskier choices that increase health costs later on. While these effects are sometimes discussed theoretically (e.g. Ehrlich and Becker (1972)), they are almost uniformly assumed away empirically, and here I follow this convention. One could easily relax this assumption by accounting for the impact of price changes on the costs of the insured pool in the average cost curve estimates. 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.

$S \in [0,1]$.[18] After learning S, individuals know the distribution of θ given S. They use this information to decide choose to be insured and face the budget constraint in (5) or uninsured and face the budget constraint in (6). Formally, let $D(\tilde{s})$ denote the marginal price that a type S is willing to pay for insurance. This solves

$$E\left[u\left(y(\theta) - x\left(m^I(\theta);\theta\right) - D(S) - p_U, m^I(\theta);\theta\right)|S\right] = E\left[u\left(y(\theta) - m^U(\theta) - p_U, m^U(\theta);\theta\right)|S\right] \quad (7)$$

All S such that $p_I - p_U \leq D(S)$ will choose to purchase insurance, whereas types S for which $D(S) > p_I - p_U$ 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.[19] Without loss of generality, assume that $\tilde{s}$ is ordered so that demand, $D(S)$, is decreasing in S. This means that if insurance prices are p_I and p_U, a fraction s will purchase insurance where s solves $D(s) = p_I - p_U$.

Cost of Insured Population Following Einav et al. (2010), define the average cost imposed on the insurer when a fraction s of the market owns insurance by

$$AC(s) = E\left[m^I(\theta) - x\left(m^I(\theta);\theta\right)|S \leq s\right] \quad (8)$$

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 expenditures and out-of-pocket spending for those with signal[20] s:

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

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

[18] In practice, it is not essential that this signal is one dimensional. Rather, the uni-dimensionality follows from its indexing of the ordering of willingness to pay for insurance in the population.

[19] Appendix 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.

[20] This relies on the assumption noted above that individuals' choices of m and c are not affected by prices p_U and p_I beyond their impact on insurance choice. 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.

a fraction s of the market by setting prices $p_I(s)$ and $p_U(s)$:

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

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. The analysis in Section 3.1 illustrates how to conduct welfare analysis for budget neutral ($G(s) = 0$), and Section 3.2 provides an approach to non-budget neutral settings in which there is a net subsidy to those in the market ($G(s) \neq 0$).

Ex-Ante Welfare To begin with derivation of the ex-ante welfare analysis, let $c^I(\theta, p_I)$ and $c^U(\theta, p_U)$ denote the consumption choices of an insured and and uninsured type θ when prices are given by p_I and p_U. These are given by:

$$c^I(\theta, p_I) = y(\theta) - p_I - x\left(m^I(\theta);\theta\right)$$

$$c^U(\theta, p_U) = y(\theta) - p_U - m^I(\theta)$$

Next, let $W(s)$ denote the ex-ante expected utility when prices are given by $p_U(s)$ and $p_I(s)$ so that a fraction s of the market purchases insurance,

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

The ex-ante expected utility has two components, depending on whether the eventual signal realization, $\tilde{s}$, leads the individuals to choose to be insured (first term) or uninsured (second term).

Expected utility conditional on observables, X, can be defined analogously by replacing the expectations $E[\circ|S = \tilde{s}]$ with $E[\circ|X = x, S = \tilde{s}]$. In both the unconditional and conditional cases, the expectation integrates over the distribution of S. This means that, in any particular example considered by the researcher, 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. (2019), 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 below.

The ex-ante welfare impact of expanding the insurance market by a small amount starting with a fraction s insured is given by $W'(s)$, where

$$\begin{aligned} W'(s) &= -sp_I'(s)\,E\left[u_c\left(c^I(\theta,p_I(s)),m^I(\theta);\theta\right)|S\leq s\right] \\ &\quad -(1-s)\,p_U'(s)\,E\left[u_c\left(c^U(\theta,p_U(s)),m^U(\theta);\theta\right)|S\geq s\right] \end{aligned} \tag{13}$$

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, $E\left[u_c\left(c^I(\theta,p_I(s)),m^I(\theta);\theta\right)|S\leq s\right]$. The second term captures the welfare cost of having higher prices faced by the uninsured ($p_U'>0$). This price increase occurs with a chance $1-s$ and is valued using the average marginal utility of income, $E\left[u_c\left(c^U(\theta,p_U(s)),m^U(\theta);\theta\right)|S\geq s\right]$.

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. I first consider the case of budget neutral policies, and then consider non-budget neutral policies.

3.1 Budget Neutral Policies

For budget neutral policies, as in the subsidies and mandates in the Einav et al. (2010) framework, I characterize individuals' ex-ante willingness to pay out of their own income to expand the insurance market. This is given by $W'(s)/E[u_c]$, which the marginal utility impact of a larger insurance market, $W'(s)$, normalized by the marginal utility of income, $E[u_c]$.[21]

[21] To 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

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

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*

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

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

$$EA(s) = \underbrace{(1-s)(C(s) - D(s) - sD'(s))}_{\textit{Transfer from Uninsured to Insured}} \beta(s) \tag{15}$$

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

$$\beta(s) = \frac{E\left[u_c\left(c^I(\theta, p_I(s)), m^I(\theta); \theta\right) | S \leq s\right] - E\left[u_c\left(c^U(\theta, p_U(s)), m^U(\theta); \theta\right) | S \geq s\right]}{E[u_c]} \tag{16}$$

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 traditional 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 $(1-s)(C(s) - D(s) - sD'(s))$. This term reduces to the transfer in equation (2) when demand equals marginal cost, $D(s) = C(s)$, as in the stylized example in Section 2. 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 $\beta(s)$.

---

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

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

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

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 $EA(s) < 0$.

Relatedly, it is important to note that the model remains valid if the underlying heterogeneity generates advantageous as opposed to adverse selection. This could occur if S reflects a preference shock or liquidity shock such that those who prefer not to purchase insurance are also the latently higher risk population. In this case, the difference in marginal utilities between insured and uninsured remains the key sufficient statistic required for measuring ex-ante expected utility.

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 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. For non-budget neutral settings, one can 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 comparisons of MVPFs across policies characterize the cost to the government of moving welfare across different sets of policy beneficiaries. Moreover, when two policies have the same distributional incidence, they can provide a Pareto-based guidance on optimal policy: if the MVPF of lowering health insurance prices is higher than the MPVF of a tax cut to an identical population, then everyone's 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 for the case when $p_U(s) = 0$.

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

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

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 $1+\frac{C(s)-D(s)}{s(-D'(s))}$.[22]

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 by 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. (2019). 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.

