Using price variation to estimate welfare in insurance markets*

Liran Einav, Amy Finkelstein, and Mark R. Cullen†

April 2008

Please don't circulate or cite without permission.
Preliminary and incomplete draft. Some results may change.
Comments are extremely welcome.

Abstract. We show how standard consumer and producer theory can be applied to estimating welfare in insurance markets with selection. The key observation is that the same variation in prices needed to trace out the demand curve in any applied welfare analysis can also be used to trace out how costs vary as market participants endogenously respond to the price of insurance. With estimates of both the demand and cost curves, welfare analysis is straightforward. Moreover, since endogenous costs are the distinguishing feature of selection models, the analysis of the cost curve also provides a direct test for the existence of selection. We discuss the data required to implement this approach, and then apply it using individual-level data from a large private employer in the United States on the health insurance options, choices and medical expenditures of its employees and their dependents. We detect adverse selection in this market and estimate that its efficiency cost, if these choices occurred in a free market setting, would be about 0.2% and 2% of the surplus that could be generated from efficient pricing. We estimate that the social cost of the subsidy needed to achieve the efficient outcome is about an order of magnitude higher than the social welfare gain from correcting the market failure.

JEL classification numbers: C13, C51, D14, D60, D82, I11.

Keywords: Asymmetric information; adverse selection; health insurance; efficiency cost.

*We are grateful to Felicia Bayer, Brenda Barlek, Chance Cassidy, Fran Filpovits, Frank Patrick, and Mike Williams for innumerable conversations explaining the institutional environment of the company, to Colleen Barry, Susan Busch, Linda Cantley, Deron Galusha, James Hill, Sally Vegso, and especially Marty Slade for providing and explaining the data, to Tatyana Deryugina, Sean Klein, and James Wang for outstanding research assistance, and to Kate Bundorf, Raj Chetty, Peter Diamond, and Hanming Fang for helpful comments. The data were provided as part of an ongoing service and research agreement between the company and Yale, under which Yale faculty and staff perform jointly agreed-upon onging and ad-hoc research projects on workers' health, injury, disabilty and health care, and Mark Cullen serves as medical director for the corporation. We gratefully acknowledge support from the National Science Foundation grant #SES-0643037 (Einav), the Alfred P. Sloan Foundation (Finkelstein), and the John D. and Catherine T. MacArthur Foundation Network on Socioeconoimc Status and Health and Alcoa Inc. (Cullen).

†Einav: Department of Economics, Stanford University, and NBER, leinav@stanford.edu; Finkelstein: Department of Economics, MIT, and NBER, afink@mit.edu; Cullen: Occupational and Environmental Medicine Program, Yale School of Medicine, mark.cullen@yale.edu.

1 Introduction

The welfare loss from selection in private insurance markets is a classic result in economic theory. It provides the textbook economic rationale for the near-ubiquitous government intervention in insurance markets. Yet there has been relatively little empirical work devoted to quantifying the inefficiency that selection causes in a particular insurance market or the welfare consequences of alternative potential policy interventions. This probably reflects the considerable challenges posed by empirical welfare analysis in markets with hidden information.

In this paper, we show how standard consumer and producer theory – familiar to any student of intermediate micro – can be applied to welfare analysis of insurance markets with selection. The key feature of selection models is that firms' costs depend on which consumers purchase their products. Because market costs are endogenous to equilibrium price, empirical welfare analysis requires not only the usual estimation of how demand varies with price, but also the estimation of how the costs of the (endogenous) market participants vary with price.

This suggests a straightforward empirical approach to welfare analysis of selection in insurance markets. The same pricing variation that is needed to estimate the demand curve (or willingness to pay) in any welfare analysis – be it the consequences of tax policy, the introduction of new goods, or selection in insurance markets – can also be used to trace out how costs vary as the set of market participants changes. With these two curves in hand, welfare analysis of the inefficiency caused by selection – or of the consequences of a range of alternative potential public policy interventions – is simple and familiar.

This approach has several appealing features. For one thing, the shape of the estimated cost curve provides a direct test of whether selection exists, and whether it is adverse or advantageous. Specifically, rejection of the null hypothesis of a constant (i.e. horizontal) marginal cost curve allows us to reject the null hypothesis of no selection, while the slope of the marginal cost curve tells us whether the resultant selection is adverse (downward sloping) or advantageous (upward sloping). This is quite important, since the existence of selection is a necessary precursor to analysis of its welfare effects. Importantly, our "cost curve" test of selection is unaffected by the existence (or lack thereof) of moral hazard. This is a distinct improvement over the current "industry standard" – the important and widely used "bivariate probit" or "positive correlation" test of Chiappori and Salanie (2000) – which jointly tests for the existence of either adverse selection or moral hazard (but not for each separately).[1]

Beyond detecting selection, the proposed approach offers three key attractive features for empirical welfare analysis. First, it does not require the researcher to make assumptions about preferences or the structure of information in the market. Second, it is relatively straightforward to implement, and likely to be widely applicable. Cost data are likely to be much easier to obtain in insurance markets than in other product markets, since they involve information on accident occurrences or insurance claims, rather than insight into the underlying production function of the firm. In

[1] Of course, the "cost curve" test has an additional data requirement that the bivariate probit test does not, namely pricing variation that is exogneous to individual demand and insurable costs.

addition, the omnipresent regulation of insurance markets offers many potential sources of the exogenous pricing variation needed to estimate the demand and cost curves. Third, it is fairly general as it does not rely on specific institutional details. This suggests that it may be informative to compare estimates of the welfare cost of adverse selection obtained by this approach across different insurance markets, or different populations.

The chief limitation to our approach is that counterfactual analysis of policies that would introduce different products than those observed in the data – as opposed to merely changing the prices of existing products – is not feasible. Such analysis requires estimation of the structural primitives underlying the estimated demand and cost curves in the insurance market, as has been done by a few recent papers (Einav, Finkelstein and Schrimpf, 2007; Hosseini, 2007; and Lustig, 2007). These papers specify and estimate a structural model of insurance demand that is derived from the choices of optimizing agents and recover the underlying (privately known) information about risk type and preferences. This allows for rich, out of sample, counterfactual welfare analysis. However, it requires the researcher to make critical assumptions about the nature of both the utility function and individuals' private information. These modeling choices can have non trivial effects on the welfare estimates. Moreover, they are often specific to the particular market studied, making it difficult to meaningfully compare welfare estimates across different insurance markets. Technical estimation challenges further impairs the ability of researchers to readily adapt these approaches to other insurance market, or even to other data sets in the same market.

Given these trade-offs, we see our approach as highly complementary to – rather than competitive with – these earlier papers. The trade-off is a familiar one in economics. It is somewhat analogous to the trade-offs in demand estimation between product-space approaches (e.g. the Almost Ideal Demand System of Deaton and Muelbauer, 1980; see e.g. Hausman (1997) for an application) and characteristic-space approaches (Lancaster, 1966; see e.g. Berry, Levinsohn, and Pakes (1995) for an application). The latter can evaluate welfare from new goods, while the former can only do this after these goods have been introduced. There is a similar trade-off inherent in the two standard approaches to estimating the deadweight loss from taxation; Harberger (1964) offers a local approximation of the welfare cost of a small tax change based on the estimated demand curve, while Hausman (1981) estimates the underlying primitives of the utility function and can therefore calculate the exact deadweight loss from any tax. In a similar spirit, recent work by Chetty (forthcoming) shows how the optimal level of insurance benefits can be inferred from estimates of key behavioral elasticities with respect to these benefit levels, holding other aspects of the insurance policy fixed; however, analysis of other features of the optimal insurance contract (beside the benefit level) requires specifying the underlying primitives behind the estimated behavioral elasticities (as in, e.g., Acemoglu and Shimer 2000).

The rest of the paper proceeds as follows. We begin in Section 2 by illustrating the spirit of the proposed approach with a simple, hypothetical, numerical example. Section 3 sketches the framework more systematically. It provides some graphical intuition for the efficiency costs of selection in insurance markets, and shows how this framework translates naturally into a series of estimating equations. We also discuss the data requirements for implementing this approach.

Section 4 illustrates our approach by applying it to the market for employer-provided health insurance in the United States. This is a market of substantial interest in its own right. The workplace is the primary source of private health insurance in the United States, covering about 90 percent of the privately insured non-elderly, or about 160 million Americans (Fronstin, 2003). More generally, government intervention in health insurance markets is widespread but also considerably varied. The government directly insures the elderly and indigent through Medicaid and Medicare (which constitute the largest – and fastest growing – entitlement program in the United States) and provides tax subsidies for employer-provided health insurance (constituting one of the largest federal tax expenditures, currently totalling about $126 billion, or about 12 percent of federal income tax revenue (President's Advisory Panel on Federal Tax Reform, 2005). The standard economic rationale for these various programs is as a counterweight to adverse selection pressures in private health insurance markets. The existing empirical evidence on this market, reviewed e.g. by Cutler and Zeckhauser (2000), is consistent with asymmetric information; however, the literature generally does not distinguish between selection and moral hazard, or analyze the welfare consequences of the detected market failure.[2] Whether public policy has the potential to produce welfare gains in health insurance markets with adverse selection, as well as the optimal form of such public policy, is an open empirical question.

We analyze individual-level data from a large private employer in the United States on the health insurance options, choices, and medical expenditures of its employers. We use the fact that, due to the organizational structure of the company, different employees doing similar jobs in different sections of the company face different employee premiums for purchasing more comprehensive relative to less comprehensive insurance. We verify that pricing appears random with respect to the characteristics that the managers setting employee premiums can likely observe. Using this price variation, we estimate a declining marginal cost curve, and thus detect adverse selection in this market. We estimate that in a competitive market the annual efficiency cost of this selection would be between $0.83 and $3.30 per employee, or between 0.22% and 0.88% of the total surplus at stake from efficient pricing. We estimate that the social cost of the public funds of the subsidy required to move from the adverse selection equilibrium to the efficient equilibrium is about an order of magnitude higher than these estimates of the welfare gain from achieving the efficient allocation.

