Measuring Ex-Ante Welfare in Insurance Markets

Nathaniel Hendren*

April, 2017

Abstract

Observed demand estimates in adversely selected markets tend to understate the ex-ante (or utilitarian) willingness to pay for insurance. This paper derives an ‘ex-ante’ demand curve that measures welfare from behind the veil of ignorance. I provide a strategy to estimate this curve using observed market demand and cost curves, combined with a risk aversion parameter. Using examples motivated by previous health insurance literature, I show that an ex-ante welfare perspective can generate different conclusions about the optimal size of insurance markets, the welfare cost of adverse selection, and the desirability of mandates relative to competitive markets.

1 Introduction

Since Akerlof (1970), it has been recognized that insurance market equilibriums may be inefficient: individuals who are willing to pay the cost they would impose on the insurance company may not obtain insurance. This is because equilibrium prices must reflect the average, not marginal, cost of the insured. Motivated by this, there is a large and growing literature estimating the difference between observed demand and costs to measure the size of this lost market surplus from efficient trades that go unmet in equilibrium. This market surplus measure is then often used to discuss the welfare impact of government policy changes, such as insurance mandates (Einav et al. (2010)).

However, insurance has value from insuring against the realization of risk. Adverse selection occurs when a portion of this risk is already known at the time of contracting. This suggests that the same knowledge of future risk that can generate adverse selection can also render market surplus as an unstable measure of welfare. In particular, the average observed demand for insurance will generally be less than the ex-ante willingness to pay for insurance (Hirshleifer (1971)). Although market surplus measures an interesting question of the value of trades that go unmet in equilibrium, it can be potentially misleading as a tool for evaluating optimal policies such as mandates and subsidies.

To illustrate this, consider the following example. Suppose individuals have $30 dollars but face a risk of losing $m dollars, where m is uniformly distributed between 0 and 10. If individuals have

*Harvard University, nhendren@fas.harvard.edu. I am very grateful to Raj Chetty, David Cutler, Liran Einav, Amy Finkelstein, Mark Shepard, and Mike Whinston for helpful comments and discussions.
a coefficient of relative risk aversion of 3, they would be willing to pay a markup of $0.50 for a full insurance policy: they are indifferent between roughly $24.50 with certainty and a uniformly distributed consumption between $20 and $30.

Figure 1, Panel A, illustrates this scenario using the demand and cost curve framework formalized in Einav et al. (2010). The horizontal axis represents the fraction, $s$, of the market purchasing insurance, and the vertical axis reflects prices and costs in the market. Each individual is willing to pay $5.50 for insurance, generating a flat demand curve of $D(s) = 5.50$. Because no one knows anything about their particular cost, each individual imposes an average cost of $5 on the insurance company, generating a flat cost curve of $AC(s) = 5$. If a competitive market were to open up in this setting, one would expect everyone ($s_{\text{CE}} = 100\%$) to purchase insurance at a price of $5$. Everyone would purchase this insurance policy, and it would generate $0.50$ of welfare, as reflected by the integral between the demand and cost curve.

**Figure 1: Example Demand and Cost Curves**

A. Before Information Revealed

![Demand and Cost Curves Before Information Revealed](image)

B. After Information Revealed

![Demand and Cost Curves After Information Revealed](image)

Now, what happens if some information is revealed at the time that the insurance market transactions occur? For simplicity, consider the extreme case that individuals have fully learned their cost. Demand will equal individuals’ known costs, $D(s) = m(s)$. Individuals 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 $s = 1$. This demand curve lies everywhere on top of the marginal cost curve, $D(s) = MC(s)$, as illustrated in Panel B. If an insurer were to try to sell insurance in this setting, they would need to set prices to cover the average cost of those who purchase insurance. But, the average cost of those choosing to purchase insurance, $AC(s)$, lies everywhere above the demand curve. Hence, the market would fully unravel. The unique competitive equilibrium would involve no one obtaining any insurance, $s_{\text{CE}} = 0\%$.

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 marginal cost curve, there are no valuable foregone trades. This reflects an extreme case of a more general phenomenon: the market demand curve does not embody the value of insurance against the realization of risk prior to the measurement of demand. This idea was arguably first developed in the classic work of Hirshleifer (1971).

The traditional approach to capture the ex-ante value of insurance would be to specify an economic structure including a utility function and information set, and then use observed outcome data (e.g. consumption or proxies for consumption) to infer the ex-ante value of insurance. If one estimated (or assumed) the underlying economic primitives, one could use the model to infer the welfare impact of alternative policies from behind the veil of ignorance. For recent implementations of this approach, see Handel et al. (2015), Section IV of Einav et al. (2016), or Finkelstein et al. (2016).¹

This paper develops a complementary “sufficient statistics” approach to measuring ex-ante welfare that can be applied within the demand and cost curve framework in Figure 1 and does not require observing the underlying outcome distributions (i.e. joint distribution of costs and preferences). To do so, I develop a new “ex-ante” demand curve that captures the ex-ante value of insurance when evaluating policies (e.g. subsidies and mandates) in markets where trades are taking place after some information is revealed. Evaluating policy proposals (e.g. insurance subsidies) by comparing the ex-ante demand curve to the marginal cost curve provides guidance on the welfare impact of these policies from the normative perspective of being behind the veil of ignorance.²

I first develop a general characterization of the ex-ante demand curve and then provide a specific implementation that rests on additional assumptions. Generically, I show that one can theoretically measure the ex-ante demand curve by combining information from the observed demand and cost curves from Einav et al. (2010) with estimates of the difference in marginal utilities between the insured and uninsured. The latter requirement is analogous to the use of marginal utilities in the Baily-Chetty formula to value social insurance (Baily (1978); Chetty (2006)).

Second, to implement this curve one needs estimates of the difference in marginal utilities between the insured and uninsured. Here, I provide a set of conditions under which one can infer this from the observed market demand and cost curves combined with an estimate of risk aversion. This risk aversion estimate can be imported from other settings or inferred within the setting from the willingness to pay for insurance relative to the variance reduction offered by the insurance product.

To illustrate the derivation of the ex-ante demand curve, return to the simple example where

¹This paper is very closely related to Handel et al. (2015) who estimate the value of insurance against reclassification risk. They document that for plausible parameter values, community rating regulations can increase welfare even though they reduce market surplus because they provide additional insurance against reclassification risk.

²The approach can also be conducted conditional on any observable information, X, if one wishes not to incorporate any redistributive value across individuals with different values of X. in this sense, the primary aim of this paper is not simply to argue for incorporating the redistributive value of insurance; rather it is about using a welfare metric that is stable with respect to the amount of information that happens to be revealed at the time demand is measured. As illustrated in Figure 1, deadweight loss does not have this stability.
all information has been revealed so that demand equals costs, $D(s) = m(s)$. 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). Assume that these prices are set subject to an aggregate resource constraint that requires the total amount of money collected to equal the total cost of the insured, $sp_I + (1 - s)p_U = sAC(s)$. One can now ask: What is the value of expanding the size of the insurance market from $s$ to $s + ds$?

To begin, suppose that the government has set prices such that a fraction $s = 0.5$ of the population chooses to purchase insurance, as illustrated in Figure 2, Panel A. Expanding the size of the insurance market 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 $dp_U = -sD'(s)ds$. Conversely, the uninsured prices must decrease by $dp_I = (1 - s)D'(s)ds$. These price changes induce a transfer from the uninsured to the insured, as indicated by the blue arrow in Figure 2, Panel B.