3.3 Insurance versus Redistribution

All of the analysis above can be conducted conditional on any observables, X, that are observed to the econometrician. It is straightforward to show that the exact same formulas in Propositions 1 and 2 continue to apply with two modifications. First, one needs to estimate the demand and cost curves separately for each value of X. Second, one needs to use the difference in marginal utilities between insured and uninsured conditional on X, as opposed to the unconditional difference in marginal utilities. Given these conditional quantities, the resulting formulas in equation (14) and (17) capture the impact of insurance policies conditional on a particular observable characteristic, X.[23] This can then be aggregated across values of X using any desired social welfare weights of the researcher or policymaker. In this sense, this approach need not incorporate any value from redistribution across different types of people as long as these types are observed to the econometrician. Instead, the approach developed here allows the econometrician to measure the welfare impact of insurance market policies using expected utility conditional on a fixed information set that does not vary depending on the information set available to individuals at the time of making insurance choices.

[22] More 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.

[23] In the limiting case, one could conceptually imagine conditioning on the information known to individuals at the time they purchase insurance, $X=\tilde{s}$. In this case, the difference in marginal utility between those who do versus do not purchase insurance is zero (for a particular value of $\tilde{s}$, either everyone purchases insurance or no one purchases insurance). Hence, $EA\,(s)=0$, and market surplus measures the marginal willingness to pay for a larger insurance market.

4 Implementation Using Market Demand and Cost Curves

To measure ex-ante welfare, one requires an estimate of the difference in marginal utilities between insured and uninsured, $\beta(s)$. 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 establish two results. First, I show that if the utility function satisfies a couple of state-independence assumptions, then one can write:

$$\beta(s) = \gamma \Delta c$$

where $\Delta c = E[c(\theta)|S \geq s] - E[c(\theta)|S < s]$ is the difference in consumption between the uninsured and insured and γ is the coefficient of absolute risk aversion. This is analogous to the methods used in the literature on optimal unemployment insurance (Baily (1978); Chetty (2006)).

However, in practice consumption is rarely observed. Therefore, I provide additional conditions[24] under which one can further write $\beta(s)$ as

$$\beta(s) = \gamma [D(s) - E[D(S)|S \geq s]] \tag{18}$$

where $D(s) - E_s[D(S)|S \geq s]$ is the difference in willingness to pay between the marginal type, s, and the average uninsured type, $S \geq s$. In this sense, Proposition 3 provides a method of estimating ex-ante willingness to pay using the willingness to pay and cost curves, combined with a measure of risk aversion. 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.

Section 4.2 formalizes the necessary assumptions required for equation (18) to hold. Section 4.2 then shows how one can potentially relax each of the required assumptions if one observes additional data elements.

[24] Although these additional conditions are restrictive, I argue they are plausible in the application considered in Section 5.

4.1 Implementation Details

I begin by using a Taylor expansion of the marginal utility function. Let $\bar{y} = E[y(\theta)]$ denote the average income of the population. Let $\bar{c}$ denote average consumption.[25] 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. Let $\bar{u}_c = u_c\left(\bar{c}, \bar{m}, \bar{\theta}\right)$ denote the marginal utility of consumption of the average type.

The marginal utility of income for an uninsured type θ is given by[26]

$$\begin{aligned}&u_c\left(c^U\left(\theta, p_U(s)\right), m^U(\theta); \theta\right) - u_c\left(\bar{c}, \bar{m}, \bar{\theta}\right) \\ &\approx \underbrace{u_{cc}\left(c^U\left(\theta, p_U(s)\right) - \bar{c}\right)}_{\text{Consumption}} + \underbrace{u_{cm}\left(m^U(\theta) - \bar{m}\right)}_{\text{Medical Spending}} + \underbrace{u_{c\theta}\left(\theta - \bar{\theta}\right)}_{\text{Preferences}}\end{aligned} \tag{19}$$

And the marginal utility of income for the types, θ, that choose to be insured is given by

$$\begin{aligned}&u_c\left(c^I\left(\theta, p_I(s)\right), m^I(\theta); \theta\right) - u_c\left(\bar{c}, \bar{m}, \bar{\theta}\right) \\ &\approx \underbrace{u_{cc}\left(c^I\left(\theta, p_I(s)\right) - \bar{c}\right)}_{\text{Consumption}} + \underbrace{u_{cm}\left(m^I(\theta) - \bar{m}\right)}_{\text{Medical Spending}} + \underbrace{u_{c\theta}\left(\theta - \bar{\theta}\right)}_{\text{Preferences}}\end{aligned} \tag{20}$$

Combining equations (19) and (20), one can write the percentage difference in marginal utilities for the insured and uninsured when a fraction s purchases insurance as

$$\begin{aligned}\beta(s) \approx\ & \frac{\bar{u}_{cc}}{\bar{u}_c}\left(E\left[c^I\left(\theta, p_I(s)\right) | S \leq s\right] - E\left[c^U\left(\theta, p_U(s)\right) | S > s\right]\right) \\ & + \frac{\bar{u}_{cm}}{\bar{u}_c}\left(E\left[m^I(\theta) | S \leq s\right] - E\left[m^U(\theta) | S > s\right]\right) \\ & + \frac{\bar{u}_{c\theta}}{\bar{u}_c}\left(E[\theta | S \leq s] - E[\theta | S > s]\right)\end{aligned} \tag{21}$$

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

[25]Formally,

$$\bar{c} = \int_\theta \left(\left[\int_{\tilde{s} \leq s} c^I\left(\theta, p_I(s)\right) d\tilde{s}\right] + \left[\int_{\tilde{s} > s} c^U\left(\theta, p_U(s)\right) d\tilde{s}\right]\right) d\theta$$

so that it equals a population-weighted average consumption of the insured and uninsured.

[26]The 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 $\beta(s^*) = \frac{1}{T}\gamma E\left[D(s^*) - D(s) | s \geq s^*\right]$. Intuitively, only $1/T$ of the willingness to pay actually comes from consumption in any given period.

versus uninsured.

Equation (21) shows how state independence assumptions can help write $\beta(s)$ using only differences in consumption. 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 preference heterogeneity in the marginal utility function does not affect the marginal utility of consumption.

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 $u_{c\theta}\left(\theta - \bar{\theta}\right) = 0$ so that there is no covariance between types and the marginal utility of consumption.

Under Assumptions 1-2, it is straightforward to see that the latter two terms in equation (21) equal zero so that the difference in marginal utilities depends only on the difference in consumption between the insured and the insured, multiplied by the curvature of the utility function. But in practice consumption is rarely observed. To that aim, I show that under additional assumptions one can use the demand curve to proxy for the difference in consumption levels.

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

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 is natural in the application to health insurance subsidies in Massachusetts in Section 5 because the demand and cost curves will be estimated for individuals at a fixed income level (150% of the Federal Poverty Line). However, future work need not impose Assumption 3 if one can observe consumption differences between the insured and uninsured in the sample.