In Section 5 we extend both the theoretical and empirical welfare analysis to account for moral hazard. As already mentioned, our test for selection based on the shape of the cost curve is robust to the presence of moral hazard; we consider this ability to test directly for selection an important contribution relative to the existing joint tests of adverse selection and moral hazard. The welfare analysis that we have just described, however, abstracted from the possibility of moral hazard. We

[2] Cutler and Reber (1998) is an important exception. They document the adverse selection induced by charging employees more on the margin for more comprehensive coverage, and estimate the welfare consequences. The general spirit of our approach has much in common with theirs. An important distinction however is that we estimate the cost curve as well as the demand curve; estimation of the cost curve is crucial for welfare analysis as it enables calculation of the (counterfactual) efficient price of insurance.

show that this proposed welfare analysis of the efficiency cost of selection becomes a lower bound on the welfare cost of adverse selection when we allow for moral hazard. With no additional data elements, we can also derive an upper bound for the welfare cost of adverse selection.[3] In our application, we calculate the upper bound on the annual welfare cost of adverse selection to be between *[results TBA]* per employee, or between *[results TBA]* of the total surplus at stake from efficient pricing. Combining these with our previous estimates – which constitute a lower bound on the welfare cost of selection in the presence of moral hazard – we conclude that the welfare cost of adverse selection in the particular choice of contracts in between *[results TBA]* of the surplus from efficient pricing, and that the social cost of the subsidy needed to correct this market failure is – at either bound – about an order of magnitude higher than the social gain from doing so.

The last section concludes by discussing a wide range of settings in which the approach we propose could be possibly applied. We view this as a promising direction for further work.

2 An example

We start by illustrating the spirit of our approach to empirical welfare analysis in selection markets with a simple example. Consider a population of individuals making a binary choice of whether to fully insure or not to insure at all. Each individual is characterized by two parameters: his willingness to pay for insurance π and his expected costs to the insurer c. Suppose individuals are uniformly drawn from a discrete distribution of three types, such that $(\pi, c) \in \{(2,1),(4,3),(6,5)\}$. Note that these types exhibit adverse selection in the sense that individuals who value insurance more are expected to cost more to the insurance company. The competitive (i.e. zero profit) equilibrium price would be $p = 4$, at which price is equal to average costs. Because $\pi > c$ for all types, an efficient allocation requires that everyone purchases insurance, or that $p \leq 2$. That is, the adverse selection results in under-provision of insurance, as pointed out in the seminal paper by Akerlof (1970).

Of course, the econometrician does not directly observe an individual's willingness to pay (π), or his individual-specific cost c. However, these can be recovered, and welfare analysis performed, if there exists data on the fraction insured and the average costs of the insured at different (exogenously generated) insurance prices. For example, consider data on insurance coverage and costs for three different prices of $p = 2, 4, 6$. Given the assumptions above, the data available to the econometrician would consist of $(p, Q, AC) = \left\{(2,1,3), \left(4, \frac{2}{3}, 4\right), \left(6, \frac{1}{3}, 5\right)\right\}$, where AC is the average costs of the insured. For example, the case of $p = 4$ will result in insurer share of $\frac{2}{3}$ (individuals with $\pi = 4$ or $\pi = 6$ will purchase, but individuals with $\pi = 2$ will not), and average costs of those who purchase insurance of $\frac{3+5}{2} = 4$. Similarly the case of $p = 2$ will result in insurer's market share of 1 and average costs of 3, and the case of $p = 6$ will result in insurer share of $\frac{1}{3}$ and average costs of 5.

Using these data, and in particular assuming that prices are exogenous with respect to both

[3] With additional information - for example exogenous variation in the amount of coverage offered in addition to its price – we can obtain an exact estimate of the welfare cost, in the presence of moral hazard.

demand and insurable costs, we can immediately see that the competitive equilibrium price (i.e. where price is equal to average cost) is $p = 4$. We can also back out the cost c of the marginal individual whose allocation is affected when the price moves. For example, when the price changes from $p = 2$ to $p = 4$ the marginal costs are given by $\frac{AC(p{=}2)Q(p{=}2)-AC(p{=}4)Q(p{=}4)}{Q(p{=}2)-Q(p{=}4)} = \frac{3\cdot 1-4\cdot\frac{2}{3}}{1-\frac{2}{3}} = 1$. Likewise, the willingness to pay π for the marginal individual is equal to the price, 2,[4] and the mass of such individuals is equal to the change in market share associated with this price change: $Q(p = 2) - Q(p = 4) = 1 - \frac{2}{3} = \frac{1}{3}$.Using such estimates of the expected cost and willingness to pay for insurance of the marginal individual, we can now compute total surplus for any given price. For example, we can conclude that it is inefficient for the marginal individual at $p = 2$ to not have insurance, that each such individual would gain a surplus (i.e. $\pi - c$) of $2 - 1 = 1$, and that there is a mass of $\frac{1}{3}$ such individuals in the market. Thus, the efficiency cost of adverse selection in such a market would be $\frac{1}{3}$ per market participant.

Consider now the possibility of moral hazard, which we use to mean that the expected total (insurable) costs change in response to coverage. In such a case, each type would be associated with three parameters: willingness to pay for insurance π, expected costs to the insurer c_H if covered, and expected insurable costs c_L if the individual is not covered. Moral hazard implies that behavior changes in response to insurance, so that $c_H > c_L$.[5] The surplus from insurance is now bounded between $\pi - c_H$ and $\pi - c_L$, depending on the utility costs associated with adjusting behavior.

The data described above allow us to back out π and c_H . The surplus computed earlier (i.e. $\pi - c_H$) is therefore a lower bound on the surplus from insurance (and thus on the welfare cost of under-insurance from adverse selection). To preview the intuition for this result (which we discuss in considerably more detail in Section 5 below), consider the limiting case in which individuals get no utility from the increased expected costs associated with higher insurance coverage; then $c_H - c_L$ is entirely social waste and the surplus from insurance is exactly $\pi - c_H$. With data on insurable costs of the uninsured we can repeat a similar exercise to obtain an estimate of c_L of the marginal individual. This will provide an upper bound on the surplus from insurance. Again, this can be seen by considering the limiting case in which the utility gain from the increased expected costs associated with higher coverage is equal to the increase in expected costs $c_H - c_L$, so the surplus from insurance is exactly $\pi - c_L$.

For ease of exposition, in what follows we first present the framework and empirical application assuming no moral hazard. As just explained, the resultant welfare estimate of the cost of adverse selection will be a lower bound if there is moral hazard. In Section 5 we then extend both the theory and the empirical application to allow for moral hazard, and show how to derive an upper

[4] This is not completely precise. Given the example, all we would know is that the willingness to pay by the marginal guy is $\pi < 4$. We would know that $\pi = 2$ with more continuous variation in price, or if we knew that the support of the willingness-to-pay distribution is 2, 4, and 6.

[5] Note that this cost difference arises due to behavioral differences under the two contracts. Both c_H and c_L are calculated under the assumption that the individual is covered by H. c_H denotes the cost to the insurer of covering the individual with H if he behaves as if he had the H contract, while c_L denotes the cost to the insurer of covering the individual with H if he behaves as if he had the L contract.

bound.

3 Theoretical framework

3.1 Preliminaries

Setup and notation We consider a situation where there is a given population of individuals, who are allowed to choose from exactly two insurance contracts available, one that offers high coverage (denoted by H) and one that offers less coverage (denoted by L). It is conceptually straightforward to extend the analysis to more contracts, but substantially complicates the graphical illustrations. To further simplify the exposition, we assume that contract L is no insurance and is available for free, and that contract H is full insurance; these are merely normalizations. We take the characteristics of the insurance contracts as given, although allow the price of insurance to be determined endogenously. This seems a reasonable characterization of many insurance markets; it is often the case that the same set of contracts are offered to observably different individuals, with variation across individuals only in pricing, and not in offered coverage. Our analysis is therefore in the spirit of Akerlof (1970) rather than Rothschild and Stiglitz (1976), who endogenize the level of coverage as well.

Define the population by a distribution $G(\zeta)$, where ζ stands for a vector of consumer characteristics. A key characteristic of the analysis is that we do not need to specify the nature of ζ; it could describe multi-dimensional risk factors, information about risk type, and/or preferences. Denote the (relative) price of contract H by p, and denote by $v_H(\zeta_i, p)$ and $v_L(\zeta_i)$ consumer i's (with characteristics ζ_i) expected utility from buying coverages H and L, respectively. Naturally, assume that $v_H(\zeta_i, p)$ is strictly decreasing in p, and that $v_H(\zeta_i, p = 0) > v_L(\zeta_i)$.

For now we ignore moral hazard and denote the expected monetary cost associated with the insurable risk for individual i by $c(\zeta_i)$; we thus assume that costs do not depend on coverage choice. This allows us to simplify the presentation of the analysis. However, it is an important assumption that is likely to be unrealistic in many contexts. In Section 5 we therefore extend the model and the empirical results to accommodate the presence of moral hazard. As can be anticipated based on the simple example we worked through in Section 2, the estimates of the efficiency cost of adverse selection that we obtain here by abstracting from moral hazard are a lower bound for the efficiency cost of adverse selection when moral hazard is accounted for. In Section 5 we discuss how to obtain an upper bound.

Demand for insurance We assume that each individual makes a discrete choice of whether to buy insurance or not. Since we take as given that there are only two available contracts and their associated coverages, demand is only a function of the (relative) price p. We assume that firms cannot offer different prices to different individuals. To the extent that firms can make prices contingent on observed characteristics, one should think of our analysis as applied to a set of individuals that only vary in unobserved (or unpriced) characteristics. We assume that if individuals choose to buy insurance they buy it at the lowest price it is available, so it is sufficient to characterize demand for insurance as a function of the lowest price p.

Given the above assumptions, individual i chooses to buy insurance if and only if $v_H(\zeta_i, p) \geq v_L(\zeta_i)$. Define $\pi(\zeta_i) \equiv \max\{p : v_H(\zeta_i, p) \geq v_L(\zeta_i)\}$. That is, $\pi(\zeta_i)$ is the highest price of insurance at which individual i is willing to buy insurance. Aggregate demand for insurance is therefore given by

$$D(p) = \int 1\left(\pi(\zeta) \geq p\right) dG(\zeta) = \Pr\left(\pi(\zeta_i) \geq p\right), \tag{1}$$

and assume that primitives imply that $D(p)$ is strictly decreasing, continuous, and differentiable.