From a market surplus perspective, these price changes have no welfare impact – they’re just a transfer. But, from behind the veil of ignorance, these transfers are valued using the individual’s marginal utility of income. A fraction $s$ of the market will be insured and will face a price increase of $dp_I$, which will generate a welfare impact of $s(1 - s)D'(s)E[u’|Insured]ds$, where $E[u’|Insured]$ is the average marginal utility of income for the fraction $s$ of the market that is insured. The remaining fraction $1 - s$ of uninsured will face a price decrease of $dp_U$, which will generate a welfare impact of $(1 - s)sD'(s)E[u’|Uninsured]ds$, where $E[u’|Uninsured]$ is the average marginal utility of income for the fraction $1 - s$ of the market that is uninsured. Summing these two effects and normalizing by the average marginal utility of income of the insured to generate a money-metric utility measure yields the ex-ante value of expanding the size of the insurance market by $ds$:

$$EA(s) = s(1 - s)D'(s)\frac{E[u’|Insured] - E[u’|Uninsured]}{E[u’|Insured]}$$  \hspace{1cm} (1)

The first term, $s(1 - s)D'(s)$, can loosely be interpreted as the size of the transfer depicted by the blue arrow in Figure 2, Panel B. The slope of the demand curve, $D'(s)$, is the source of uncertainty against one wishes to be insured: a higher slope indicates a greater exposure to risk, which in this framework is captured by the extent to which expanding the size of the insurance market affects insurance prices. The second term, $\frac{E[u’|Insured] - E[u’|Uninsured]}{E[u’|Insured]}$, is the percentage difference in marginal utilities between the insured and uninsured population. Weighting the transfer by the difference in marginal utilities recovers the ex-ante value of insurance. This result is akin to the Baily-Chetty condition in optimal unemployment insurance that measures the value of more generous social insurance using the marginal utility of the beneficiaries (e.g. unemployed) relative to non-beneficiaries (e.g. employed). (Baily (1978); Chetty (2006)). Here, the beneficiaries of lower insurance prices are those who choose to purchase insurance. To the extent to which those who

---

3Such a setting would require individuals who purchase insurance to pay $p_I = $6.25 and those who do not purchase insurance to pay $p_U = $1.25, so that the marginal price of insurance is $p_I - p_U = $5 and the aggregate resource constraint holds.
have learned they have high demand for insurance at the time demand is measured (e.g. because they have a medical condition) also have a higher marginal utilities of income, transfers from the uninsured to the insured increase welfare from behind the veil of ignorance. In this sense, Equation (1) provides a general formula for recovering the ex-ante welfare of a larger (or smaller) insurance market.

Figure 2: Recovering Ex-Ante Demand

A. Marginal Increase in Fraction Insured

B. Transfer from Uninsured to Insured

C. Baily-Chetty Valuation of Transfer using Marg. Utilities

D. Generating Ex-Ante Demand

While weighting the transfer by the difference in marginal utilities provides a general expression for the ex-ante value of a larger insurance market, it is not readily implementable since generally one does not observe marginal utilities. This is analogous to the problems faced when implementing the Baily-Chetty condition in other settings, such as unemployment insurance. Building on the literature on optimal social insurance, I provide conditions under which one can approximate the difference in marginal utilities for the insured versus uninsured using Taylor expansions of the marginal utility function combined with (a) market level demand curves, $D(s)$, and (b) an estimate.
of risk aversion. As will be illustrated below, this implementation will rely on several assumptions that can be generalized or relaxed depending on the setting of interest. But, to illustrate in this simple example, recall the identity \( D(s) = m(s) \). Taking a first-order Taylor expansion of the marginal utility function, the percentage difference in marginal utilities is given by:

\[
\frac{E[u'|\text{Insured}] - E[u'|\text{Uninsured}]}{E[u'|\text{Insured}]} \approx -\frac{u_c}{u_c} (D(s) - E[D(s')|s \geq s'])
\]

where \(-\frac{u_c}{u_c}\) is the coefficient of absolute risk aversion and \(D(s) - E[D(s')|s \geq s']\) is the difference between the demand level when a fraction \(s\) are insured and the average willingness to pay of the uninsured, \(E[D(s')|s \geq s']\).\(^4\) In this example, the coefficient of relative risk aversion is 3, so that the coefficient of absolute risk aversion is approximately 3/25, where 25 is the average consumption in the population. Combining, this suggests the ex-ante value of insurance from expanding the market when exactly 50% have insurance is \(EA(0.5) = 0.5 \times 0.5 \times (-10) \times (3/25) \times (5 - 2.5) = 0.75\). From behind the veil of ignorance, individuals are willing to pay a markup of $0.75 to expand the size of the insurance market when 50% of the population has insurance, as illustrated in Figure 2, Panel C.

Figure 2, Panel D calculates \(EA(s)\) for all values of \(s \in [0,1]\). Adding this ex-ante value of insurance yields the "ex-ante demand curve", \(D(s) + EA(s)\), depicted by the solid red line in Figure 2, Panel D. The ex-ante welfare impact of expanding the insurance market from no one having insurance, \(s = 0\), to everyone having insurance, \(s = 1\), is $0.50, as indicated by the shaded area between the \(D(s) + EA(s)\) and the marginal cost curve, \(MC(s)\). In this sense, using the ex-ante demand curve recovers the ex-ante value of insurance depicted in Figure 1, Panel A, despite the fact that the market exists after some information is revealed. The ex-ante demand curve, \(D(s) + EA(s)\), facilitates graphical welfare analysis of policies that affect the observed market but from retain the normative perspective of being behind the veil of ignorance.

The remainder of this paper provides a generalized formula for the ex-ante demand curve and illustrates an application to settings motivated by estimates from the value of health insurance from Einav et al. (2010). The generalized formula in the next section relaxes the assumptions made above such as no moral hazard and perfect knowledge of risk at the time demand is measured. The basic formula in equation (1) continues to hold with a modification to the transfer term to account for the fact that demand may not equal marginal cost. I then provide general conditions under which one can also use the market demand curves combined with a coefficient of risk aversion (as in equation (2)) to measure the ex-ante demand curve.

\(^4\)As discussed in Section 3, in practice one can either import an estimate of risk aversion from another setting (e.g. CRRA of 3 or CARA of 5x10^-4), or one can estimate it internally using the willingness to pay for insurance against remaining risk. In particular, the coefficient of absolute risk aversion is approximately 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, \(-\frac{u_c}{u_c} \approx 2 \times \frac{\text{Markup}_{\text{insurance}}}{\text{Variance Reduction}_{\text{expenses}}}.\) In this simple example, there is no remaining risk so that demand does not reveal anything about risk aversion; but in realistic empirical applications one in principle can estimate this risk aversion coefficient internally.
estimate the demand and cost curves for a more versus less generous health insurance policy. From behind the veil of ignorance, the optimal size of the insurance market is larger than the deadweight loss minimizing size of the market. For a benchmark specification of a coefficient of absolute risk aversion of $5 \times 10^{-4}$, ex-ante welfare is maximized when 77.7% of the market owns insurance as opposed to the surplus-maximizing fraction of 75.6% in Einav et al. (2010). Moreover, the welfare cost from adverse selection is $14.25 versus $9.55, so that lost market surplus captures 67% of the ex-ante welfare cost of adverse selection.