Combining these Assumptions yields the main result.

Proposition 3. *Suppose Assumptions 1-2 hold. Then,*

$$\beta(s) \approx \gamma \Delta c \tag{22}$$

where the $\approx$ *denotes a first-order Taylor approximation,* $\Delta c = E\left[c^U(\theta, p_U(s)) | S > s\right] - E\left[c^I(\theta, p_I(s)) | S \leq s\right]$ *is the difference in consumption of the uninsured and insured, 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,* $\bar{\theta}$*.*

Moreover, suppose Assumption 3 holds and that the insurance product provides full insurance, $x^I(m; \theta) = 0$*. Then,*

$$\beta(s) \;\; \approx \;\; \gamma\left(D(s) - E\left[D(S) | S \geq s\right]\right) \tag{23}$$

where $D(s) - E[D(S) | S \geq s]$ *is the difference between the willingness to pay of the marginal type, s, and the average uninsured type. Combining with Proposition 1, the ex-ante component of willingness to pay is given by*

$$EA(s) \approx (1-s)\left(C(s) - D(s) - sD'(s)\right)\gamma\left(D(s) - E\left[D(S) | S \geq s\right]\right) \tag{24}$$

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

Proof. Imposing Assumptions 1-2 to equation (21) yields the result in equation (22). Appendix (G) provides the derivation of equation (23). □

When $C(s) = D(s)$ as in the stylized example, equation (24) reduces to equation (4). In this more general setup, equation (23) 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 do rely on assumptions that should be made clear. Here, I discuss the potential 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:

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

where $\frac{\bar{u}_{cm}}{\bar{u}_c} = \frac{u_{cm}\left(\bar{c},\bar{m};\bar{\theta}\right)}{u_c\left(\bar{c},\bar{m};\bar{\theta}\right)}$ measures how the marginal utility of consumption varies with the level of medical spending (holding c and θ constant) and $E\left[m(\theta)\,|S \geq s\right] - E\left[m(\theta)\,|S < s\right]$ 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,

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

The term $\frac{\bar{u}_{c\theta}}{\bar{u}_c}\left(E\left[\theta|S \geq s\right] - E\left[\theta|S < s\right]\right)$ 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. If individuals have heterogeneous risk aversion, it would be natural that the more risk averse are those who purchase insurance. But, it is important to note this does not mean Assumption 2 does not hold. This is because those with a higher amount of risk aversion (i.e. greater curvature of utility) could also have a lower average marginal utility of income than those with a lower degree of risk aversion (i.e. lower slope of utility). For example, suppose that individuals have CRRA preferences of the form $u(c) = k^{\theta}c^{1-\theta}$. This utility function exhibits the same willingness to pay for insurance for a type θ but will have $u_c = k^{\theta}(1-\theta)\,c^{-\theta}$. For small k this is decreasing in θ; but for large k this is increasing in θ. As a result, the fact that individuals with more risk aversion purchase insurance does not immediately introduce a source of bias in the benchmark implementation in equation (24).

Assumption 3 Heterogeneous income or liquidity shocks, $y(\theta)$, could be a driver of insurance demand. If incomes differ between the insured and uninsured, then one can estimate a modified formula for $\beta(s)$ in the case when $x(m;\theta) = 0$ as

$$\beta(s) = \gamma\left(D(s) - E\left[D(S)\,|S \geq s\right] + E\left[y(\theta)\,|S \leq s\right] - E\left[y(\theta)\,|S > s\right]\right)$$

where $E\left[y(\theta)\,|S \leq s\right] - E\left[y(\theta)\,|S > s\right]$ 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 (23) 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 $D(s) - E\left[D(S)\,|S \geq s\right]$.

A key advantage of the empirical implementation in the next Section is that the demand and cost estimates will be constructed for a population conditional on their income level (in particular, those near 150% of the federal poverty line). But more generally, if income is likely to differ in other contexts between the insured and uninsured, it would suggest a value to directly observing the consumption of the insured and uninsured. One could then multiply this difference by the coefficient of absolute risk aversion to estimate of $\beta(s)$.

5 Application: Health Insurance Subsidies for Low-Income Adults

Should there be a mandate for health insurance? Finkelstein et al. (2019) estimate demand and cost curves for health insurance for low-income adults in Massachusetts. Their results show that a private market fully unravels as a result of adverse selection and uncompensated care externalities.[27] What is the welfare cost of this unraveling? Should the government intervene and impose a mandate or fully subsidize insurance? Can market surplus and expected utility lead to different conclusions about the optimal government intervention? In this section, I use the estimates from Finkelstein et al. (2019) to illustrate the methods

[27] In addition to the application provided here, Appendix I also illustrates how to apply the approach to analyze the welfare impact of different prices for more versus less generous insurance at a large employer using estimates from Einav et al. (2010).

developed in this paper.[28] From an ex-ante perspective, the optimal insurance prices are roughly 30% lower than than the prices that maximize market surplus; mandates would increase expected utility despite reducing market surplus.

Background Using administrative data from Massachusetts' subsidized insurance exchange, Commonwealth Care, Finkelstein et al. (2019) exploit discontinuities in the subsidy schedule to estimate willingness to pay and costs of insurance among low-income adults between 133% and 350% of the Federal Poverty Line (FPL). As subsidies decline and prices rise, insurance take-up falls. And, as prices rise, the average cost of the insureds increases as well, indicating a presence of adverse selection. Figure 4 (Panel A) depicts the resulting willingness to pay and cost curves for those with incomes at 150% FPL.[29] Relative to the numbers presented in Finkelstein et al. (2019), these curves are scaled 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 also presents the cost of these marginal enrollees that is paid by the insurer, $C^{gross}(s)$. I use the term "gross" costs because part of these costs are not net resource costs, due to the presence of uncompensated care discussed further below. The figure presents clear evidence of adverse selection: the marginal enrollees tend to be lower-cost (i.e. $C^{gross}(s)$ slopes downward). Finkelstein et al. (2019) also find that enrollee willingness to pay is far below individuals' own expected costs paid by the insurer, $D(s) < C^{gross}(s)$. 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 adjust for the presence of uncompensated care, 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 $C(s)$.