Supply and equilibrium We assume that there are $N \geq 2$ identical risk neutral insurance providers, who set prices in a Nash Equilibrium (a-la Bertrand). We assume that when multiple firms set the same price, individuals who decide to purchase insurance at this price choose each firm randomly. We also assume that there is a loading factor in providing insurance, which we denote by a fixed cost F for each insurance contract sold.[6] This loading factor can be thought of as arising from transaction costs of handling and processing insurance claims, or of attracting customers and maintaining the company more generally. It is not essential for any of the welfare analysis but, as we shall see, allows for the (interesting and presumably practically relevant) possibility that it may not be socially efficient for all individuals to have insurance, even if they are all risk averse.

The foregoing assumptions imply that the average (expected) cost curve in the market is given by

$$AC(p) = \frac{1}{D(p)} \int (F + c(\zeta)) 1\left(\pi(\zeta) \geq p\right) dG(\zeta) = F + E\left(c(\zeta) | \pi(\zeta) \geq p\right), \tag{2}$$

and the marginal (expected) cost curve[7] in the market is given by

$$MC(p) = F + E\left(c(\zeta) | \pi(\zeta) = p\right). \tag{3}$$

It is easy to verify that the above assumptions guarantee the existence and uniqueness of equilibrium.[8] In particular, the equilibrium is characterized by the lowest price that implies zero profits, that is:

$$p^* = \min\{p : p = AC(p)\}. \tag{4}$$

3.2 Welfare analysis

Measuring welfare We measure consumer surplus by the certainty equivalent. The certainty equivalent of an uncertain outcome is the amount that would make an individual indifferent between

[6] When the lower coverage isn't "no coverage," one should think of F as the (fixed) incremental costs for providing the higher coverage.

[7] Note that there could be multiple marginal consumers. Because price is the only way to screen in our setup, all these consumers will together average to form the marginal cost curve.

[8] This is a similar result to the "buyers' equilibrium" in the (richer and more complex) setting analyzed by Wilson (1980).

obtaining this amount for sure and obtaining the uncertain outcome. An outcome with a higher certainty equivalent therefore provides higher utility to the individual. This welfare measure is attractive as it can be measured in monetary units. Total surplus in the market is the sum of certainty equivalents for consumers and profits of firms. As is standard in partial equilibrium applied welfare analysis, we ignore income effects associated with price changes; equivalently, we assume that utility is quasi-linear in all other goods (which may be a reasonable approximation for insurance premiums, which tend to be a small fraction of the individual's income).

Denote by $ce_H(\zeta_i)$ and $ce_L(\zeta_i)$ the certainty equivalent of consumer i from an allocation of contract H and L, respectively; under the assumption that all individuals are risk averse, the willingness to pay for insurance is given by $\pi(\zeta_i) = ce_H(\zeta_i) - ce_L(\zeta_i) > 0$. We can write consumer welfare as

$$CS = \int \left[(ce_H(\zeta) - p) 1 (\pi(\zeta) \geq p) + ce_L(\zeta) 1 (\pi(\zeta) < p) \right] dG(\zeta) \tag{5}$$

and producer welfare as

$$PS = \int (p - c(\zeta) - F) 1 (\pi(\zeta) \geq p) \, dG(\zeta). \tag{6}$$

Total welfare will then be given by

$$TS = CS + PS = \int \left[(ce_H(\zeta) - c(\zeta) - F) 1 (\pi(\zeta) \geq p) + ce_L(\zeta) 1 (\pi(\zeta) < p) \right] dG(\zeta). \tag{7}$$

It is now easy to see that it is socially efficient for individual i to purchase insurance if and only if

$$ce_H(\zeta_i) - c(\zeta_i) - F \geq ce_L(\zeta_i) \tag{8}$$

In other words, it is socially efficient for individual i to purchase insurance only if his willingness to pay $\pi(\zeta_i) = ce_H(\zeta_i) - ce_L(\zeta_i)$ is at least as great as the expected social cost of providing the insurance, $c(\zeta_i) + F$.

Graphical illustration of welfare cost of adverse selection Figure 1 and Figure 2 provide a graphical analysis of two canonical cases. The relative price (or cost) of contract H is on the vertical axis. Quantity (i.e. share of individuals in the market with contract H) is on the horizontal axis; the maximum possible quantity (i.e. 1) is denoted by Q_{max}. The demand curve denotes the relative demand for the H contract. Likewise, the average cost (AC) curve and marginal cost (MC) curve denote the average and marginal incremental costs (i.e. expected claims) to the insurer from the H contract relative to coverage with the L contract (which we have normalized to 0).

Figure 1 describes a case of adverse selection. The key feature of adverse selection is that the individuals who value insurance the most (i.e. have the highest willingness to pay) are those who have the highest expected costs. This is equivalent to a declining MC curve; as the price falls, individuals with lower willingness to pay are brought into the market, and bring down average costs. The essence of the private information problem is that firms cannot charge individuals based on their (privately known) marginal cost, but are instead restricted to charge a uniform price, which in equilibrium implies average cost pricing. Since average costs are always higher than marginal costs, adverse selection creates under-insurance, a familiar result first pointed out by Akerlof (1970). This under-insurance is illustrated in Figure 1. The equilibrium share of individuals who buy contract H is at Q_{eqm} (i.e. where the AC curve intersects with the demand curve), while the efficient number of insurance buyers (defined where the MC curve intersects with the demand curve) is at $Q_{eff} > Q_{eqm}$.[9]

The welfare loss due to adverse selection is represented by the shaded region CDE in Figure 1; this represents the lost consumer surplus from individuals who are not insured in equilibrium (because their willingness to pay is less than the average cost of the insured population) but whom it would be efficient to insure (because their willingness to pay exceeds their own marginal cost). One could similarly evaluate the welfare consequence of mandatory social insurance. Mandating everyone to buy H generates welfare equal to the area ABE minus the area EGH. This can be compared to welfare at the competitive equilibrium (area ABCD), welfare under the first best (area ABE), welfare from mandating everyone to buy L (normalized to zero), or the net welfare gain of the subsidy required to achieve the efficient equilibrium. The relative welfare rankings of these alternatives is an open empirical question. A primary purpose of the proposed framework is the develop an empirical approach to assessing welfare under alternative potential policy interventions (including the no intervention option).

Advantageous selection The original theory of selection in insurance markets emphasized the possibility of adverse selection, and the resultant efficiency loss from under-insurance (Akerlof, 1970; Rothschild and Stiglitz, 1976). Consistent with this theory, the empirical evidence points to several insurance markets, including health insurance and annuities, in which the insured have higher average costs than the uninsured. However, a growing body of empirical evidence suggests that in many other insurance markets, including life insurance and long-term care insurance, there exists "advantageous selection"; those with more insurance have lower average costs than those with less or no insurance. Cutler, Finkelstein, and McGarry (2008) provide a review of the evidence of adverse and advantageous selection in different insurance markets.

Figure 2 therefore describes the case of advantageous selection. In contrast to adverse selection, with advantageous selection the individuals who value insurance the most are those who have the least expected claims. This translates to upward sloping MC and AC curves. Once again, the source of market inefficiency is that consumers vary in their marginal cost, but firms are restricted to uniform pricing, and in equilibrium price based on average cost. However, with advantageous selection the resultant market failure is one of over-insurance rather than under-insurance ($Q_{eff} < Q_{eqm}$; see, Figure 2 or e.g., de Meza and Webb, 2001). Intuitively, insurance providers have an additional incentive to reduce price, as the infra-marginal customers whom they acquire as a result are relatively good risks. The resultant welfare loss is given by the shaded area CDE, and represents the excess of MC over willingness to pay for individuals whose willingness to

[9] Note that because of the assumption of a loading factor $F > 0$, it is possible for $Q_{eff} < Q_{max}$ (as drawn).

pay exceeds the average costs of the insured population. Once again, we can also easily evaluate welfare of different situations in Figure 2: mandating insurance (the area ABE minus the area EFH), mandating no insurance (normalized to zero), competitive equilibrium (ABE minus CDE), and first best (ABE).

These graphical analyses illustrate that the demand and cost curves are sufficient statistics for welfare analysis of equilibrium and non-equilibrium pricing of the existing contracts. This in turn is the essence of our empirical approach. We estimate the demand and cost curves, but remain agnostic about the underlying preferences that determine the demand curve and the underlying nature of the individuals' information or behavior that gives rise to the cost curves. As long as individuals' revealed choices can be used for welfare analysis, the precise *source* of the selection (i.e. the ζ) is not germane for analyzing the efficiency consequences of the resultant selection, or the welfare consequences of public policies that change the equilibrium price (e.g., by mandating or subsidizing a particular policy). In Appendix A we provide a specific example that illustrates this point. The source of selection – – for example, whether it is driven by unobserved preferences for insurance such as risk aversion, or by heterogeneity among individuals as to how much they know about their risks – may of course, be of independent interest if, for example, one of the counterfactuals stipulates changing the information structure.[10]

3.3 Estimation

Applying our framework requires data that allows estimation of the demand curve $D(p)$ and the average cost curve $AC(p)$. The marginal cost curve can be directly backed out from these two curves and does not require further estimation. To see this, note that

$$MC(p) = \frac{\partial TC(p)}{\partial D(p)} = \frac{\partial \left(AC(p) \cdot D(p)\right)}{\partial D(p)} = \left(\frac{\partial D(p)}{\partial p}\right)^{-1} \frac{\partial \left(AC(p) \cdot D(p)\right)}{\partial p} \tag{9}$$

With these three curves – $D(p)$, $AC(p)$, and $MC(p)$ – in hand, we can straightforwardly compute welfare of various allocations as described above.

As is standard, estimating the demand curve requires data on prices and quantities (i.e. insurance coverage), and price variation that is exogenous to demand which can be used to trace out the demand curve. To estimate the $AC(p)$ curve we need, in addition, data on the expected costs of insurees, such as data on subsequent risk realization and how it translates to insurer costs. With such data we can then use the very same variation in prices to trace out the $AC(p)$ curve. That is, we do not require a separate source of variation. Note that the $AC(p)$ curve estimates the average costs of those who (endogenously) choose insurance at a given price.

[10] Recent evidence suggests that the underlying source of advantageous selection may, in fact, differ across insurance markets. Fang, Keane, and Silverman (2006) suggest that differences in expected costs and in demand across individuals with different cognitive ability is the primary source of the advantageous selection they document in the Medigap market. By contrast, Finkelstein and McGarry (2006) suggest that risk aversion may be an important source of the advantageous selection they document in long-term care insurance.