The distinction between ex-ante welfare and deadweight loss is generally increasing in the square of the size of the premium and risk. This explains why the distinction between observed and ex-ante demand is small in the baseline Einav et al. (2010) setup. To illustrate how the results differ for larger insurable risks, I consider parameterizations that scale the demand and cost curves for top-up insurance in Einav et al. (2010). In particular, I consider a “medium risk” scenario that scales the estimates by 4x and large risk scenario that scales them by 10x. The latter more closely reflects the size of insurable risk associated with a full health insurance policy, as opposed to a top-up insurance policy. Lost market surplus captures 32% of the ex-ante welfare cost of adverse selection in the medium risk scenario and 18% of the ex-ante welfare cost of adverse selection in the large risk scenario. In this sense, ex-ante welfare diverges from market surplus as a measure of welfare when the size of the premium variation begins to comprise a significant fraction of the individuals’ income.

In addition to measuring the welfare cost of adverse selection, the model also sheds light on the desirability of allowing competitive markets versus mandates. Competitive markets may suffer from adverse selection; but mandates may require some individual to purchase insurance who don’t value it. For the medium and large risk scenarios above, deadweight loss measures of welfare prefer competitive markets relative to mandates. But, from behind the veil of ignorance, mandates deliver higher welfare than do competitive markets. In this sense, an ex-ante welfare perspective can lead to different conclusions about the desirability of government intervention in insurance markets.

The remainder of the paper proceeds as follows. Section 2 presents the general modeling structure and derives a general representation of the ex-ante demand curve. Section 3 provides conditions under which one can estimate the ex-ante demand curve using only market-level demand and cost curves combined with a measure of risk aversion. Section 4 calibrates the ex-ante demand curve using examples motivated from estimates in Einav et al. (2016). Section 5 concludes.

## 2 General Model

This section sets up a general model that will derive the demand and cost curves of the Einav et al. (2010) from an utility function that will be used to measure ex-ante (utilitarian) welfare. While the language will generally refer to a health insurance context, it is straightforward to amend the model to capture other insurance settings, such as unemployment insurance. The main result, provided in Proposition 1 derives an ex-ante demand curve that can be used to measure of the ex-ante welfare
impact of policy changes (e.g. subsidies and mandates) that affect the size of the insurance market.

2.1 Setup

Individuals face uncertainty over a future shock, captured by a random variable $\theta$. After learning $\theta$, individuals choose their non-medical consumption, $c$, and medical spending, $m$. Individuals have a utility function over these choices, $u(c, m; \theta)$, that is affected by the shock, $\theta$, which forms the source of uncertainty for the individuals. In addition to affecting their utility, the shock can also affect the individual’s income, $y(\theta)$.

There exists an insurance contract at price $p_I$ that allows individuals to obtain medical services at cost $x(m; \theta)$ yielding the budget constraint

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

where $y(\theta)$ is the individual’s income. Conversely, uninsured individuals must pay the full price of $m$, yielding a budget constraint

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

where $p_U$ is a penalty paid by individuals that are uninsured. Let $\{c^I(\theta), m^I(\theta)\}$ denote the choice of consumption and medical spending of a type $\theta$ if she is insured, and $\{c^U(\theta), m^U(\theta)\}$ if she is uninsured.\(^5\)

At the time individuals make the decision to be insured or uninsured, individuals may know something about their particular type $\theta$, which I denote by a signal $\bar{s} \in [0, 1]$. For simplicity, I follow Einav et al. (2010) and assume that only the relative price of insurance, $p_I - p_U$, affects demand. Appendix A provides a general statement of Proposition 1 when demand is affected differentially by increases of $p_U$ as opposed to decreases in $p_I$.

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

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

so that all $\bar{s}$ such that $p_I - p_U \leq D(\bar{s})$ will choose to purchase insurance, whereas types $\bar{s}$ for which $D(\bar{s}) > p_I - p_U$ will choose to remain uninsured and pay the penalty $p_U$. Without loss of generality, one can assume that $\bar{s}$ is ordered so that demand, $D(\bar{s})$, is decreasing in $\bar{s}$.

\(^5\)The notation $m^I(\theta)$ implies that $m^I(\theta)$ is not a function of the price, $p_I$. 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. For now, I adopt the common assumption that $m^I(\theta)$ does not depend on $p_I$, but it is straightforward to relax this assumption: if $m^I$ depends on $p_I$, then the cost to an insurer of raising/lowering their prices would also include a component from the impact of these price changes on the costs of their insured pool.

Similarly, I make the simplifying assumption that $m^U(\theta)$ does not depend on $p_U$. However, in contrast to the assumption that $m^I(\theta)$ does not depend on $p_I$, this assumption is without loss of generality because of the envelope theorem: $m^U(\theta)$ is fully paid by the individual so that behavioral responses of $m^U$ do not affect welfare measures of either the individual or other parties.
Following Einav et al. (2010), define the average cost of insurance when a fraction \( s \) of the market owns insurance, \( AC (s) \), where

\[
AC (s) = E [m^I (\theta) - x (m^I (\theta); \theta) | \bar{s} \leq s]
\] (4)

Let \( MC (s) \) characterize how this average cost changes as the size of the market expands, \( MC (s) = \frac{4}{s^3} [sAC (s)] \), where \( sAC (s) \) is the total cost of the insured. Given the assumption that individuals' choices are not affected by prices \( p_U \) and \( p_I \), this marginal cost is the net difference between expenditures and out-of-pocket spending for the marginal type, \( s \):

\[
MC (s) = E [m^I (\theta) - x (m^I (\theta); \theta) | \bar{s} = s]
\] (5)

Finally, I let \( p_I (s) \) and \( p_U (s) \) denote the prices of insurance and being uninsured that satisfy the demand curve and resource constraints of the economy when a fraction \( s \) of the market owns insurance. Hence,

\[
D (s) = p_I (s) - p_U (s)
\] (6)

and the resource constraint requires the total amount of resources collected from the insured and uninsured equals the total cost of the insured:

\[
sp_I + (1 - s)p_U = sAC (s)
\] (7)

The prices \( p_I (s) \) and \( p_U (s) \) are defined implicitly as solutions to equations (7) and (6).

### 2.2 Ex-Ante Welfare

Ex-ante welfare is defined as the population average utility level in the population, \( u (c, m; \theta) \). Let \( W (s) \) denote this average utility when prices are such that a fraction \( s \) of the market owns insurance. This is given by

\[
W (s) = \int_0^s E \left[ u (y (\theta) - p_I (s), m^I (\theta)) | \bar{s} = s' \right] ds' + \int_s^1 E \left[ u (y (\theta) - m^U (\theta) - p_U (s), m^U (\theta); \theta) | \bar{s} \right] d\bar{s}
\] (8)

The first term captures the welfare of the insured and the second term captures the welfare of the uninsured. The aim of this paper is to use \( W (s) \) as the welfare measure for evaluating policies such as subsidies and mandates that affect the size of the insurance market, \( s \).

If one observed or estimated the utility function, one could directly measure \( W (s) \). This would be analogous to the approach to measuring welfare taken by Finkelstein et al. (2016) and Handel et al. (2015). Here, I instead follow the “sufficient statistics” approach of Einav et al. (2010) and build a measure of \( W (s) \) from the demand and cost curves.