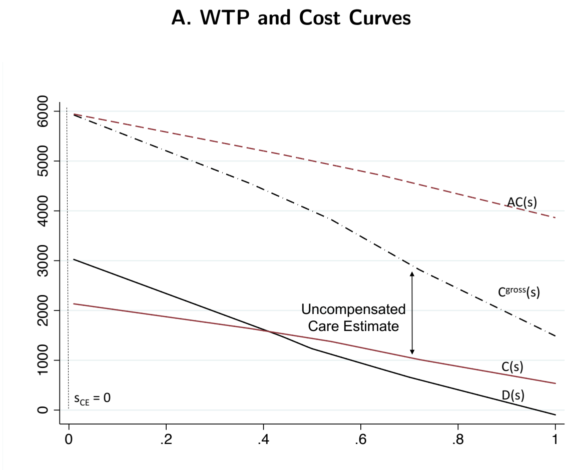
Market Surplus Panel B of Figure 4 conducts a standard welfare analysis that focuses on market surplus, comparing observed willingness to pay to costs, $D(s) - C(s)$. Market surplus is maximized when the demand curve intersects the cost curve, which occurs when 41% of the

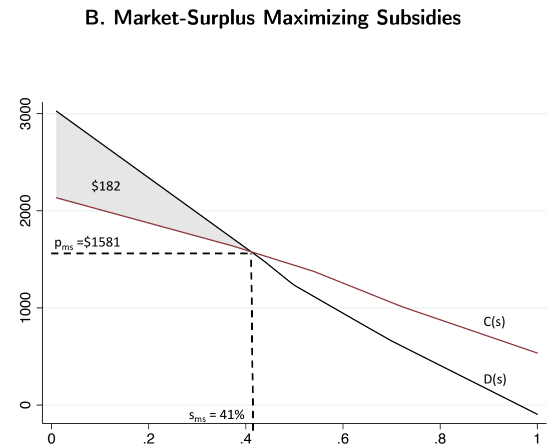
[28]In addition to the application provided here, Appendix I also illustrates how to apply the approach to analyze the welfare impact of different prices for more versus less generous insurance at a large employer using estimates from Einav et al. (2010).

[29]150% FPL corresponds to roughly $16K in income for an individual with no children.

market owns insurance. The marginal price that leads to this allocation is \$1581.[30] 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.

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





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 $\gamma = 5x10^{-4}$ (e.g. similar to estimates in Handel et al. (2015)).[31] Appendix H discussed how the estimates change with alternative risk aversion assumptions.

Figure 5 presents the ex-ante demand curve from Proposition 1 using $\gamma = 5x10^{-4}$ and equation (24). 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 $D'(0.5) = -3405$.[32] The average willingness to pay with those whose demand is below $D(0.5) = 1232$ is 559. Combining

[30] This 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.

[31] As discussed in Appendix H, Handel 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 estimate is likely a lower bound on the size γ.

[32] Finkelstein et al. (2019) estimate a piece-wise linear demand cure. To obtain smooth estimates of the slope of demand, I regress the estimates of $D(s)$ from Finkelstein et al. (2019) on a 10th order polynomial in s. The results are similar for other smoothed functions.

using equation (24), the ex-ante willingness to pay for a larger insurance market is

$$\begin{aligned} EA(s) &= (1-s)\left(C(s) - D(s) - sD'(s)\right)\gamma 2\left(D(s) - E\left[D(S)\,|S \geq s\right]\right) \\ &= .5\left(1438 - 1232 + 0.5 * 3405\right)\left(5x10^{-4}\right)(1232 - 559) \\ &= 321 \end{aligned}$$

The median individual is willing to pay \$1,232 for insurance at the time one observes them in the market. But, prior to learning their willingness to pay for insurance, individuals are willing to pay to have lower insurance prices. 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: Ex-Ante Welfare of Health Insurance for Low-Income Adults

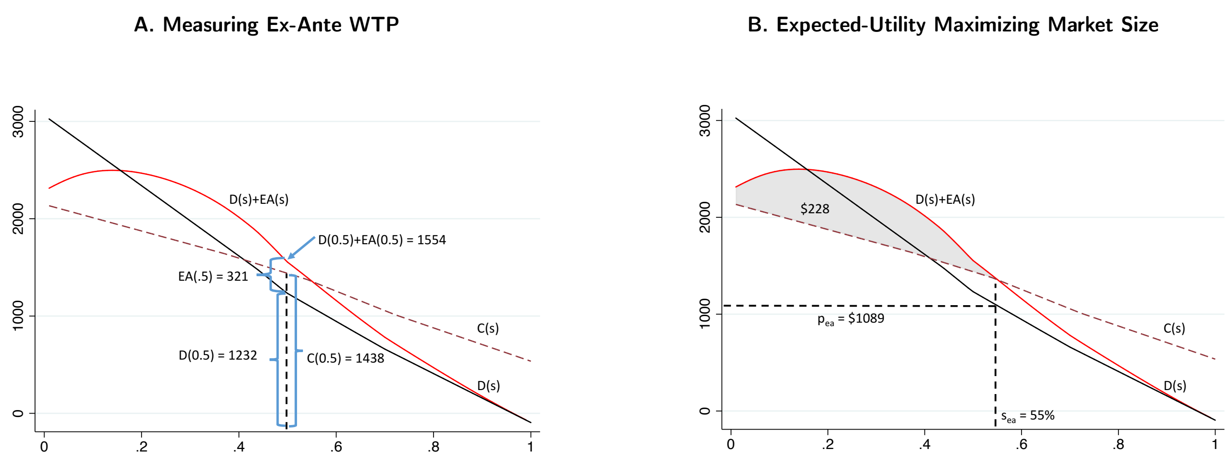


Figure 5 presents the Ex-ante WTP curve, $D^{Ex-Ante}(s)$, for all values of s.[33] Expected utility is maximized when $W'(s) = 0$, or $D(s) + EA(s) = C(s)$. This occurs when 55% of the market owns insurance and the marginal price of insurance is \$1,089. This contrasts with