As mentioned, the basic framework made a number of simplifying assumptions for expositional purposes which do not limit the ability to apply this approach more broadly. It is straightforward to apply the approach to the case where the high coverage contract provides less than full coverage and/or where the low coverage contract provides some coverage; in such settings (which includes our application below) we must simply be clear that the cost curve of interest is derived from the *difference* in average costs to the insurance company associated with H coverage rather than with L. Likewise, while it was simpler to show the analysis graphically with only two coverage choices, estimation with more than two coverage choices is straightforward, and does not require the policies to be vertically "rankable." The data requirements would simply extend to having price, quantity, and costs for each contract, as well as pricing variation across all relevant relative prices so that the full demand system can be estimated**.**

A Direct Test of Selection This framework provides a direct test of selection based on the slope of the estimated marginal cost curve. A rejection of the null hypothesis of a constant marginal cost curve (i.e. slope of zero) allows us to reject the null of no selection. Moreover, the sign of the slope of the estimated cost curve informs us of the nature of any selection; a downward sloping cost curve indicates adverse selection, while an upward sloping curve indicates advantageous selection.[11]

A very nice property of this selection test is that it allows a distinct test for selection that is not affected by the existence of moral hazard (or lack thereof). By contrast, the influential and widely used bivariate probit test (and other variants; see, e.g., Chiappori and Salanie, 2000), which compares realized risks of individuals with different insurance contracts, jointly tests for the existence of either selection or moral hazard (but not each separately). Exogenous pricing variation – which is not required for the bivariate probit test – is the key to a distinct test for selection. It allows us to analyze how the risk characteristics of the sample who selects a given insurance contract varies as we vary the price of that contract.

Counterfactual Welfare Analysis The key for any counterfactual analysis that uses the approach we propose is that insurance contracts are taken as given, and only their prices vary. More structure would be required if we were to analyze the welfare effects of introducing insurance contracts not observed in the data. Thus, for example, the estimates can be used to analyze the effect of a wide variety of standard government interventions in insurance markets which change the price of insurance. These include mandatory insurance coverage, taxes and subsidies for insurance, and regulations of private insurance markets such as minimum standards for offered products, regulation of the allowable price level, or regulation of allowable pricing differences across observably different individuals. However, analysis of the welfare consequences of introducing new products

[11] Conceptually, adverse selection refers to a monotonically declining marginal cost curve, and advantageous selection to a monotonically increasing marginal cost curve. In practice, most empirical tests of selection look globally at average costs under different insurance contracts rather than locally at the marginal costs for the marginal market participant (see, e.g., Finkelstein and Poterba (2004) for a case of adverse selection, or Fang, Keane, and Silverman (2006) for a case of advantageous selection). As long as the marginal cost curve is monotone, the inferences are valid.

(i.e.., products with coverage levels not observed in the data) is not feasible.

A key strength of our approach to estimating welfare is its simplicity and transparency. It relies on estimating the demand and average cost curves, from which everything else can be derived and quantified. With sufficient price variation, no functional form assumptions are needed for the prices to trace out the demand and average cost curves. For example, if the main objective is to estimate the efficiency cost of selection, then price variation that spans the range between the market equilibrium price (point C in Figures 1 and 2) and the efficient price (point E in figures 1 and 2) allows us to estimate the welfare cost of selection (area CDE) non-parametrically (that is, without any functional form assumptions regarding the shape of the demand or average cost curves). With pricing variation that does not span these points, the area CDE can still be estimated, but this will require some functional form assumption (such as, for example, linear demand and cost curves, as in our baseline specification below). Of course, as long as the only "good" variation is in prices, any other model, which specifies individuals' utilities directly, will also be implicitly imposing functional form assumptions on the demand and cost curves, which are the objects on which the efficiency costs rely. To the extent that such models could use theory as a guidance for a more plausible functional form of the demand and cost curves, we could also use this guidance for our approach in specifying more plausible functional forms of these curves, but without fully estimating the underlying utility parameters that give rise to them.

It is also worthwhile to observe that we could make some progress toward estimating the efficiency cost of selection with fewer data requirements. We use Figure 1 (adverse selection) for this discussion (it is easy to imagine an analogous discussion which uses Figure 2). Suppose we observe only the relative price of insurance. If we are willing to assume that the price we observe is the competitive equilibrium price P_{eqm}, we can obtain a (presumably not very tight) upper bound of the welfare cost of selection, given by $P_{eqm}Q_{max}$ (the rectangle IJKO in the figure).[12] If we also observe the market share of contract L, denoted $(Q_{max} - Q_{eqm})$, this upper bound can be tightened to $P_{eqm}(Q_{max} - Q_{eqm})$ (the rectangle CJKL in the figure). Finally, if we also have data on the average insurable costs of the individuals choosing contract L, denoted AC_L, we can further tighten up the upper bound to be $(P_{eqm} - AC_L)(Q_{max} - Q_{eqm})$ (which is equal to the area CJGD in the figure). Anything further will probably require price variation, which provides more information about the marginal cost and marginal willingness-to-pay for individuals currently not covered by H.

4 Empirical application: Employer-provided health insurance

4.1 Data and environment

We use detailed, individual-level data on the U.S.-based workers (and their dependents) at a large multinational producer of aluminium and related products. In 2004, the primary year of our

[12] This upper bound is what we used in Einav, Finkelstein, and Schrimpf (2007) to define the Maximum Money at Stake (MMS) concept, as a way to quantify the relevant size of an insurance market.

analysis, the company had approximately 45,000 active employees in the U.S. working at about 300 different plant or office locations in 39 different states.[13]

The data contain detailed information on the menu of health insurance options available to each employee, the premium associated with each option, employees' health insurance choices from these menus, and their subsequent medical insurance claims.[14] Crucially, the data also contain plausibly exogenous variation in the prices of the insurance contracts offered to otherwise similar individuals within the company. As discussed above, such pricing variation is crucial for implementing our approach; we therefore defer a detailed discussion of the nature and source of this variation to the next section.

Finally, the data contain rich demographic information. We observe the employees' human resources record and income as reported on her tax filing, as well as information from their health insurance filings on the number and ages of other insured individuals. The resulting data include age, race, gender, annual salary, job tenure at the company, and information on the nature of the employee's job, as well as the number and ages of insured dependents. We observe this information both contemporaneously and for several previous years (if the employee was employed at the company).

As a result, we suspect that we observe virtually everything about the employees that the administrators setting insurance prices can observe without direct personal contact, and perhaps some characteristics that they might not be able to observe (such as detailed health care utilization and expenditure information from previous years). This is important because it allows us to examine whether the variation in contract pricing across individuals appears random (at least with respect to these observables).

Table 1 provides some descriptive statistics on the employees. We make a number of sample restrictions for purposes of data purity, which brings the original sample of about 45,000 active workers down to about 37,000 active workers. Column (1) presents descriptive statistics for this entire sample, and columns (2)-(5) present statistics for smaller cuts of the data that are relevant for the analysis below. Approximately one-third of the employees are salaried (as opposed to hourly workers). The vast majority (over 95 percent) of salaried workers were offered a new set of health insurance options in 2004, which are the options we study below. By contrast, only about half of the hourly workers were offered the new benefits in 2004, as many of them were under union contracts that specified benefits for the entire contract duration. For this reason, and because (as discussed below) the pricing variation is cleaner for the salaried workers, we limit our analysis to these employees. For comparison, column (6) presents statistics from the 2005 census on characteristics of employees in the United States, and column (7) repeats the exercise for those employees who work in manufacturing industries.

[13] Some of these locations are quite small. There are only 46 plants in 22 states with 250 or more active employees.

[14] We have detailed data at the individual claim level, but for the purpose of this paper we largely aggregate claims by a given employee over the entire 2004 calendar year, which is the coverage period.

4.2 Empirical strategy and relationship with the theoretical framework

In 2004, in an effort to control health care spending, the company introduced a new set of health insurance options to virtually all its salaried employees and about one-half of its hourly employees.[15] Our primary empirical exercise examines the choice between two possible levels of PPO coverages in 2004. Following our earlier nomenclature, we will use the notation H and L to reflect the High and Low coverage, respectively. These two coverage plans vary only in their consumer cost sharing rules (specifically, the level of the deductible and the out-of-pocket maximum). Although employees could choose from among several other options (in particular, an HMO, a no coverage option, and a high deductible option with a tax preferred savings element), we focus on these two most frequent choices, which approximately three-fifths of the salaried employees chose. Columns 3 and 4 of Table 1 compare the demographic characteristics of the subset of employees choosing the two most common options to the overall sample; they look quite similar. We show later that the pricing variation in the relative price of the High and Low coverage options does not affect the probability the employee chooses one of these "outside" options, so that we are not unduly concerned about sample selection.

The individuals in our data selected their coverage during the open enrollment period at the end of 2003. Employees who did not make an active choice from the new benefit options were defaulted into the Low coverage option.[16] Given the evidence of the important role of defaults in affecting benefit selection (e.g., Madrian and Shea, 2001), it seems likely that having the Low coverage as the default option may affect our demand estimates below. We note, however, that this is not a problem for our analysis per se, but is just another example of why our welfare analysis is specific to its context (including the particular options offered, the workers in the sample, and other features of the institutional environment such as the defaults).[17] We also suspect that default may be less important in our setting than in others. This is because, as mentioned, 2004 was the first year in which the new benefits were offered. These new benefits came with much effort by the company to

[15] Over the subsequent several years, most of the remaining hourly employees were transitioned to the new health insurance options as their union contracts expired. Consistent with the goal of reducing health spending, the new set of PPO contract choices contained plans with higher consumer cost sharing than the old set of PPO options. This variation in the contracts offered is not well suited to the approach developed here, which relies on variation in the pricing of the same set of contract offerings. Busch et al. (2006) study the effect of the change in plan options between 2003 and 2004 on the use of preventive care. Einav, Finkelstein, McKnight, and Cullen (in progress), use the staggered timing across employees in the transition from one set of contract offerings to another to study the impact of consumer cost sharing on medical expenditures and health outcomes.

[16] An exception to this statement applies to the approximately 20 percent of employees already enrolled in an HMO or opted out of coverage; for such individuals, if no active choice was made, these enrollment choices were continued. Enrollment in other options could not be continued since the set of benefits available changed in 2004.