Following Einav et al. (2010), I use the demand function, \( D (s) \), to capture the impact on the utility of the uninsured, \( E [u (y (\theta) - m^U (\theta) - p_U, m^U (\theta); \theta) | \bar{s}] = E [u (y (\theta) - D (\bar{s}) - p_U, m^U (\theta); \theta) | \bar{s}] \).
This yields an alternative expression for $W(s)$ that does not require keeping track of the uninsured utility:

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

Now, consider a hypothetical policy change that expands the size of the market through increasing $p_U$ and decreasing $p_I$. For any $s$, these prices, $p_I(s)$ and $p_U(s)$, must satisfy the resource constraint (Equation (7)) and are consistent with demand (Equation (6)).

The marginal welfare impact of expanding the size of the market through this policy is given by

$$W'(s) = -sp'_I(s) E[u'(y(\theta) - p_I(s), m^I(\theta)) | \tilde{s} \leq s]$$

$$- (1 - s)p'_U(s) E[u_c(y(\theta) - D(\tilde{s}) - p_U(s), m^I(\theta); \tilde{s}) | \tilde{s} \geq s]$$

The first term captures the welfare increase from lower prices for the insured ($p'_I < 0$) and the second term captures the welfare cost of having higher prices faced by the uninsured ($p'_U > 0$). These transfers have no impact on the total amount of economic surplus in the market; but from an ex-ante perspective these transfers are weighted by the marginal utilities of income if insured or uninsured. Combining equation (9) with the resource constraint yields the main result.

**Proposition 1.** The marginal welfare impact of expanding the size of the insurance market from $s^*$ to $s^* + ds$ is given by

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

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

$$EA(s^*) = (1 - s^*) \left( MC(s^*) - D(s^*) - s^*D'(s^*) \right) \beta(s^*)$$

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

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

**Proof.** See Appendix B
the uninsured to the insured, which is given by \((1 - s^*)(MC(s^*) - D(s^*) - s^*D'(s^*))\), and the percentage difference between the marginal utilities of the insured and uninsured, \(\beta(s^*)\). The size of the transfer, \((1 - s^*)(MC(s^*) - D(s^*) - s^*D'(s^*))\) is fully identified from the demand and cost curves in the framework of Einav et al. (2010). This term is identical to the transfer in equation (1) when demand equals marginal cost, \(D(s) = MC(s)\), as in the example in the introduction.

2.3 Welfare Analyses

Optimal size of the insurance market  What is the optimal size of the insurance market from behind the veil of ignorance? Let \(s^{Ex-Ante}\) denote the size of the market that maximizes ex-ante welfare, \(W(s)\). This equates the ex-ante demand, \(D(s) + EA(s)\), equal to marginal cost, \(MC(s)\),

\[
D\left(s^{Ex-Ante}\right) + EA\left(s^{Ex-Ante}\right) - MC\left(s^{Ex-Ante}\right) = 0
\]

For comparison, one can also measure total economic surplus of the observed market as in Einav et al. (2010), which is given by the difference between the demand and marginal cost curves. Let \(s^{Observed}\) denote the size of the market that maximizes economic surplus of the observed market, which solves

\[
D\left(s^{Observed}\right) = MC\left(s^{Observed}\right)
\]

We assume for simplicity that these solutions are both unique. Proposition 1 suggests that as long as the marginal utilities of the insured are higher than the uninsured, the optimal size of the insurance market is larger than is suggested by economic surplus.

\[\text{Corollary 1. Suppose } \beta\left(s^{Observed}\right) > 0. \text{ Then } s^{Ex-Ante} > s^{Observed}. \text{ Conversely, if } \beta\left(s^{Observed}\right) < 0 \text{ then } s^{Ex-Ante} < s^{Observed}.\]

To see this, note that at \(s = s^{Observed}\) we have \(EA(s) = (1 - s) sD'(s) \beta(s) > 0\) whenever \(\beta\left(s^{Obseved}\right) > 0\). Hence, from behind the veil of ignorance, welfare can always be improved by expanding the insurance market beyond the size that maximizes economic surplus whenever \(\beta\left(s^{Observed}\right) > 0\). In this sense, the optimal size of the insurance market from behind the veil of ignorance necessarily will involve marginal deadweight loss, \(MC\left(s^{Ex-Ante}\right) > D\left(s^{Ex-Ante}\right)\).

Welfare cost from adverse selection  What is the welfare cost of adverse selection from behind the veil of ignorance? Since Akerlof (1970), it has been recognized that competitive markets may lead to an inefficient provision of insurance because prices must reflect average, not marginal costs. Let \(s^{CE}\) denote the size of the market that satisfies \(D\left(s^{CE}\right) = AC\left(s^{CE}\right)\), which I assume for simplicity to be unique. The ex-ante demand curve defined in Proposition 1 can help measure the ex-ante welfare loss from a competitive equilibrium insuring a fraction \(s^{CE}\) of the market, as opposed to having an optimal size of the insurance market, \(s^{Ex-Ante}\).

\[\text{Corollary 2. The ex-ante welfare loss from adverse selection in a competitive equilibrium is given}\]
One can also compare the welfare cost of a competitive equilibrium allocation, \( s^{CE} \), from an ex-ante welfare perspective, \( W^{CE} \), to the welfare cost from adverse selection as measured by total market surplus, as in Einav et al. (2010). This is given by

\[
M^{CE} = \int_{s^{CE}}^{s_{\text{Observed}}} [D(s) - MC(s)] \, ds
\]

It is straightforward to show that as long as the cost curves slope down (i.e. there is adverse selection) and \( \beta(s) > 0 \), then \( W^{CE} > M^{CE} \).

In short, Proposition 1 provides a method for characterizing the welfare impact of insurance mandates and subsidies from behind the veil of ignorance. The key requirements are the market-level demand and cost curves, combined with an estimate of the difference in marginal utilities between the insured and uninsured. If the insured have higher (lower) marginal utilities than the uninsured, then the welfare impact of expanding the insurance market is higher (lower) than is suggested by comparing solely the observed demand and cost curves.

### 2.4 Insurance versus Redistribution

The formula in Proposition 1 calculates welfare from behind a single veil of ignorance – i.e. before any information about \( \theta \) is known. The approach provided here can also be amended to facilitate welfare analysis after some observable information, \( X \), has been revealed about \( \theta \). 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.

To capture this, 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.\(^7\) 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 (13):

\[
\beta(s, x) = \frac{E[u_c(y(\theta) - p_I(s), m^{I}(\theta); \theta) | \hat{s} \leq s_x, X = x] - E[u_c(y(\theta) - D(\hat{s}) - p_U(s), m^{I}(\theta); \theta) | \hat{s} \geq s_x, X = x]}{E[u_c(y(\theta) - p_I(s), m^{I}(\theta); \theta) | \hat{s} \leq s_x, X = x]}
\]

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 \( p_U(s) = MC(s) - D(s) - s \frac{\partial D}{\partial s} \), 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

\(^7\)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.
given by
\[
EA(s, x) = (1 - s_x) \left( MC(s) - D(s) - s \frac{\partial D}{\partial s} \right) \beta(s_x, x)
\]
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.

For simplicity, the remainder of the paper focuses on the case of measuring ex-ante welfare from behind a complete veil of ignorance.

3 Implementation Using Market-Level Demand and Cost Curves

The key requirement for constructing the ex-ante demand curve is an estimate of the difference in marginal utilities between insured and uninsured, \( \beta(s) \). This section shows how one can make several assumptions common in the literature on social insurance (e.g. Baily (1978); Chetty (2006)) to write \( \beta(s) \) as a function of only market level demand curves combined with a measure of risk aversion. These assumptions are certainly not without loss of generality, and I discuss their potential violations below. But, they provide a way of estimating the ex-ante demand curve with little or no additional data requirements relative to what is required in Einav et al. (2010). In this sense, it provides a benchmark method for understanding whether in a given context there is an important distinction between ex-ante and observed demand.