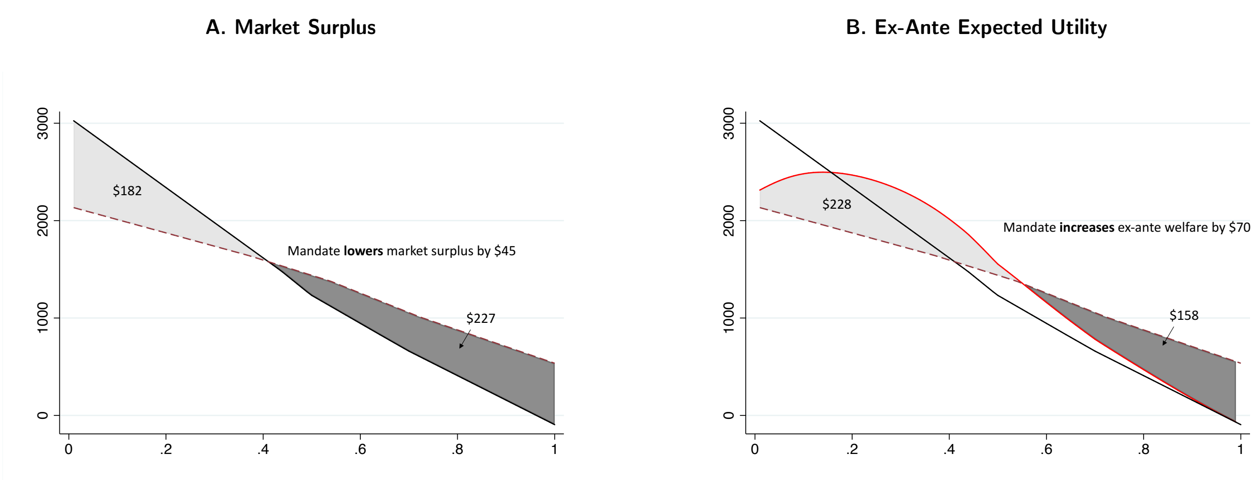
[33]It is perhaps surprising that $EA(s) < 0$ for low values of s. Mathematically, this is because for low values of s, $C(s) - D(s) < sD'(s)$. Economically, this means that expanding the size of the insurance market actually generates a Pareto improvement, as it can lower prices for both the insured 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.

the market surplus-maximizing size of the market of 41% and the optimal price is roughly 30% lower than the 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 different conclusions about optimal insurance prices and the welfare cost of adverse selection.

Mandates In addition to the optimal insurance prices, one can also ask whether the government should simply impose a mandate ($s = 1$), versus allow the competitive market to fully unravel ($s = 0$). 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 surplus 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.

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



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 market surplus.

Appendix H shows that this conclusion that mandates increase ex-ante expected utility holds as long as the coefficient of absolute risk aversion exceeds roughly $1.8x10^{-4}$. This is generally far below the estimates of risk aversion obtained in health insurance settings. For example, $1.8x10^{-4}$ is roughly 4 standard deviations below the mean in the distribution of risk aversion estimates in Handel et al. (2015). In this sense, an ex-ante welfare perspective can lead to different conclusions about the desirability of government mandates.

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 subsidies. Recall, the MVPF equals the marginal willingness to pay of the beneficiaries for insurance subsidies per dollar of government expenditure. This is given by the formula:

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

where $\beta\left(s\right)=\gamma\left(D\left(s\right)-E\left[D\left(S\right)|S\geq s\right]\right)$ is the difference in marginal utilities between the insured and uninsured.

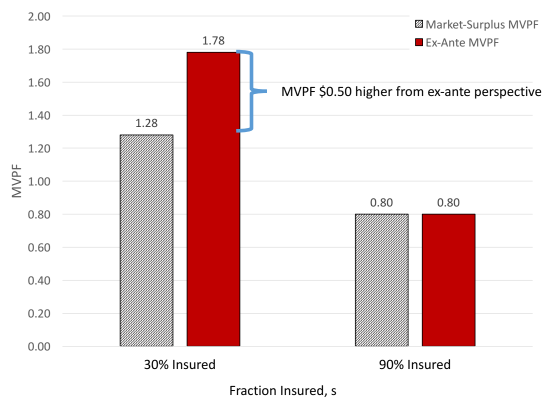
Figure 7 calculates this MVPF for two values of s that corresponds to the range of take-up estimates in Finkelstein et al. (2019). The price variation in Finkelstein et al. (2019) leads to between 30-90% of the market choosing to purchase insurance, and therefore I construct the MVPF of more generous insurance subsidies at these endpoints of this observed variation.

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

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

From a market surplus perspective, the willingness to pay for additional subsidies 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. Accounting for this, the results suggest that the marginal willingness to pay of the beneficiaries for insurance subsidies per dollar of government expenditure is actually 1.78, not 1.28.

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



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 $D(s^*)$ and $E[D(s)|s \geq s^*]$ is smaller when a larger fraction of the market has insurance – in turn, this means that the consumption of the insured relative to the average in the population population is smaller when a larger fraction of the market is insured. As a result, the MVPF is similar (around 0.8) from both perspectives.

6 Conclusion

This paper develops a set of tools to measure the impact of insurance market policies on ex-ante measures of welfare. These measures differ from market surplus because they measure expected utility before individuals learn their willingness to pay for insurance. An ex-ante welfare perspective can lead to different conclusions than a market surplus perspective. Policies that maximize ex-ante expected utility often involve lower insurance prices, a greater value of mandates, and a higher value of insurance subsidies.

Future work could measure the welfare consequences of contract distortions, such as the exclusion of high cost drugs for chronic conditions. It could also expand beyond the binary case considered here to consider multiple contracts. One could also explore the role of income and liquidity shocks. The Massachusetts application in Section 5 considered a market for insurance for those with a particular income level (150% FPL). But in other settings, one might imagine that income and liquidity is a key driver of willingness to pay for insurance. For example, it could be that income shocks lead individuals to forego insurance under the Affordable Care Act. In this case, the insured may have a lower marginal utility of income than the uninsured ($\beta(s) < 0$). If true, imposing a mandate may not increase ex-ante expected utility even if it increases market surplus. As noted in Section 4, this is readily tested with data on consumption of insured and uninsured.

The intuition and approach likely extend to settings where prices are not observed, such as approaches to valuing social insurance. Many approaches use labor supply responses to value social insurance programs (e.g. Keane and Moffitt (1998); Gallen (2014); Dague (2014)). Such approaches capture the value of insurance against only the risk that remains after choosing labor supply. Other approaches use changes in consumption around a shock to infer willingness to pay (e.g. Gruber (1997); Meyer and Mok (2013)). But consumption should change when information about the event is revealed, not when the event occurs. And the approaches here could be extended to measure ex-ante expected utility in such settings.[34]

Many macroeconomic welfare measures face similar conceptual issues. This includes the famous calculations of the welfare cost of business cycles in Lucas (2003). When consumption responds to information over time, the variance of consumption changes may under-state ex-ante welfare. Using consumption data, one could extend the tools in this paper to measure the ex-ante welfare cost of business cycles.

References

Baily, M. N. (1978). Some aspects of optimal unemployment insurance. *Journal of Public Economics 10(3)*, 379–402. 1, 2, 4

Cabral, M. (2017). Claim timing and ex post adverse selection. *Review of Economic Studies 84*(1). 1

[34]The approach could also be generalized to allow for behavioral bias in insurance choices, as in Spinnewijn (2017).