[17] Generalizing our estimate of the efficiency cost of adverse selection to a setting with some other default would require specification and modeling of the underlying features of the individual behavior that generates the default sensitivity, just as generalizing it to a setting with other contract options would require modeling of the underlying primitives behind the estimated demand curve. In both cases, it would be useful to also have "good" variation in the default option (and these other contract options), so that results are not driven only by modelling assumptions.

advertise and explain the new options to its employees, making it likely that most individuals were "active" choosers. This is also consistent with the fact that only a small fraction (about 20% of those faced with the new benefits) of the employees are in (by choice or default) the default, Low coverage.

The relative employee premium for the H contract (relative to the L contract) varied across individuals. We discuss this variation below. For each individual we observe the relative price she faced, which coverage she chose, and subsequent medical claims. Define the relative price individual i faces to be $p_i = p_i^H - p_i^L$, where p_i^j is employee i's (pre-tax) annual contribution if she chose coverage j. Let D_i be equal to 1 or 0 if individual i chose coverage H or L, respectively. Finally, let m_i be individual i's total medical spending during 2004.

In our theoretical section we defined (for simplicity) H to be full coverage and L to be no coverage. This implied that m_i was the cost to the insurance company from covering individual i. In our empirical setting, however, H is not full coverage and L still provides coverage, requiring a minor adjustment. In particular, let $c^H(m_i)$ and $c^L(m_i)$ imply the cost to the insurance company from medical expenditures totalling m_i under coverages H and L, respectively. Thus, the incremental costs to the insurance company from the higher coverage relative to the lower coverage, holding m_i constant, is given by $c_i \equiv c(m_i) = c^H(m_i) - c^L(m_i)$. The AC curve is computed by calculating the average c_i for all individuals who choose H coverage at a given relative price p, and seeing how this average varies as the relative price varies.

Figure 3 illustrates the major differences in consumer cost sharing between the two coverage options – and the construction of c_i – graphically. Figure 3(a) shows the annual out-of-pocket spending (on the vertical axis) associated with a given level of total medical spending m (on the horizontal axis) for each coverage option. We graph the rules for family coverage which are the focus of our baseline analysis; therefore total medical spending (here and in the subsequent empirical analysis) refers to medical spending of the employee and all covered dependents.[18] The more comprehensive coverage has no deductible and a 10 percent coinsurance rate up to a maximal out-of-pocket spending of $5,000. The lower coverage has a $500 dollar deductible, and a 10 percent coinsurance rate for each dollar above the deductible up to a maximal out-of-pocket spending of $5,500. Figure 3(b) shows the implied difference in out-of-pocket spending between the Low and High coverage, for a given level of annual medical spending m_i; by construction, this is equal to c_i, the incremental cost to the insurance company from the High coverage of a given level of medical spending.[19] We compute $c_i \equiv c(m_i) = c^H(m_i) - c^L(m_i)$ for each employee with High coverage.

[18] These rules are the same for the two other multi-person coverage tiers ("employee and spouse," or "employee and children"); the deductible and out-of-pocket maximum in the "employee only" coverage tier are half the amounts shown here.

[19] Figure 3 abstracts from a few details. First, we have described the cost-sharing rules for "in network" spending; the plans also specify a separate deductible, a higher (30 percent) co-insurance rate, and a separate out-of-pocket maximum for "out of network" services; these additional deductibles and out-of-pocket maximums differ between the Higher and Low coverages. We will account for this in subsequent analysis, but since these are a small fraction of overall expenditures we suspect they are unlikely to affect the results. Second, both plans (identically) specify certain expenditures that are fully covered (such as various types of preventive care); we take these "free" care categories into

Note that while $c^H(m_i)$ is observed directly in the data for individuals with High coverage, $c^L(m_i)$ must be computed counterfactually from the rules of the Low coverage plan. For consistency, we therefore calculate both $c^H(m_i)$ and $c^L(m_i)$ from plan rules. For our baseline sample (described below), we obtain a correlation between actual out-of-pocket spending and predicted out-of-pocket spending for individuals with the High coverage of over 90 percent.[20]

The nature of the differences in the plans make them particularly well suited to our analysis. The fact that they are identical in all aspects except for the consumer cost sharing means that we do not have to worry about differences in other plan features – such as the network of doctors covered – which might well differ in unobservable ways for workers in different parts of the country. Relatedly, the simple and clear plan details facilitates the calculation of (counterfactual) insurable claims under the plan not chosen which, as just discussed, is necessary for computing c_i. The nature of the plan differences is also important for understanding the margin on which we may detect selection (or moral hazard). As shown in Figure 3(b), the difference in out-of-pocket spending between the plans mainly occurs because of the difference in deductible (\$500 vs. 0); as a result, for the first \$500 of spending, the out-of-pocket difference is \$0.90 for each dollar spent. For annual medical expenditure that exceeds \$500, however, there is no difference in the marginal out-of-pocket cost of additional spending until total medical spending reaches \$50,000, which is very rarely the case (occurring for less than one percent of the sample). In terms of selection, this suggests that the differences in the plans could matter for the insurance decisions of anyone with positive expected expenditures. In terms of moral hazard, this suggests that if individuals are forward looking and have perfect foresight then differences in behavior for people covered by the different plans should be concentrated among employees who expect to spend less than \$500.

For our baselines specification, we estimate the demand and average cost functions using OLS, assuming linear (in prices) demand and cost curves. That is, we estimate the following two equations

$$D_i = \alpha + \beta p_i + \epsilon_i \tag{10}$$

$$c_i = \gamma + \delta p_i + u_i \tag{11}$$

where D_i, c_i, and p_i are defined earlier. Following the theoretical framework, the demand equation is estimated on the entire sample, while the (average) cost equation is estimated on the sample of

account in calculating the out-of-pocket expenses associated with a particular employee's annual medical expenditures under a particular plan. Finally, we have described the plans as if the cost sharing rules were a function only of the total spending by the members covered by the plan. In practice, as is typical of most health insurance plans (see e.g., Eichner, 1998; or Kowalski, 2007), the cost sharing rules in a plan that covers more than one family member can vary depending on how the spending is distributed among family members. In particular, a given individual in a family can exhaust her deductible or reach her out-of-pocket maximum either by spending the requisite amount that is required by single coverage or by having the cumulative spending of other members of the family reach the family deductible or the requisite amount of out-of-pocket maximum for non-single coverage. We account for the composition of spending within the family in generating the predicted consumer cost sharing under the different plans. In practice, however, it does not have much effect on our prediction ability or on our cost analysis.

[20] The correlation is not 100 percent because of minor but subtle issues regarding coverage, which are mentioned in the oprevious footnote.

individuals who chose the High coverage.

Using these estimates, we can construct our predicted demand and average cost curves:

$$D = \alpha + \beta p \tag{12}$$

$$AC = \gamma + \delta p \tag{13}$$

from which we can also back out – as in equation (9) – the marginal cost curve:

$$MC = \frac{1}{\beta}\left(\frac{\partial\left(\alpha+\beta p\right)\left(\gamma+\delta p\right)}{\partial p}\right) = \frac{1}{\beta}\left(\alpha\delta + \gamma\beta + 2\beta\delta p\right) = \frac{\alpha\delta}{\beta} + \gamma + 2\delta p \tag{14}$$

With these three curves in hand, we can compute all three curves as a function of D and estimate any area between them. Appendix B shows the algebra behind the calculation of the efficiency cost of adverse selection (measured by the area of triangle CDE in Figure 1) as well as additional points and areas that we use below in deriving various welfare calculations from our empirical estimates of the demand and cost curves. Essentially, the conceptual exercise is to quantify what would have been the efficiency costs of adverse selection if the non-market setting in which our data is generated is taken to a competitive market, assuming that the rest of the features (except pricing) remain the same.

4.3 Variation in prices

An essential element in the analysis is that there is variation across workers in the relative price they face for the High coverage option, and that this variation is unrelated to the workers' willingness to pay for High coverage or to her insurable costs. The business structure of the company provides a credible source of such "good" variation across different workers in the company. Each year, the company headquarters designs a set of different possible pricing menus for employee benefits. The coverage options (as opposed to their prices to the employee) are the same across all the menus. Seven different menus were available in 2004; each menu specified the price tag (i.e. premium) to the employee for various employee benefits, such as the life insurance, disability insurance, and health insurance. For our purposes, the key element of interest is the incremental premium the employee must pay for the High coverage option relative to the Low coverage option, $p = p_H - p_L$; we refer to this relative price of High coverage as the "price" in everything that follows.

There were 6 different prices in 2004 (as two menus were identical in this respect), ranging (for family coverage) from \$384 to \$659.[21] Employees always have a choice of four different "tiers" of health insurance coverage – "employee only," "employee plus spouse," "employee plus children," and "family." For any health insurance coverage option in any menu, the family price is triple the "employee only" price, 1.58 times the "employee plus children" price, and 1.43 times the "employee

[21] The price to the employee of the High coverage options was in the ballpark of \$1,500 for family coverage, although of course ranged across the different menus. The incidence of being offered a menu with a lower average price level (across different options) may well be passed on to employees in the form of lower wages (Gruber, 1994). This is one reason why it makes sense to focus the analysis on the difference in prices for the different coverage options, rather than the level of prices.

plus spouse" price. Our baseline analysis limits the sample to family coverage, which is the most common tier (slightly over half of employees); however, we also present results for all coverage tiers combined, and they are quite similar.[22] Columns 4 and 5 of Table 1 compare the demographic characteristics of workers with family coverage to workers in all four coverage tiers; not surprisingly, they differ. We assume throughout that the choice of coverage tier is unrelated to the pricing variation; this assumption is made more palatable by the fact that coverage tier options are limited by the demographic composition of the family, and that the price multiplier across coverage tiers is the same in all of the menus.

Which price menu a given employee faces is determined by the president of her business unit. The company is divided into approximately forty business units, each of which has essentially complete independence to run their business in the manner they see fit, provided that they do so ethically and safely, and at or above the company's normal rate of return; failure on these dimensions can result in the replacement of the unit's president. Business units are typically organized by functionality – such as primary metals, flexible packaging materials, rigid packaging materials, or home exterior – and are independent of geography; there are often multiple business units in the same state, and even at the same job site (i.e. plant or office location). The number of active employees in a business unit ranges from the low teens (in "government affairs") to close to 6,000 (in "primary metals"). On average (median) there are about 1,000 (500) active employees in a business unit. The business unit president may choose different price menus for workers within his unit based on their location and their employment type (salaried or hourly worker and, if hourly, which union if at all the worker is in). As a result, workers doing the same job in the same state – or even (in principle, although as we shall see not in practice in our sample) at the same location – may face different prices for their health insurance benefits due to their business unit affiliation. The business unit's president has the ultimate authority over benefit setting, although human resource directors and various other parties provide input, and in the case of unionized employees the president must negotiate the benefits with the relevant union.