3.1 Measuring Marginal Utilities

To begin, suppose that utility is state-independent so that the marginal utility of income only depends on the level of consumption. In this case, a Taylor expansion of \( u_c \) around a given consumption level \( \bar{c} \) yields
\[
\begin{align*}
    u_c \left( y(\theta) - D(s) - p_U, m^I(\theta) ; \theta \right) & \approx u_c + u_{cc} | C(s) - \bar{c} | \\
\end{align*}
\]
where \( C(s) = y(\theta) - D(s) - p_U \) is the level of consumption of the insured individual with demand \( D(s) \). More generally, the marginal utility of consumption could also depend on the level of \( m \) or the state of nature, \( \theta \). By ruling this out, the level of consumption, \( C(s) \), provides information about the marginal utility of consumption.

Second, I assume that the curvature of utility, \( u_{cc} \), is common across individuals with the same level of demand, \( D(s) \). Let \( \gamma = -\frac{u_{cc}(s)}{u_c(s)} \) denote the coefficient of absolute risk aversion. Under
these two assumptions, one can write

\[ \beta (s^*) = \gamma E [C (s) - C (s^*) \mid s \geq s^*] \]  

(15)

Finally, note two terms can affect the level of consumption: \( y (\theta) \) and \( D (s) \). Suppose that income does not vary systematically across the level of insurance demand, \( s \). In this case, differences in consumption equal differences in insurance demand, \( \frac{d}{ds} E [y (\theta) - D (\hat{s}) - p_U (s) - \hat{c} | \hat{s}] = \frac{d}{ds} D (\hat{s}) \). This yields:

\[ \beta (s^*) = \gamma E [D (s^*) - D (s) \mid s \geq s^*] \]  

(16)

The difference in marginal utilities can be approximated by the difference in the willingness to pay for insurance for the marginal type and the average uninsured individual, multiplied by the coefficient of absolute risk aversion.

Under these three assumptions, the slope of the demand curve provides information about the difference in marginal utilities for the insured and uninsured. Loosely, it measures how much information about risk has been revealed. Steeper demand curves correspond to greater differences in marginal utilities between insured and uninsured, which in turn suggest a greater ex-ante value of insurance.

Substituting the approximation for \( \beta (s) \) in equation (16) into the general expression in equation (12) yields

\[ EA (s^*) = (1 - s^*) (MC (s^*) - D (s^*) - s^* D' (s^*)) \gamma E [D (s^*) - D (s) \mid s \geq s^*] \]  

(17)

Equation (17) illustrates how \( EA (s) \) is increasing in the square of the size of the demand and cost curves. Replacing \( MC (s) \) and \( D (s) \) with \( aMC (s) \) and \( aD (s) \) yields \( a^2 EA (s) \). This suggests that the distinction between observed demand, \( D (s) \), and the ex-ante demand curve, \( D (s) + EA (s) \), is increasing in the square of the size of the insurable risk. In settings in which the insurable risk is small, the distinction between ex-ante and observed demand is likely to be small; but in settings where the costs constitute a larger fraction of income, the distinction between ex-ante and observed

---

\(^8\)To see this, note that:

\[ \beta (s^*) = \frac{E [u_c (y (\theta) - p_I (s^*), m_I (\theta); \theta) \mid s \leq s^*] - E \left[ \frac{\partial D (s, p_U (s^*))}{\partial p_U} u_c (y (\theta) - D (s) - p_U (s^*), m_I (\theta); \theta) \mid s \geq s^* \right]}{E [u_c (y (\theta) - p_I (s^*), m_I (\theta); \theta) \mid s \leq s^*] - E \left[ u_c (y (\theta) - D (s) - p_U (s^*), m_I (\theta); \theta) \mid s \geq s^* \right]} \]

\[ \approx \frac{E [u_c (y (\theta) - D (s^*, p_U (s^*)), m_I (\theta); \theta) \mid s \leq s^*] - E \left[ u_c (y (\theta) - D (s, p_U (s^*)), m_I (\theta); \theta) \mid s \geq s^* \right]}{u_c (s^*)} \]

\[ \approx \frac{u_{cc} (s^*) E [u_{cc} (y (\theta) - D (s^*, p_U (s^*))) \mid s \leq s^*] - E \left[ u_{cc} (y (\theta) - D (s, p_U (s^*))) \mid s \geq s^* \right]}{u_c (s^*)} \]

\[ \approx \frac{u_{cc} (s^*) E [D (s, p_U (s^*)) - D (s^*, p_U (s^*))) \mid s \geq s^*]}{u_c (s^*)} \]

\[ \approx \gamma E [D (s^*, p_U (s^*)) - D (s, p_U (s^*)) \mid s \geq s^*] \]
demand can be quite large, as will be illustrated in Section 4.

**Measuring Risk Aversion**  In order to translate the slope of the demand curve into differences in marginal utilities, one requires an estimate or risk aversion, \( \gamma \). This can imported from another setting; for example, one might wish to assume that the coefficient of absolute risk aversion is \( 5 \times 10^{-4} \) as will be done below. But, it is also possible to infer the risk aversion coefficient from the willingness to pay individuals have to insure against their remaining risk. For example, in a CARA-Normal model with normally distributed risks, the coefficient of absolute risk aversion is given by

\[
\gamma \approx 2 \frac{\text{Markup}}{\text{Variance Reduction}} \approx 2 \frac{D(s) - MC(s)}{\text{var}(m^U(\theta)) - \text{var}(x(m^I(\theta); \theta))}
\]

(18)

Using this formula, one can take an estimate of the variance reduction in expenditures induced by insurance, \( \text{var}(m^U(\theta)) - \text{var}(x(m^I(\theta); \theta)) \), and use this to form an internal estimate the coefficient of absolute risk aversion that is consistent with the observed willingness to pay for insurance. In this sense, one need not necessarily rely on an external measure of risk aversion, but rather can instead rely on an internal estimate of the variance reduction offered by insurance combined with the markup individuals are willing to pay over their average costs, \( D(s) - MC(s) \), to arrive at a risk aversion coefficient that can be used to measure the ex-ante value of insurance.

### 3.2 Violations of Equation (13)

Generating an estimate of the difference in marginal utilities, \( \beta(s) \), required several assumptions that may be violated in practice. Here, I discuss several potential violations, how they could bias the results, and how one could potentially generalize the implementation to account for the violations.

To begin, note that the baseline specification of \( \beta(s) = \gamma E[D(s^*) - D(s) | s \geq s^*] \) assumes that a $1 higher demand for insurance leads to $1 less consumption. However, one could imagine that individuals could borrow and/or save, so that $1 higher demand does not correspond to a $1 lower consumption. If individuals can spread this demand payment over \( T \) periods through borrowing and saving, then the impact of the additional dollar on consumption will be

\[
\beta(s) = \frac{1}{T} \gamma E[D(s^*) - D(s) | s \geq s^*]
\]

so that the difference in marginal utilities between the insured and uninsured types will be smaller than is suggested by the difference in willingness to pay, \( E[D(s^*) - D(s) | s \geq s^*] \).