Chetty, R. (2006). A general formula for the optimal level of social insurance. *Journal of Public Economics 90*(10), 1879–1901. 1, 2, 4

Dague, L. (2014). The effect of Medicaid premiums on enrollment: A regression discontinuity approach. *Journal of Health Economics 37*, 1–12. 1, 6

Ehrlich, I. and G. S. Becker (1972). Market insurance, self-insurance, and self-protection. *Journal of Political Economy 80(4)*, 623–648. 17

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, 4, 2, 2, 3, 17, 3, 3, 3.1, 27, 28, B, G, I, 37, I

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, 2

Feldman, R. D. and B. E. Dowd (1982). Simulation of a health insurance market with adverse selection. *Operations Research 30*(6), 1027–1042. 1

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

Finkelstein, A., N. Hendren, and M. Shepard (2019). Subsidizing health insurance for low-income adults: Evidence from massachusetts. *American Economic Review 109*(4), 1530–67. 1, 2, 3, 3.2, 5, 5, 32, 5, D, I

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

Finkelstein, A., K. McGarry, and A. Sufi (2005). Dynamic inefficiencies in insurance markets: Evidence from long-term care insurance. *American Economic Review 95*(2), 224–228. 1

Gallen, T. S. (2014). Measuring the value of Medicaid using TennCare. *Working Paper*. 1, 5, 6

Gruber, J. (1997). The consumption smoothing benefits of unemployment insurance. *American Economic Review 87(1)*, 192–205. 1, 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, 2, 15, 4.1, 5, 31, 5, H

Harsanyi, J. C. (1978). Bayesian decision theory and utilitarian ethics. *The American Economic Review 68*(2), 223–228. 1

Hendren, N. (2013). Private information and insurance rejections. *Econometrica 81*(5), 1713–1762. 1

Hendren, N. (2014). The inequality deflator: Interpersonal comparisons without a social welfare function. *NBER Working Paper, http://www.nber.org/papers/w20351*. 3.2

Hendren, N. (2016). The policy elasticity. *Tax Policy and the Economy 30*. 1, 8, 3.2

Hendren, N. (2017). Knowledge of future job loss and implications for unemployment insurance. *American Economic Review 107*(7). 1

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

Keane, M. and R. Moffitt (1998). A structural model of multiple welfare program participation and labor supply. *International economic review*, 553–589. 1, 6

Lucas, R. E. (2003). Macroeconomic priorities. *American economic review 93*(1), 1–14. 6

Meyer, B. D. and W. K. Mok (2013). Disability, earnings, income and consumption. Technical report, National Bureau of Economic Research. 1, 6

Panhans, M. (2018). Adverse selection in aca exchange markets: Evidence from colorado. *The American Economic Journal: Applied Economics*. 1

Saez, E. and S. Stantcheva (2016). Generalized social marginal welfare weights for optimal tax theory. *American Economic Review 106*(1), 24–45. 1

Spinnewijn, J. (2017). Heterogeneity, demand for insurance, and adverse selection. *American Economic Journal: Economic Policy 9*(1), 308–43. 34

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

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

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

$$dp_U = -\frac{s}{1-s}dp_I$$

Now, consider the demand identity:

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

differentiating and re-arranging yields:

$$\begin{aligned} dp_I - dp_U &= D'(s)\,ds \\ dp_I &= D'(s)\,ds + dp_U \\ dp_I &= D'(s)\,ds - \frac{s}{1-s}dp_I \\ \frac{1}{1-s}dp_I &= D'(s)\,ds \\ dp_I &= (1-s)\,D'(s)\,ds \end{aligned}$$

and

$$\begin{aligned} dp_U &= dp_I - D'(s)\,ds \\ dp_U &= -sD'(s)\,ds \end{aligned}$$

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 imported from another setting as in Section I. 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 by 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 the it is straightforward to show[35] that the coefficient of absolute risk aversion is given by:

$$\gamma(\tilde{s}) \quad = \quad 2\frac{D(\tilde{s}) - C(\tilde{s}) + (1 - p(\tilde{s}))\, E\left[m^I - m^U | \tilde{s}\right]}{V} \tag{25}$$

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

[35] To see this, suppress notation w.r.t. θ and condition all expectations on $\tilde{s}$. Let $\left(\bar{c}, \bar{m}, \bar{\theta}\right)$ 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 $\tilde{s}$.

offered by the insurance:

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

and $(1 - p(\tilde{s}))\, E\left[m^I - m^U|\tilde{s}\right]$ is a correction term to account for moral hazard. $E\left[m^I(\theta) - m^U(\theta)|\tilde{s}\right]$ 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 $\tilde{s}$, $E\left[m^I(\theta) - m^U(\theta)|\tilde{s}\right]$ 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 states Proposition 1 in 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

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

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

$$\frac{V'(s^*)}{E\left[u_c\left(y(\theta) - p_I(s), m^I(\theta);\theta\right)|s \leq s^*\right]} \approx \underbrace{p_I(s^*) - p_U(s^*) + EA(s^*)}_{\text{Ex-Ante Demand}} - MC(s^*) \tag{26}$$

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

$$EA(s^*) = \underbrace{\frac{1 - s^*}{1 + s^*\left(\frac{\partial D}{\partial p_U} - 1\right)}\left(MC(s^*) - (p_I(s^*) - p_U(s^*)) - s^*\frac{\partial D}{\partial s}\right)\beta(s^*)}_{\text{Transfer from Uninsured to Insured}} \tag{27}$$

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

$$\beta(s) = \frac{E\left[u_c\left(y(\theta) - p_I(s), m^I(\theta); \theta\right) | \tilde{s} \leq s\right] - E\left[\frac{\partial D(\tilde{s}, p_U(s))}{\partial p_U} u_c\left(y(\theta) - D(\tilde{s}, p_U(s)), m^I(\theta); \theta\right) | \tilde{s} \geq s\right]}{E\left[u_c\left(y(\theta) - p_I(s), m^I(\theta); \theta\right) | \tilde{s} \leq s\right]} \tag{28}$$

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

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

so the marginal welfare impact is given by

$$\begin{aligned} W'(s) &= -sp_I'(s) E\left[u'\left(y(\theta) - p_I(s) - x^I\left(m^I(\theta); \theta\right), m^I(\theta); \theta\right) | \tilde{s} \leq s\right] \\ &\quad - (1-s) p_U'(s) E\left[u_c\left(y(\theta) - m^U(\theta) - p_U(s), m^U(\theta); \theta\right) | \tilde{s} \geq s\right] \end{aligned}$$