A priori, it struck us as more plausible that the pricing variation across salaried workers in different business units is more likely to be exogenous – reflecting idiosyncratic characteristics of the business unit presidents rather than differences in the workers in the different business units – than the pricing variation across hourly workers. Many of the jobs that salaried workers do are quite similar across business units. Thus, for example, accountants, paralegals, administrative assistants, electrical engineers, or metallurgists working in the same state may face different prices because their benefits were chosen by the president of the "rigid packaging" business unit, rather than by the president of "primary metals." By comparison, the nature of the hourly workers' work (which often involves the operation of particular types of machinery) is more likely to differ across different units, and may depend on what the business unit is producing. For example, the work of the potroom operators stirring molton metal around in large vats in the "primary metals" business unit is likely to be different from the work of the furnace operators in the "rigid packaging" unit.

[22] The breakdown of our sample across coverage tiers is as follows: family coverage (52 percent), employee + spousal coverage (25 percent), employee only coverage (16 percent), and employee + children coverage (7 percent).

The available data are consistent with this basic intuition. Table 2 compares mean demographic characteristics of salary workers with family coverage who face different relative prices. Importantly, we observe all of the characteristics of the employee that the business unit president (or his human resource director) is likely to observe. In general, the results look quite balanced. There is no substantive or statistically significant difference in the average age, average (log) wages, the fraction of males, the fraction of whites, or in the average (log) 2003 medical spending across employees who faced different prices. The two possible exceptions to this general pattern are job tenure and log 2003 medical expenditures when restricted to employees in the most common plan in 2003 (to avoid potential differences in spending arising from moral hazard effects of different 2003 coverages); average differences in these characteristics across employees who faced different prices are marginally statistically significant.[23] A joint $F-test$ of all of the coefficients leaves us unable to reject the null that they are jointly uncorrelated to price; when we examine the five contemporaneous characteristics (age, job tenure, gender, race, and wages) we estimate an $F-stat$ of 1.80 ($p-value = 0.14$).[24] Moreover, if we include state fixed effects in the analysis – so that we allow prices to be non-randomly assigned across states and use only the *within state* differences in the prices faced by salaried workers due to differences in which business unit they are in – the value of the $F-stat$ declines to 1.11 ($p-value = 0.37$). There is still substantial variation in prices across salaried workers within state. For example, for salaried workers with family coverage, the overall standard deviation in price is \$60 and the within state standard deviation is \$48.[25] We present results below both with and without state fixed effects and show that they are quite similar.

We also present results when all four coverage tiers are pooled – again with and without state fixed effects – and find similar results to those shown in Table 2 for family coverage. Since, as discussed, prices vary by coverage tier, in any analysis that pools coverage tiers we include (de-meaned) indicator variables for the coverage tier. To try to make the price variable roughly comparable across coverage tiers we multiple the price for the 16 percent of employees with "employee only" coverage by two, to reflect that fact that the coverage (deductible and out-of-pocket maximum) on the three other coverage tiers is double.[26] The covariates appear similarly balanced in the pooled sample of salary workers in all four coverage tiers[27]. By contrast, similar analysis of covariates for hourly workers suggests statistically significant differences across employees who face different prices; the $p-value$ on the joint test of covariates is 0.10 for the baseline sample with family coverage and no

[23] The $p-value$ on each is 0.08. We should note, of course, that with seven different covariates the $p-value$ should be adjusted upward to take account of the multiple hypothesis testing, so that this is in fact a conservative estimate of the $p-value$.

[24] When we also include 2003 spending for those in the same plan as a sixth covariate (so that our sample size falls by about 25 percent) we calculate an F stat of 1.05 (p value of 0.41).

[25] While there are multiple business units within locations so that in principle there could be within-location variation in prices across salaried workers, in practice the multiple business units in the same location always chose the same pricing menu for their salaried workers.

[26] Results from pooling the three multi-person coverage tiers, which do not call for price adjustment, are quite similar.

[27] For the whole sample of all four coverage tiers, the $F-stat$ on the five contemporaneous demographic characteristics is 1.23 ($p-value = 0.32$) without state fixed effects and 1.08 ($p-value = 0.39$) with state fixed effects.

state fixed effects, and less than 0.05 for the three other specifications. As noted, this is not surprising given the institutional environment, and motivates the restriction of our analysis to salaried employees.

The only variables that should be controlled for in any of the analyses of how demand or costs vary with price are variables that are priced. For this reason, in the analysis below, we do not include controls for any demographic characteristics of the employees or their families in any of the specifications, as prices do not vary with these characteristics. The fact that individuals of, say, different incomes or different ages may have different expected medical costs – and that this may affect which plan they choose – is part of the endogenous selection we wish to study, rather than control for. By contrast, we control for the coverage tier when we pool coverage tiers – since prices do vary by coverage tier. In the same spirit, we show results with and without state fixed effects to allow for the possibility that prices - which can vary by location and therefore by state within a business unit – may vary non-randomly across states (for example, reflecting different health care utilization patterns or costs).

Table 3 shows the raw data for our key variables for the baseline sample of 3,779 salaried workers with family coverage. The relative price of the high coverage ranges from \$384 to \$659, with about three-quarters of the sample facing the lowest price. Market share of the High coverage option is generally declining with price, and ranges from 0.67 to 0.49.[28] Average costs of the (endogenously selected) sample of individuals who select High coverage is generally increasing with price, as we would expect with adverse selection. These average costs range from \$355 to \$374. Recall that cost is defined as the difference in costs to the insurer associated with a given employee's family's medical spending if those expenditures were insured under the High coverage option relative to the Low coverage option; as illustrated in Figure 3, this difference is a non-linear function of the family's medical spending and should almost never exceed \$500.[29]

4.4 Results

The first two panels of Table 4 summarize our main results from the demand and cost analysis; the corresponding welfare implications are shown in the top half of Table 5. We will return to the bottom panels of Tables 4 and 5 later, in Section 5. The four columns report results for four main alternative specifications: family coverage (3,779 workers) and all four coverage tiers (7,263 workers), with and without state fixed effects. The results are quite similar across specifications. Note that, for our welfare analysis, both the slope and the intercept of the demand and cost curves will be relevant.

[28] The description of the range of observed market shares (and the subsequent statement about the range of observed costs) ignores the data for the price that only 7 individuals face; needless to say, this data point has very little effect on our estimates below.

[29] Although Figure 3(b) suggests that the maximum possible value of the cost variable is \$500, as explained in footnote 4.2 this figure – but not our calculations – abstracts from the rules pertaning to how spending is distributed among family members. It is therefore in principle possible for the difference in cost to be as high as \$700. We calculate a cost above 500 for 12 observations out of the 2,466 in the sample with High coverage, and a cost of \$700 for one of these individuals.

The top panel of Table 4 reports results from our estimate of demand in equation (10). The coefficients of the demand curve is highly stable across specifications. We estimate that a \$100 increase in price reduces the probability that the employee chooses the higher coverage plan by 6 to 7 percentage points; these estimates are usually statistically significant at conventional levels. Relative to the average share of High coverage of about 65 percent, these results imply that a \$100 increase in price is associated with about a 10 percent decline in market share.[30]

Panel B of Table 4 reports the results from our cost estimates based on equation (11). As mentioned, as we are interested in the average cost curve from providing the High coverage, we estimate the cost equation on the sub-sample of individuals who (endogenously) choose the High coverage option. The cost estimates are also reasonably stable across specifications. The coefficient δ on relative price consistently shows that costs are rising with price. The indicates the presence of adverse selection. That is, the average cost of individuals who purchased High coverage is higher when the price is higher. In other words, when the price selects those who have, on average, higher willingness to pay for High coverage, the average costs of this group are also higher.[31]

As noted, this represents a new test for the existence of selection as well as its nature (adverse or advantageous). Our use of price variation allows a key improvement over existing tests: we provide a direct test of selection, rather than a joint test for selection or moral hazard.

Beyond detecting adverse selection, however, the estimate of the average cost curve alone does not allow us to form even an approximate guess of the associated efficiency cost of adverse selection. The point estimates suggest that a dollar increase in the relative price of the High coverage is associated with an increase in the average cost of the (endogenous) sample selecting High coverage at that price of about 4 to 9 cents. What this implies for the efficiency cost is, however, unclear, without knowledge of the demand curve. A central theme of this paper is that we can use the combined estimates from the demand curve and the cost curve to move beyond detecting selection to quantifying its efficiency cost and, relatedly, to calculating the welfare benefits from alternative public policy interventions in the market.

Figure 4 shows how to translate the empirical estimates in Table 4 into the theoretical welfare analysis shown in Figure 1. That is, Figure 4 is the empirical analog to Figure 1. It graphs our implied demand and cost curve estimates from our baseline specification in column (1) of Table 4; it also shows the marginal cost curve implied by these estimates (see equation (14)). In this

[30] This so-called "semi-elasticity" has been estimated in several other health insurance contexts. Differences in the contract choices in these different settings mean that comparisons of these semi-elasticities across different contexts are not very meaingful. Nonetheless, we note for completeness and casual interest that our estimate of a semi-elasticity (with respect toa \$100 increase in premium) around -10 is somewhat larger than the typical semi-elasticities estimated (which tend to be around -3 to -4) although by no means the highest in the literature; Chernew et al. (2007) and Bundorf et al. (2008) provide useful summaries of the existing studies.

[31] It is common in analysis of health care spending to analyze the log of spending due to the highly skewed nature of spending. We do not do so here, however, since the relevant cost curve (for pricing and for quilibrium analysis) is the average cost curve, not the average log costs. Moreover, as a practical matter our cost variable is not skewed. Since the costs we are interested in are the *difference* in costs to the insurance company for a given level of medical spending, it is (by construction) bounded from above by \$500 for most employees, and never exceeds \$700.

specification, the implied welfare cost of adverse selection (i.e., triangle CDE in Figure 1) is $1.09 per employee.[32] It should be readily apparent from the figure that, holding the cost curve constant, shifting and/or rotating the demand curve could generate very different welfare costs (that is, areas for the triangle CDE). This underscores the observation that merely estimating the slope of the cost curve (i.e., detecting adverse selection) is not by itself informative about the likely magnitude of the resultant inefficiency.