Second, those with higher demand for insurance might have lower (or higher) incomes, \( y(\theta) \).

---

\( ^9 \) In practice, whether or not individuals are able to save may be less of a concern if one uses “internal” measures of risk aversion from equation (18). The value of \( \gamma \) that is identified in equation (18) is the willingness to pay for a reduction in variance in out-of-pocket expenditures. If individuals are able to save, then the full variance reduction offered by insurance is actually smaller than \( \text{var}(m^U(\theta)) - \text{var}(x(m^I(\theta); \theta)) \). In this sense, the \( \gamma \) estimated in equation (18) reflects the willingness to pay for a reduction in out of pocket spending, analogous to the willingness to pay for a reduction in the price of insurance. In this sense, ex-ante willingness to pay estimates that rely on internal measures of risk aversion are robust to cases in which the individuals are able to save.
Indeed, one potential model of insurance demand could be that those with low willingness to pay are those who faced liquidity shocks. In this case, the insured could even have a lower marginal utility of income than the uninsured. Here, one can measure $\beta(s)$ using equation (15) instead of equation (16). The key requirement however, is that the researcher must measure how consumption varies with demand, $s$. This is a more significant empirical requirement than the benchmark implementation that infers how consumption varies from the demand curve.

Third, risk aversion could be heterogeneous. If one knew how risk aversion varied with the level of consumption, one could add a covariance between risk aversion and the marginal utility of consumption, as proposed by Andrews and Miller (2013). In standard models, one would expect those with the highest degree of risk aversion to have not only the highest demand for insurance but also the highest marginal utility of consumption. In this case, the estimate of $\beta(s)$ in equation (16) would understate the true difference in marginal utilities, and therefore understate the magnitude of the ex-ante demand curve.

Lastly, one there could be a violation of state dependence, so that those with high demand for insurance have a lower (or higher) marginal utility of consumption even conditional on having a same level of consumption. Empirically implementing the general formula in Proposition 1 would be quite difficult in this case. But, heuristically if those with higher demand for insurance also have a higher (smaller) marginal utility of income even conditional on consumption levels, then the estimate of $\beta(s)$ in equation (16) would understate (overstate) the true difference in marginal utilities, and therefore understate (overstate) the ex-ante demand curve.

3.3 Summary

In summary, the implementation in equation (16) provides a method for estimating $\beta(s)$ using market level demand curves combined with estimates of risk aversion. Although the validity of this formula relies on strong assumptions, it provides a benchmark for evaluating the potential degree to which an observed demand curve, $D(s)$, differs from the willingness to pay for a larger insurance market from behind the veil of ignorance, $D(s) + EA(s)$. Going forward, I illustrate how to apply this formula in settings motivated by existing work.

4 Illustration using Three Examples

This section illustrates how to construct ex-ante demand curves using estimates from Einav et al. (2010). 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. To illustrate how the ex-ante demand curve varies with the size of the insurable risk, I consider three variations of their estimates. I begin by comparing the welfare cost of adverse selection from an ex-ante perspective, $W^{Ex-Ante}$, as opposed to a comparing market surplus, $M^{CE}$.
4.1 Welfare Cost of Adverse Selection

Small Risk: Demand and Cost Curves from Einav et al. (2010) The first example comes directly from the demand and cost curve estimates in Einav et al. (2010). Figure 3, Panel A presents these estimates. The solid black line presents the estimated demand curve. A $100 increase in the price of insurance leads to a 7pp reduction in the fraction of the market purchasing the more generous policy. Higher prices also lead to a higher average cost of the insured: A $100 increase in price that reduces the market size by 7pp leads to a higher average cost of the insured population of roughly $15.5, as illustrated by the long-dashed blue line. Taking the derivative of total costs yields the marginal cost curve, illustrated by the dashed red line.

A competitive equilibrium in this environment would result in roughly $s^{CE} = 61.7\%$ of the market purchasing the more generous policy, reflected by the intersection between the average cost curve and the demand curve in Panel A. This occurs with a price of $D(s^{CE}) = AC(s^{CE}) = 463.5$. But, those that are indifferent to purchasing insurance at a price of $463.5$ on average impose a cost on the insurance company, $MC(s^{CE})$, that is less than their willingness to pay. Aggregating across these potential trades for which demand is above marginal cost, the aggregate lost surplus from adverse selection is $M^{CE} = 9.57$. To place this in perspective, this is roughly 3.3% of the size of the market of $286 (61.7\%$ of $463.5$). In this sense, the welfare cost from adverse selection is small.

Figure 3 Panel B presents the ex-ante demand curve, $D(s) + EA(s)$ in red. To calculate $EA(s)$, I use the formula in equation (17). I take a value for the coefficient of absolute risk aversion of $\gamma = 5 \times 10^{-4}$, similar to the mean estimate in Handel et al. (2015). To see how the ex-ante demand curve is calculated, note that the linearity of demand implies $E[D(s^*)] = D(0) - (1 - s^*)MC(s^*)$, so that

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

For example, at $s^* = 61.7\%$ (the competitive equilibrium value), the ex-ante component of demand is given by

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

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

$$EA(0.617) = 39.4$$

as reflected by the difference between the observed demand curve, $D(s)$, and the ex-ante demand curve, $D(s) + EA(s)$, in Figure 3, Panel B.

\[^{10}\text{In principle, one can also calculate an ‘in-sample’ measure of the coefficient of absolute risk aversion that is consistent with the extent to which the insurance product reduces the variance of out-of-pocket costs and (b) the markup that the marginal individual is willing to pay, } D(s) - AC(s).\]
The ex-ante demand curve intersects the marginal cost curve when a fraction \( s^{Ex-Ante} = 77.7\% \) of the market is insured, fairly similar to the size of the market that maximizes total surplus, \( s^{Observed} = 75.6\% \). Integrating between the ex-ante demand curve and marginal cost curve between \( s^{CE} \) and \( s^{Ex-Ante} \) yields an ex-ante welfare cost of adverse selection of \( W^{Ex-Ante} = $14.25 \). 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, \( \frac{M^{CE}}{W^{Ex-Ante}} \approx 0.67 \).

**Figure 3: Ex-Ante Demand in Three Examples**

A. Demand and Cost Curves from Einav et al. (2010)

B. Ex-Ante Demand in Einav et. al. (2010)

C. Ex-Ante Demand in Medium Risk Scenario

D. Ex-Ante Demand in Large Risk Scenario

Medium Risk: Scaling Demand and Cost Curves from Einav et al. (2010) by 4x The top-up insurance setting of Einav et al. (2010) considers an insurance product that insures against a relatively small fraction of total medical expenses. As noted in Section 3, the ex-ante component...
of demand is increasing in the square of the size of the insurable risk. To illustrate this, the next two examples derive the ex-ante demand curve for larger risk scenarios and illustrate a greater potential divergence between observed and ex-ante demand.

To begin, I consider a 4x scaling of the demand and cost curves in Einav et al. (2010). This scaling has no impact on the fraction of the market that is insured in a competitive equilibrium, \( s_{CE} = 61.7\% \) or on the size of the market that maximizes market surplus, \( s_{\text{Observed}} = 75.6\% \). However, it does significantly change the ex-ante demand curve. From behind the veil of ignorance, the optimal size of the insurance market is \( s_{\text{Ex-Ante}} = 81.6\% \), and the ex-ante welfare cost of adverse selection in a competitive equilibrium is \( W_{\text{Ex-Ante}} = $120.62 \), which is more than three times as large as the lost market surplus of \( M_{CE} = $38.26 \). In this sense, lost market surplus captures only 32% of the ex-ante welfare cost of adverse selection.