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. (2019)). 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 $(1-s)\,p_U'(s)\,E\left[u_c\left(y(\theta) - p_I(s), m^I(\theta);\theta\right)|\tilde{s} \leq s\right]$ and then dividing by $E[u_c]$ yields

$$\begin{aligned}&\frac{W'(s)}{E[u_c]}\\ &= -\left[sp_I'(s) + (1-s)\,p_U'(s)\right] + (1-s)\,p_U'(s)\\ &\times\left(\frac{E\left[u_c\left(y(\theta) - p_I(s) - x^I\left(m^I(\theta);\theta\right), m^I(\theta);\theta\right)|S \leq s\right] - E\left[u_c\left(y(\theta) - m^U(\theta), m^U(\theta);\theta\right)|S \geq s\right]}{E[u_c]}\right)\end{aligned}$$

or

$$\begin{aligned}&\frac{W'(s)}{E[u_c]}\\ &= -MDWL + (1-s)\,p_U'(s)\,\beta(s)\end{aligned}$$

where

$$\beta(s) = \left(\frac{E\left[u_c\left(y(\theta) - p_I(s) - x^I\left(m^I(\theta);\theta\right), m^I(\theta);\theta\right)|S \leq s\right] - E\left[u_c\left(y(\theta) - m^U(\theta), m^U(\theta);\theta\right)|S \geq s\right]}{E[u_c]}\right)$$

Now, 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

$$sp_I(s) + (1-s)\,p_U(s) = sAC(s) \tag{29}$$

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 (29) 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

$$\frac{W'(s)}{E[u_c]} = D(s) - C(s) + (1-s)\left(D(s) - C(s) + s\frac{\partial D}{\partial s}\right)\beta(s)$$

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 Insurance Versus Redistribution: Conditioning on $X = x$

The approach provided here can also be amended to facilitate welfare analysis after some observable information, X, has been revealed about θ. For example, perhaps one does not wish to incorporate the value of insurance to the extent to which it redistributes across those with different incomes or health conditions.

This appendix shows how one can make adjustments to the baseline formula for $EA(s)$ by conditioning on the observable characteristics, $X = x$. To see how this can work, suppose prices, p_U and p_I , are charged uniformly to people with different values of X and that a fraction s of the market purchases insurance.[36] Let s_x denote the fraction of the population with characteristics $X = x$ that are uninsured. (note that $s = E_X[s_x]$ is the total fraction of the market insured). Next, let $\beta(s, x)$ denote the difference in marginal utilities between the insured and uninsured given by a generalized version of equation (16):

$$\beta(s,x) = \frac{E\left[u_c\left(y(\theta) - p_I(s), m^I(\theta); \theta\right) | \tilde{s} \le s_x, X = x\right] - E\left[u_c\left(y(\theta) - D(\tilde{s}) - p_U(s), m^I(\theta); \theta\right) | \tilde{s} \ge s_x, X = x\right]}{E\left[u_c\left(y(\theta) - p_I(s), m^I(\theta); \theta\right) | \tilde{s} \le s_x, X = x\right]}$$

Now, note that the aggregate impact on p_U of expanding the size of the insurance market is determined by the aggregate resource constraint, and hence we continue to have $p'_U(s) = C(s) - D(s) - s\frac{\partial D}{\partial s}$, where $\frac{\partial D}{\partial s}$ is the slope of the aggregate demand curve (across all X). Combining, the ex-ante welfare value of expanding the insurance market for those with characteristics $X = x$ is given by

$$EA(s,x) = (1 - s_x)\left(C(s) - D(s) - s\frac{\partial D}{\partial s}\right)\beta(s_x, x) \tag{30}$$

and aggregating across all values of X using equal weights on those with different X characteristics yields an ex-ante welfare value of $E_X[EA(s, X)]$. This approach aggregates welfare from behind a set of "veils of ignorance" – one for each value of X. In the limiting case where X incorporates all information about s, then there is no difference in marginal utilities across s conditional on X, $\beta(s_x, x) = 0$. Hence, there would be no additional ex-ante value to the insurance ($EA(s, x) = 0$). This is simply another way of saying that market surplus treats all sources of differences in demand as redistribution as opposed to having potential insurance value.

[36] If prices, p_U and p_I, are charged differentially to those with different X characteristics, then one can simply conduct welfare analysis by conditioning on X everywhere in Proposition 1.

Analogous derivations show that the MVPF conditional on $X = x$ generalizes to

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

G Proof of Proposition 3

To begin, consider the Taylor expansion in equation (21),

$$\begin{aligned}\beta(s) \approx\; & \frac{\bar{u}_{cc}}{\bar{u}_c}\left(E\left[c^I(\theta, p_I(s))\,|S \le s\right] - E\left[c^U(\theta, p_U(s))\,|S > s\right]\right) \\ & + \frac{\bar{u}_{cm}}{\bar{u}_c}\left(E\left[m^I(\theta)\,|S \le s\right] - E\left[m^U(\theta)\,|S > s\right]\right) \\ & + \frac{\bar{u}_{c\theta}}{\bar{u}_c}\left(E\left[\theta|S \le s\right] - E\left[\theta|S > s\right]\right)\end{aligned}$$

Now, note that Assumptions 1 and 2 imply that the latter two terms equal zero. This implies $\beta(s) = \gamma\Delta c$. To show that the demand curve can be used to proxy for the difference in marginal utility of consumption, we follow a trick provided in Einav et al. (2010) where we substitute the demand function into the utility function to note that utility of the uninsured with signal $S = \tilde{s}$ can be written as:

$$E\left[u\left(c^U(\theta, p_U(s)), m^U(\theta); \theta\right)|S = \tilde{s}\right] = E\left[u\left(c^I(\theta, D(\tilde{s}) + p_U(s)), m^I(\theta); \theta\right)|S = \tilde{s}\right]$$

so that the derivative with respect to s is given by

$$-p_U'(s)\,E\left[u_c\left(c^U(\theta, p_U(s)), m^U(\theta); \theta\right)|S = \tilde{s}\right] = -\left(p_U'(s)\right)E\left[u_c\left(c^I(\theta, D(\tilde{s}) + p_U(s)), m^I(\theta); \theta\right)|S = \tilde{s}\right]$$

Hence, one can replace the marginal utility of consumption for the uninsured with the marginal utility of consumption for the uninsured that they would have if they chose to be insured,

$$\beta(s) = \frac{E\left[u_c\left(c^I(\theta, p_I(\theta)), m^I(\theta); \theta\right)|S \le s\right] - E\left[u_c\left(c^I(\theta, D(S) + p_U(s)), m^I(\theta); \theta\right)|S > s\right]}{E\left[u_c\right]}$$

Now, replacing the Taylor expansion terms and imposing $u_{cm} = u_{c\theta} = 0$, we have

$$\beta(s) = \frac{\bar{u}_{cc}}{\bar{u}_c}\left(E\left[c^I(\theta, p_I(\theta))\,|S \le s\right] - E\left[c^I(\theta, D(S) + p_U(s))\,|S > s\right]\right)$$

Next, note that in the case where $x^I(m;\theta) = 0$, consumption of the insured is given by $c^I(\theta, p_I(s)) = y(\theta) - p_I(s)$ so that

$$E\left[c^I(\theta, p_I(\theta))\,|S \le s\right] - E\left[c^I(\theta, D(S) + p_U(s))\,|S > s\right] = -p_I(s) + p_U(s) + E\left[D(S)\,|S > s\right]$$

so that

$$\beta(s) = \left(-\frac{\bar{u}_{cc}}{\bar{u}_c}\right)(D(s) - E[D(S)|S > s])$$

which concludes the proof.