Figure 4 also provides some useful information about the fit of our estimates, and where our pricing variation is relative to the key prices of interest for welfare analysis. The circles superimposed on the figure represent the actual data points (from Table 3), with the size of each circle being proportional to the number of individuals who faced that price. The fit of the cost curve appears quite good. The fit of the demand curve is also reasonable, although the scatter of data points led us to assess the sensitivity of the results to a concave demand curve (see Section 4.5).

We estimate the equilibrium price and market share of the High coverage option (i.e., point C) to be $353 and 0.69, respectively. We estimate the efficient price and market share (i.e., point E) to be $294 and 0.74, respectively.[33] Thus, we estimate that adverse selection lowers the market share of the High coverage by 5 percentage points relative to the efficient level of coverage. The estimated equilibrium price of $353 is slightly lower than the modal (and lowest) price in the data of $384. The company therefore appears to set price reasonably close to the equilibrium price. Of course, the fact that we observe prices varying from $384 to $659 – and this is how we identify the demand and cost curves – underscores the point made earlier that to generate pricing variation we observe a market that is not in equilibrium. Our analysis of "equilibrium" pricing – like our analysis of "efficient" pricing – is therefore based on counterfactuals. By the same token, our analysis of the efficiency cost of adverse selection in this market is not an analysis of the realized efficiency cost but rather what this efficiency cost would be if, contrary to fact, these options were offered in a competitive market setting.

Our estimate of the efficiency cost of adverse selection (triangle CDE) involves estimating demand and cost over a price range ($294 to $353) that we do not observe in the data, but that is within 25 percent of the price experienced by the majority of the individuals in the data. By contrast, it is important to emphasize that our estimate of the total surplus at stake from efficient pricing (triangle ABE) – which we use below as one possible benchmark by which to scale our estimates of the efficiency cost of adverse selection – involves considerable out-of-sample extrapolation of demand and cost at prices as high as $1,350 (almost double the highest price we observe). This

[32] Appendix B provides the (straightforward) algebra used to compute the various points and areas in Figure 4 from the estimates of the intercept and slope of the demand and cost curve in Table 4, column (1).

[33] Note that the cost curve is calculated from claims data and therefore does not account for any potential loading factor. In other words, using the notation from the Section 3, we implictly assume $F = 0$. In the absence of any loading or any other factors, it should always be efficient for risk averse individuals to buy the High coverage, so that the marginal cost curve should always be below the demand curve (i.e., the effiicent quantity should be 1). Our evidence of lower demand for High coverage may reflect the fact that (as noted) the Low coverage is the default, or mispecification of the functional form of demand. We return to this latter point in Section 4.5.

suggests the need for some caution in the use of this scaling factor.

In a similar vein, Figure 4 also provides some insight into the "out of sample" properties of our linear extrapolation of the demand and cost curves, and highlights some additional theoretical restrictions which we try imposing in the robustness analysis below. Throughout the domain of possible market shares (0 to 1), average costs lie between about \$330 and \$400; this is consistent with the a priori requirement (discussed earlier) that average costs must be, by construction, between 0 and \$700 (or more likely \$500) and suggests that our out-of-sample projection of average costs may not be unreasonable. The out-of-sample fit of the demand curve is less good. Willingness to pay for higher coverage is (theoretically) bounded above at \$700 (the maximum possible out-of-pocket savings from High coverage), and (theoretically) bounded below by 0 (any rational individual should always prefer more coverage to less if the former is offered for free). Yet, as seen in Figure 4, market share for High coverage does not go to 0 until a price of \$1,350 (as opposed to a price of \$700 or lower) and instead of reaching a market share of 1 at a price of 0 (or higher), market share is 0.94 when the price is 0. In the robustness analysis below we examine the sensitivity of our findings to functional forms that impose these theoretical constraints.

Welfare The top half of Table 5 summarizes some of the welfare implications of the cost and demand estimates in Table 4. Each of the four columns reports the estimates from the specification in the corresponding column in Table 4. The spirit of the analysis behind each of the calculations should be familiar from the description of Figure 4. The exact algebra behind each estimate is given in Appendix B.

The first row calculates the efficiency cost of adverse selection (triangle CDE). Our estimates range across the specifications from \$0.83 to \$3.30 per employee. These estimates are quite similar whether we use only those with family coverage or the full sample with all four coverage tiers; they are about three times larger when state fixed effects are included than when they are not.

One useful way to gauge this estimate of the welfare cost of adverse selection is to compare the social welfare gain from efficient pricing (triangle CDE) to the social welfare cost of achieving this efficient pricing. Such a calculation provides a guide to whether there is scope for welfare improving government intervention in the form of price subsidies. To compute the optimal subsidy, rows 2 and 4 of Table 5 report, respectively, the efficient price (point E) and the equilibrium price (point C). The difference between them (multiplied by the fraction of employees who would take it up) represents the size of the subsidy per employee needed to achieve the efficient price; this ranges across our specifications from \$40 to \$77, or from 11 to 22 percent of the equilibrium price. Given standard estimates of the marginal cost of public funds of about 0.3 (e.g., Poterba, 1996), our estimates imply that the social cost of the subsidy needed to achieve the efficient allocation is between \$11 and \$25 (row 6). Thus, we estimate that the social cost of subsidizing the market to get to the efficient allocation is about an order of magnitude higher than the social gain from doing so (row 1).

Of course, a tax subsidy is only one potential policy. Another standard public policy option – mandatory coverage – is also easily evaluated using the results of Table 4. We could compare the welfare in equilibrium (area ABCD) to the welfare if everyone were mandated to have Higher coverage (triangle ABE, minus triangle EGH). Such a calculation is sufficiently out of sample (relative to where our pricing variation is) that we did not undertake it for the baseline specification. We return to this, and perform this calculation in the robustness section, where we impose theoretical restrictions on the location of the demand curve out-of-sample.

Another way to gauge the magnitude of our estimates of the welfare cost of adverse selection is to compare them to what we knew about how large this cost could have been before we started the analysis. Here we follow the various upper bounds discussed in the end of Section 3. While there we assumed that we observed only the equilibrium price, it is easy to show that these bounds are valid as long at we are willing to assume that the observed price is at or above the efficient price so that the inefficiency generated by selection is one of under-insurance (which is consistent with the adverse selection we detect in the data). We focus on family coverage only, and the thought experiment is to assume that we observe data (on price, quantity, and costs) from only one of the rows of Table 3, so there is no price variation. We report weighted averages across these rows. Following this approach, the highest upper bound on the welfare cost of adverse selection is given by the (average) price that individuals are offered for the policy; Einav, Finkelstein, and Schrimpf (2007) term this the "maximum money at stake." This price is \$414 in our data (for family coverage; see Table 3), which bounds the per-employee welfare cost of adverse selection at \$414. Since individuals have the option to buy High coverage at this price but choose not to do so, their welfare loss from being inefficiently uncovered by this option cannot exceed \$414.Viewed from this benchmark, our estimate of the efficiency cost of adverse selection (row 1 in columns (1) and (2) of Table 5) is between 0.26% and 0.80% of this "maximum money at stake." Moreover, as discussed in Section 3, we can tighten this bound by observing that in fact about two-thirds of the sample *does* buy High coverage when offered at the observed price. Since the inefficiency created by adverse selection is one of under-insurance, the "money at stake" is limited to the approximately one-third of the sample who does not buy the higher coverage. This implies an (average) upper bound of welfare cost of \$146. As a result, our estimate of the welfare cost of adverse selection rises to between 0.74% and 2.25% of the ex ante bound. Finally, we can tighten this bound further by using our information about the average costs of those who do not purchase High coverage. Since the welfare loss is only the difference between the willingness to pay and the expected cost of those who choose Low coverage, then rather than multiplying the market share of Low coverage by the price, we can multiply it by the difference between the price and their average cost. Using Table 3 column 5, this calculation implies an upper bound of \$28 per employee, making our estimate of the welfare cost of adverse selection to be between 4% and 12% of this bound. These types of calculations suggest that, at least in this case, the bounds on the welfare cost of adverse selection that we could have calculated without using exogenous variation in price to trace out the demand and cost curves would have been extremely loose relative to the actual welfare cost we calculate.

Finally, another natural benchmark is to consider the welfare loss from adverse selection as a fraction of the total surplus at stake from efficient pricing; this is given by the ratio of triangles CDE to ABE in Figure 1 (or, analogously, Figure 4). Of course, as already noted in the context of Figure 4, calculation of the area of triangle ABE involves extrapolating our estimates of the demand and cost curves considerably out of sample of the observed price variation. As a result, some caution must be exercised in interpreting this number. On the other hand, an attractive feature of this way of scaling our estimate is that, given our assumption of linear demand and cost, this ratio is invariant to assumptions made about the loading factor, F. By contrast, our estimate of the welfare loss (triangle CDE) does depend on F. Row 7 of Table 5 presents our estimate of the total surplus from efficient pricing; it ranges across the specifications from \$322 to \$378. Row 8 shows our estimate of the efficiency cost as a fraction of the total surplus – it ranges from 0.22% to 0.88%. Thus, in all specifications we estimate the welfare cost of adverse selection in this setting to be less than 1 percent of the surplus generated by efficient pricing.

4.5 Robustness

[To be completed; Section not written yet]

5 Incorporating moral hazard

The discussion thus far has abstracted away from potential moral hazard effect of insurance. This may be a reasonable abstraction for some insurance markets, such as annuities (Finkelstein and Poterba, 2004), long-term care insurance (Gruber and Grabowski, 2007), or deductible choice in auto insurance (Cohen and Einav, 2007). However, in many other interesting insurance markets – including our empirical application to health insurance – moral hazard is likely to be an economically important feature.

In this section we therefore extend both the theoretical and empirical analysis to handle insurance markets with moral hazard. When we do so, the welfare estimate of adverse selection derived in the preceding section becomes a lower bound on the welfare cost of adverse selection. With no additional data elements, we can also derive an upper bound for the welfare cost of adverse selection. With further variation, described below, we could obtain an exact estimate of the cost of adverse selection in a setting with moral hazard.

5.1 Theory

Setup and notation Conceptually, moral hazard is easily incorporated into the existing theoretical framework. We illustrate this by returning to the original framework in which we defined H to be full coverage and L to be no coverage. Recall that this framework can be easily generalized to two partial coverage contracts of different levels of comprehensiveness (as in the empirical application) as well as to more than two coverage options.