**Large Risk: Scaling Demand and Cost Curves from Einav et al. (2010) by 10x** Finally, I consider a “large risk” example that scales the demand and cost curves in Einav et al. (2010) by a factor of 10. This yields average costs around $5,000, which more closely approximates the costs associated with the extensive margin purchases of health insurance. As shown in Panel D of Figure 3, there is a larger divergence between ex-ante and observed demand. Ex-ante welfare is maximized when a fraction \( s_{\text{Ex-Ante}} = 85.6\% \) of the market is insured, and the ex-ante welfare cost from adverse selection is \( W_{\text{Ex-Ante}} = $427 \). This contrasts with the lost surplus from adverse selection of \( M_{CE} = $77 \), which captures only 18% of the total ex-ante welfare cost of adverse selection. In short, the ex-ante welfare perspective can generate significantly different conclusions than what would be inferred solely from comparing observed demand and cost curves, especially when considering insurance against large risks.

**4.2 Markets versus Mandates**

Do government insurance mandates deliver higher welfare than competitive insurance markets? Can the government significantly improve welfare by optimally subsidizing competitive markets? This section illustrates how the ex-ante demand curve provides guidance to these questions from the ex-ante welfare perspective.

To illustrate, Figure 4 presents the observed and ex-ante demand curve, along with the marginal cost curve, for the medium risk example. Panel A presents the observed demand curve; Panel B presents the ex-ante demand curve. To calculate the (ex-ante) welfare impact of a mandate relative to a competitive equilibrium, one can integrate between the demand (or ex-ante demand) curve and marginal cost curve between \( s_{CE} \) and 1. From both a market surplus and ex-ante welfare perspective, expanding the size of insurance market starting at \( s_{CE} \) increases welfare. However, expanding all the way to 100% coverage involves covering some individuals who value their insurance less than costs.

In the medium risk example, a full mandate reduces market surplus: Expanding the market to \( s_{\text{Observed}} \) increases market surplus by $38.26, but expanding from \( s_{\text{Observed}} \) to 100% reduces market
surplus by $117.84, as shown in Panel A of Figure 4. Mandates generate a net loss of surplus of $79.58 relative to competitive markets.

However, from behind the veil of ignorance mandates deliver higher ex-ante welfare than the competitive market allocation. Expanding the market from $s^{CE}$ to $s^{Ex-Ante}$ increases ex-ante welfare by $120.62$, and further expanding to 100% coverage reduces ex-ante welfare by $94.70$. From behind the veil of ignorance, individuals would be willing to forego $25.92$ of consumption in all states of the world in order to have an insurance mandate instead of a competitive equilibrium, despite the fact doing so lowers total market surplus. In this sense, an ex-ante welfare perspective can lead to different conclusions about the desirability of insurance mandates.

**Figure 4: Markets versus Mandates: Medium Risk Example**

A. Total Market Surplus Comparison

<table>
<thead>
<tr>
<th>$s^{CE}$</th>
<th>$s^{Ex-Paid}$</th>
<th>$s^{Ex-Ante}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$38.26$</td>
<td>$117.84$</td>
<td></td>
</tr>
</tbody>
</table>

Mandates have lower total surplus than competitive markets

B. Ex-Ante Welfare Comparison

<table>
<thead>
<tr>
<th>$s^{CE}$</th>
<th>$s^{Ex-Ante}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$120.62$</td>
<td>$94.70$</td>
</tr>
</tbody>
</table>

Mandates have higher ex-ante welfare than competitive markets

5 Conclusion

This paper provides reduced-form methods to infer the ex-ante value of insurance from demand and cost curve estimates generated from points in time when individuals may have knowledge about their future risk. The results generally suggest insurance is more valuable from an ex-ante perspective, so that measures of deadweight loss understate the ex-ante value of insurance. The methods can be used in future empirical to quantify the potential divergence between ex-ante and observed demand. In settings in which the observed risk is small, or in which premiums are already heavily subsidized, one would expect the distinction to be small. But, in settings where the risk is a non-trivial fraction of the individuals income, the observed and ex-ante welfare perspectives can lead to different conclusions about policy interventions, such as subsidies and mandates.
References


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, 1, 2, 2.1, 2.1, 2.2, 2.2, 2.3, 2.3, 3, 4, 4.1, 4.1, 4.1


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

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.

**Proposition.** The marginal welfare impact of expanding the size of the insurance market from $s^*$ to $s^* + ds$ is given by

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

(19)
where $EA(s^*)$ is the ex-ante value of expanding the size of the insurance market,

$$EA(s^*) = \frac{1 - s^*}{1 + s^* (\frac{\partial D}{\partial EC} - 1)} \left( MC(s^*) - (p_I(s^*) - p_U(s^*)) - s^* \frac{\partial D}{\partial s} \beta(s^*) \right)$$  \hspace{1cm} (20)

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

$$\beta(s) = \frac{E[u_c(y(\theta) - p_I(s), m^I(\theta); \theta) | \bar{s} \leq s] - E[u_c(y(\theta) - D(\bar{s}, p_U(s)), m^I(\theta); \theta) | \bar{s} \geq s]}{E[u_c(y(\theta) - p_I(s), m^I(\theta); \theta) | \bar{s} \leq s]}$$  \hspace{1cm} (21)

### B Proof of Proposition 1

This Appendix walks through the proof of Proposition 1. Let $p_I(s)$ and $p_U(s)$ satisfy the resource constraint (7) and the demand constraint, $p_I(s) = D(s, p_U(s))$ when fraction $s$ of the market purchasing insurance when facing those prices. Ex-ante welfare, $V(s)$, is given by

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

Substituting the demand function $E[u(y(\theta) - m^U(\theta) - p_U(s), m^U(\theta); \theta) | \bar{s}] = E[u(c - D(\bar{s}, p_U(s)), m^I(\theta); \theta) | \bar{s}]$ yields:

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

so the marginal welfare impact is given by

$$V'(s) = -sp_I(s) E[u'(y(\theta) - p_I(s), m^I(\theta); \theta) | \bar{s} \leq s] - (1 - s) p_U(s) E[\frac{\partial D(\bar{s}, p_U(s))}{\partial p_U} u_c(y(\theta) - D(\bar{s}, p_U(s)), m^I(\theta); \theta) | \bar{s} \geq s]$$

Now,

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

so that

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

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. 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 \((1 - s) p'_I(s)\) \(E\left[u_c\left(y(\theta) - p_I(s), m^I(\theta); \theta\right) \mid \bar{s} \leq s\right]\) and then dividing by \(E\left[u_c\left(y(\theta) - p_I(s), m^I(\theta); \theta\right) \mid \bar{s} \leq s\right]\) yields

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

or

\[
\frac{V'(s)}{E\left[u_c\left(y(\theta) - p_I(s), m^I(\theta); \theta\right) \mid \bar{s} \leq s\right]} = -MDWL + (1 - s)p'_U(s) \\
\times \left(E\left[u_c\left(y(\theta) - p_I(s), m^I(\theta); \theta\right) \mid \bar{s} \leq s\right] - E\left[\frac{\partial D(\bar{s},p_U(s))}{\partial p_U}u_c\left(y(\theta) - D(\bar{s}, p_U(s)), m^I(\theta); \theta\right) \mid \bar{s} \geq s\right]\right)
\]

where \(p'_U(s) > 0\). So, at an optimum where marginal utilities are higher amongst the insured, we have a positive MDWL as long as \(\frac{\partial D(\bar{s},p_U(s))}{\partial p_U} \approx 1\).