H Alternative Risk Aversion Estimates

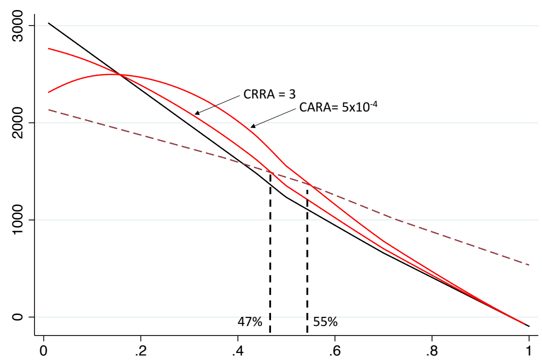
To what extent is the conclusion that mandates increase ex-ante expected utility robust to alternative plausible risk aversion parameters? To begin, Figure A1 Presents estimates of the ex-ante willingness to pay curve under the assumption that the coefficient of relative risk aversion is 3. I translate the coefficient of relative risk aversion of 3 into a coefficient of absolute risk aversion by dividing by $16,335. This income level of $16,335 is the income of single adults at 150% FPL, which is the population considered in the Massachusetts example.

Using this coefficient of absolute risk aversion of roughly $1.8x10^{-4}$, I reconstruct the ex-ante demand curve. 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 $s = 0$.

In sum, the conclusion that mandates increase ex-ante welfare continues to hold for this alternative risk aversion calculation. However, a coefficient of relative risk aversion of around 3 is about the point at which the ex-ante welfare impact of the mandate is closest to zero. Mandates increase ex-ante expected utility as long as the coefficient of relative risk aversion is greater than 3, or coefficient of absolute risk aversion exceeds $1.8x10^{-4}$.

In health insurance contexts, most estimates of risk aversion exceed a coefficient of absolute risk aversion of $1.8x10^{-4}$, with estimates often around $5x10^{-4}$. For example, Handel et al. (2015) estimate a mean coefficient of absolute risk aversion $4.39x10^{-4}$, with an SD of $0.66x10^{-4}$. This suggests the estimate of $1.8x10^{-4}$ is roughly 4 standard deviations below the mean estimate. Moreover, their estimate is for a higher income population. Assuming absolute risk aversion is declining in income, this suggests that the low income population in the Massachusetts example would have even higher risk aversion than in Handel et al. (2015). Therefore, the conclusion that mandates increase ex-ante expected utility appears to hold for a wide range of plausible risk aversion parameters. However, the results here characterize when mandates increase ex-ante expected utility. Mandates increase ex-ante expected utility if and only if the coefficient of absolute risk aversion for this population exceeds $1.8x10^{-4}$.

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



I More vs. Less Generous Employer-Provided Health Insurance

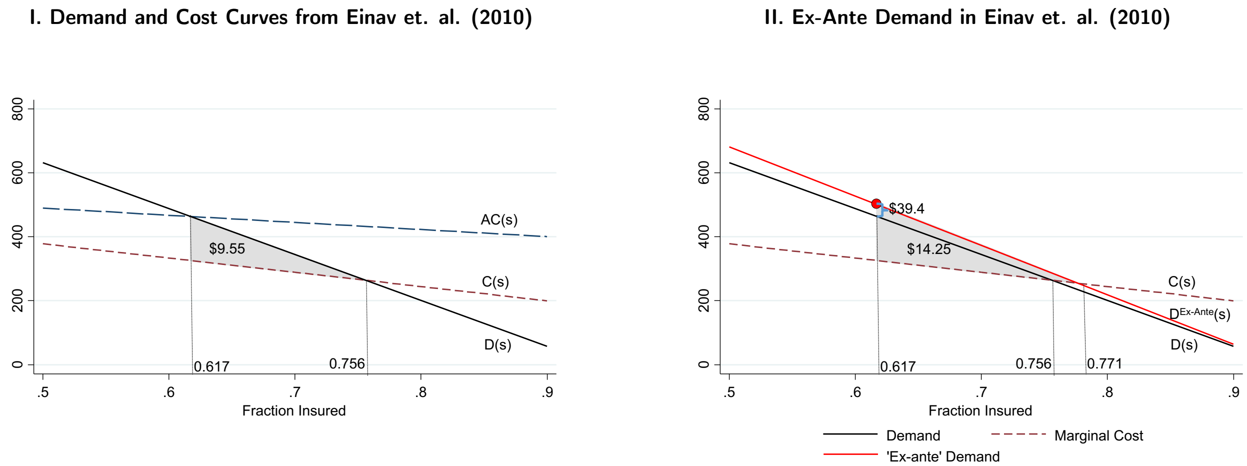
Einav et al. (2010) use variation in prices across business units of Alcoa to estimate demand and cost curves for a more generous health insurance policy relative to a less generous policy. Figure A2, Panel I presents their demand and cost curve estimates. A competitive equilibrium in this environment 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 $D\left(s^{CE}\right) = AC\left(s^{CE}\right) = \$463.5$. But, those that are indifferent to purchasing insurance at a price of \$463.5 on average impose a cost on the insurance company, $C\left(s^{CE}\right)$, 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 I of Figure A2.

How does this compare to an ex-ante measure of the welfare cost of adverse selection? Panel II of Figure A2 presents the estimated $D^{Ex-Ante}\left(s\right)$ assuming $\gamma = 5 \times 10^{-4}$.[37] 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%.

[37]To illustrate its calculation at $s = 0.617$, note that the linearity of demand in Einav et al. (2010) implies $E\left[D\left(s\right) - D\left(S\right)|S \geq s\right] = D'\frac{(1-s)}{2}$. So, equation (24) becomes

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

Figure A2: 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 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 A2, 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? A key distinction in the Einav et al. (2010) example relative to the Finkelstein et al. (2019) example is that the former considers a top-up market

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

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

Plugging in $MC\,(0.617) = \$325.88$, $D\,(0.617) = \$463.5$, and $D' = -1435.97$, along with $\gamma = 5x10^{-4}$, 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, as opposed to a full insurance policy. As a result, this distinction partially reflects a general phenomenon that the ex-ante adjustment tends to be increasing in the size of the insurable loss.

To see why the size of the insurable risk can matter, 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 (24) 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.[38]

[38] Similarly, 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.