What is \(\frac{\partial D(\bar{s},p_U(s))}{\partial p_U}\)? Higher penalties should increase the WTP people have for insurance, \(\frac{\partial D}{\partial p_U} > 0\). We have \(D(s,p_U(s)) = p_I(s)\) so that

\[
\frac{\partial D}{\partial s} + \frac{\partial D}{\partial p_U} p'_U(s) = p'_I(s)
\]

or

\[
\frac{\partial D}{\partial s} = p'_I(s) - \frac{\partial D}{\partial p_U} p'_U(s)
\]

so that as \(s\) increases, it influences demand through the differences in prices, \(p_I\) and \(p_U\). If \(\frac{\partial D}{\partial p_U} \approx 1\), then \(\frac{\partial D}{\partial s} = p'_I(s) - p'_U(s)\): the increase in the size of the market can come from two sources: lower prices or increase penalties. Note \(E\left[u\left(y(\theta) - m^U(\theta) - p_U, m^U(\theta); \theta\right) \mid \bar{s}\right] = E\left[u\left(y(\theta) - D(\bar{s}, p_U), m^I(\theta); \theta\right) \mid \bar{s}\right]\)

so that differentiating with respect to \(p_U\)

\[
E\left[u_c\left(y(\theta) - m^U(\theta) - p_U, m^U(\theta); \theta\right) \mid \bar{s}\right] = E\left[\frac{\partial D}{\partial p_U} u_c\left(y(\theta) - D(\bar{s}, p_U), m^I(\theta); \theta\right) \mid \bar{s}\right]
\]

Note that \(D\) only depends on \(s\) (not \(\theta\)) so that we can solve

\[
\frac{\partial D}{\partial p_U} = \frac{E\left[u_c\left(y(\theta) - m^U(\theta) - p_U, m^U(\theta); \theta\right) \mid \bar{s}\right]}{E\left[u_c\left(y(\theta) - D(\bar{s}, p_U), m^I(\theta); \theta\right) \mid \bar{s}\right]}
\]

so that marginal utility when insured is lower than marginal utility when uninsured if and only if \(\frac{\partial D}{\partial p_U} > 1\).

Now, how can we approximate the difference in marginal utilities? We can take an approximation that assumes (a) \(\frac{\partial D(\bar{s},p_U(s))}{\partial p_U}\) is roughly constant in \(\bar{s}\) and (b) \(u_c\) is locally linear in \(s\). Then, Taylor expanding \(u_c\) we have

\[
\frac{\partial D(\bar{s},p_U(s))}{\partial p_U} u_c\left(y(\theta) - D(\bar{s}, p_U(s)), m^I(\theta); \theta\right) \approx \frac{\partial D(\bar{s},p_U(s))}{\partial p_U} \left[u_c\left(y(\theta) - D(\bar{s}, p_U(s)), m^I(\theta); \theta\right) - u_c \frac{\partial D}{\partial s}(s - \bar{s})\right]
\]
Taking expectations,
\[
E \left[ \frac{\partial D(s, p_U(s))}{\partial p_U} u_c \left( y(\theta) - D(s, p_U(s)), m^I(\theta); \theta \right) | \theta \geq s \right] = E \left[ \frac{\partial D(s, p_U(s))}{\partial p_U} \left( u_c \left( y(\theta) - D(s, p_U(s)), m^I(\theta); \theta \right) - u_{cc} \frac{\partial D}{\partial s} (s - \tilde{s}) \right) | \theta \geq s \right]
\]

where \( \frac{\partial D}{\partial p_U} \) is the difference of the impact of raising the price of insurance versus lowering the mandate penalty on demand and \( \frac{\partial D}{\partial s} \) is the slope of the demand curve holding \( p_U \) constant. In practice, we might assume \( \frac{\partial D}{\partial p_U} = 1 \), that marginal utilities are constant for the insured (who all pay the same price). In this case,
\[
E \left[ \frac{\partial D(s, p_U(s))}{\partial p_U} u_c \left( y(\theta) - D(s, p_U(s)), m^I(\theta); \theta \right) | \theta \geq s \right] = E \left[ u_c \left( y(\theta) - D(s, p_U(s)), m^I(\theta); \theta \right) - u_{cc} \left( D(s, p_U(s)) - D(s, p_U(s)) \right) | \theta \geq s \right]
\]

So, the difference between insured and uninsured marginal utilities is given by
\[
\left( E \left[ u_c \left( y(\theta) - p_I(s), m^I(\theta); \theta \right) | \theta \geq s \right] - E \left[ \frac{\partial D(s, p_U(s))}{\partial p_U} u_c \left( y(\theta) - D(s, p_U(s)), m^I(\theta); \theta \right) | \theta \geq s \right] \right) = \frac{u_{cc}}{u_c} E [D(s) - D(s) | \theta \geq s]
\]

So, the welfare impact is approximately
\[
\frac{V'(s)}{E \left[ u_c \left( y(\theta) - p_I(s), m^I(\theta); \theta \right) | \theta \geq s \right]} = -MDWL + (1 - s) p'_U(s) \left( \frac{-u_{cc}}{u_c} E [D(s) - D(s) | \theta \geq s] \right)
\]

(22)

Finally, 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 \( s \): \(-sp'_I(s) - (1 - s) p'_U(s) = (p_I(s) - p_U(s)) - MC(s)\). Note that
\[
sp'_I(s) + (1 - s) p_U(s) = sAC(s)
\]

so that
\[
p_I(s) + sp'_I(s) + (1 - s) p'_U(s) - p_U(s) = MC(s)
\]
or
\[
sp'_I(s) + (1 - s) p'_U(s) = MC(s) - p_I(s) + p_U(s)
\]

Moreover, differentiating equation (23) yields \( p'_I(s) = \frac{\partial D}{\partial s} + \frac{\partial D}{\partial p_U} p'_U(s) \). To see this, note that:
\[
-s \left[ \frac{\partial D}{\partial s} + \frac{\partial D}{\partial p_U} p'_U(s) \right] - (1 - s) p'_U(s) = (p_I(s) - p_U(s)) - MC(s)
\]
\[
-p_U(s) \left[ 1 + s \left( \frac{\partial D}{\partial p_U} - 1 \right) \right] = (p_I(s) - p_U(s)) - MC(s) + s \frac{\partial D}{\partial s}
\]
\[
-p'_U(s) = \frac{1}{1 + s \left( \frac{\partial D}{\partial p_U} - 1 \right)} \left[ (p_I(s) - p_U(s)) - MC(s) + s \frac{\partial D}{\partial s} \right]
\]

Combining with equation (22) yields the desired result. Under the additional approximation that \( \frac{\partial D}{\partial p_U} = 1 \),
and replacing $D(s) = p_F(s) - p_U(s)$, this yields

$$\frac{V'(s)}{E[u_c(y(\theta) - p_F(s), m^T(\theta); \theta)|\hat{s} \leq s]} = D(s) - MC(s) + (1 - s) \left(D(s) - MC(s) + s \frac{\partial D}{\partial s}\right) \left[-\frac{u_c}{u_c} E[D(s) - D(\hat{s}) | \hat{s} \geq s]\right]$$

which concludes the proof.