The key change once we allow for moral hazard is that the expected cost of the insurable event for individual i is now a function of her insurance coverage. It is useful therefore to define two expected monetary costs for individual i which vary with the coverage choice. We denote by $c_H(\zeta_i)$ individual i's expected costs of the insurable event when she has full coverage, and by

$c_L(\zeta_i)$ individual i's expected costs of the insurable event when she has no coverage (and therefore presumably behaves differently). Naturally, we assume throughout that higher coverage leads to more (rather than less) propensity to claim, so that $c_H(\zeta_i) \geq c_L(\zeta_i)$; if moral hazard exists this inequality will be strict, while without moral hazard $c_H(\zeta_i) = c_L(\zeta_i)$.[34] As a result, we now have two marginal cost curves, MC_H and MC_L and two corresponding average cost curves AC_H and AC_L, with MC_H and AC_H always higher than MC_L and AC_L, respectively.

We emphasize that the difference between the MC_L and MC_H curves comes only from behavioral differences under the two coverages, and not from the "mechanical" differences in what the two policies cover. These mechanical differences in coverage are already captured by the curves themselves, which reflect the different cost to the insurer from providing High coverage vs. Low coverage, holding behavior fixed. The key difference is that the MC_H curve is calculated as the costs to the insurance company of reimbursing based on the H contract if the individual behaves as if she is covered by H. In contrast, the MC_L curve is calculated as the costs to the insurance company of reimbursing based on the H contract if the individual behaves as if she is covered by contract L. If coverage does not affect behavior (i.e., if there is no moral hazard), the two curves would be the same. By the same token, the difference between the two cost curves is, by definition, the magnitude of moral hazard (from the insurer's perspective). Note therefore that our framework provides a test for moral hazard as well as the aforementioned test for adverse selection; a finding of $MC_H > MC_L$ allows us to reject the null of no moral hazard.

It is well known that this behavioral response to insurance coverage introduces an additional source of market inefficiency. However, in contrast to the selection case, the government has no comparative advantage over the private sector in ameliorating moral hazard (i.e., encouraging individuals to choose socially optimal behavior). Our analysis of the welfare cost of selection therefore takes the moral hazard environment (i.e., the effect of insurance coverage on expected costs) as given.

Welfare bounds In the presence of moral hazard, we still (as before, in section 3) use the certainty equivalent to measure consumer surplus. However, since behavior changes in response to coverage, it is useful to define the certainty equivalent of a given coverage as a function not only of the coverage but also of the behavior under that coverage. Formally, let $ce_{H|j}(\zeta_i)$ and $ce_{L|j}(\zeta_i)$ be the certainty equivalent of consumer i from an allocation of contract H and L, respectively, given her behavior (insurance utilization) under coverage $j = L, H$. Note that, as before, this welfare measure is attractive as it can be measured in monetary units, and total surplus in the market is the sum of certainty equivalents for consumers and profits of firms. Of course, $ce_{H|L}$ and $ce_{L|H}$ represent hypothetical (counterfactual) certainty equivalents (in which we calculate the certainty equivalent under one contract under the assumption that behavior under that contract would be that experienced under a different contract); but as we will see in a moment, these are useful conceptual constructs.

[34] We abstract from the possibility that different contracts could be associated with different underlying behaviors that produce the same costs.

The willingness to pay for insurance, when moral hazard exists, is now given by $\pi(\zeta_i) = ce_{H|H}(\zeta_i) - ce_{L|L}(\zeta_i) > 0$. Note that this willingness to pay is the consumer surplus from higher coverage (abstracting from the price paid). We can rewrite this expression as:

$$\pi(\zeta_i) = \Delta CS(\zeta_i) = \left(ce_{H|L}(\zeta_i) - ce_{L|L}(\zeta_i)\right) + \left(ce_{H|H}(\zeta_i) - ce_{H|L}(\zeta_i)\right). \tag{15}$$

That is, we can decompose the consumer surplus from higher coverage (i.e., willingness to pay for higher coverage) into two components. The first component is the utility benefit from High coverage, conditional on behavior not changing in response to this higher coverage. This represents the consumer surplus from High coverage when insurance contracts have no effect on behavior (i.e., no moral hazard) and is positive by definition; since we hold behavior fixed, this first component is not associated with any social cost and therefore translates to social surplus. The second component is the utility benefit from changing behavior, holding coverage fixed (at H). This component of consumer surplus will be weakly less than the associated social surplus; individuals may value the change in behavior at less than the social cost of the change in behavior since they do not pay for it (they have full coverage under H).

While we do not explicitly model individuals' choices regarding behavior, it is easy to see through revealed preference arguments that this second component of the individual's utility from High coverage can be bounded. On one hand, it cannot be negative: under coverage H individual i could have chosen to behave as if she was under coverage L, but given that she decided to behave differently, it must mean she is better off. On the other hand, this second component cannot be greater than the social cost associated with the change of behavior, $c_H(\zeta_i) - c_L(\zeta_i)$. If it were, then individual i under coverage L (i.e. when she bears the full social cost of the change in behavior) would have found it better off to increase her utilization and behave as if she is under H. That is, the consumer surplus created by allocating insurance to individual i is now bounded by

$$ce_{H|L}(\zeta_i) - ce_{L|L}(\zeta_i) \leq \pi(\zeta_i) = \Delta CS(\zeta_i) \leq ce_{H|L}(\zeta_i) - ce_{L|L}(\zeta_i) + c_H(\zeta_i) - c_L(\zeta_i) \tag{16}$$

The producer surplus[35] created by allocating insurance to individual i is given by

$$\Delta PS(\zeta_i) = -F - c_H(\zeta_i) \tag{17}$$

and the total surplus from this allocation, $\Delta TS(\zeta_i) = \Delta CS(\zeta_i) + \Delta PS(\zeta_i)$, is therefore bound by

$$ce_{H|L}(\zeta_i) - ce_{L|L}(\zeta_i) - c_H(\zeta_i) - F \leq \Delta TS(\zeta_i) \leq ce_{H|L}(\zeta_i) - ce_{L|L}(\zeta_i) - c_L(\zeta_i) - F. \tag{18}$$

Note that in the absence of moral hazard, equation (18) becomes equivalent to the analogous equation (8) we presented in Section 3.

The bounds on welfare in the presence of moral hazard in equation (18) are expressed in terms of certainty equivalents, which we do not directly observe. To translate them into objects we can estimate empirically, recall that we have (or can estimate) information on $\pi(\zeta_i)$, $c_H(\zeta_i)$, and

[35] As with our definition of consumer surplus, we abstract from price, which does not affect total surplus (since it is just a transfer between consumers and producers); this explains why our producer surplus equation is negative.

$c_L(\zeta_i)$. In one extreme (the lower bound), individuals value the additional utilization at their social cost. This implies that $\pi(\zeta_i) = \left(ce_{H|L}(\zeta_i) - ce_{L|L}(\zeta_i)\right) + (c_H(\zeta_i) - c_L(\zeta_i))$ and therefore that $\Delta TS(\zeta_i) = \pi(\zeta_i) - c_H(\zeta_i) - F$. In the other extreme (the upper bound), individuals do not value at all the additional utilization. This implies that $\pi(\zeta_i) = \left(ce_{H|L}(\zeta_i) - ce_{L|L}(\zeta_i)\right) + 0$ and $\Delta TS(\zeta_i) = \pi(\zeta_i) - c_L(\zeta_i) - F$. Note that the former implies our original measure of welfare cost (in the absence of moral hazard), which is now a lower bound. The latter is now an upper bound for the welfare cost. Thus, in the presence of moral hazard, the difference between $c_H(\zeta_i)$ and $c_L(\zeta_i)$ will help us provide bounds on the welfare cots of adverse selection. Equipped with the total surplus generated by allocating insurance to each individual i, we can now characterize the set of individuals for whom allocating efficient is efficient, and compare it to the equilibrium allocation. As we show below, this set of individuals will change as we move from one extreme (the lower bound) to the other (the upper bound).

The welfare cost of adverse selection: a graphical illustration To help provide some intuition for the foregoing conclusions, Figure 5 illustrates the extension of the analysis of adverse selection in Figure 1 to incorporate moral hazard.[36] The MC and AC curves from Figure 1 correspond to MC_H and AC_H in Figure 5. In addition, the figure presents the "new" cost curves MC_L and AC_L, as defined above.

Incorporating moral hazard doesn't change the supply-side and equilibrium analysis. These still rely on the cost curves derived from the expected cost given coverage H (i.e., MC_H and AC_H). Equilibrium is still defined by where AC_H intersects with the demand curve, as AC_H represents the actual average costs realized by the insurance company in providing coverage H. The demand curve may change, however, as now rational forward-looking individuals may not only incorporate the expected reduction in the mean and variance of out-of-pocket spending into their willingness to pay for High coverage, but also their behavioral response to such coverage. Since we do not derive the demand curve from underlying primitives, the equilibrium analysis is not affected.

However, the welfare implications can change. Because expected cost are now endogenous, the previous estimate of the welfare cost of adverse selection (i.e., area CDE in Figure 1 and in Figure 5) will now be a lower bound. We can also obtain an upper bound for the welfare cost of adverse selection by adding the difference between the MC_H and MC_L curves in the relevant region. Figure 5 illustrates this point, where the welfare cost of adverse selection is bounded between our earlier estimate (area CDE) and the entire shaded region (area CJK). Where exactly within this range the welfare costs are will depend on the utility cost to individuals from the behavioral change associated with the change of coverage. The less costly (in utility terms) it is for individuals to adjust their behavior (and hence their expected costs) downward, the closer the cost of adverse selection is to the lower bound (triangle CDE). For example, if insured individuals go to the doctor only because they don't pay for it then the additional cost (the difference between MC_H and MC_L) is a social waste. In contrast, if uninsured individuals do not see the doctor even when they are very sick,

[36] A similar analysis could be carried for the analogous advantegeous selection case of Figure 2.

then the additional cost is not associated with social loss.

Intuitively, consider the limiting case in which it is costless for individuals to change their behavior (i.e., they get almost no utility from the increased expected costs associated with higher insurance coverage). In this case, the area between the MC_H and MC_L curves is all social waste because there is no utility gain from the higher costs. In this case the efficient allocation is point E (where MC_H intersects with demand) and the area between the two marginal cost curves does not contribute to lost consumer surplus, since it is pure social waste. The welfare cost from adverse selection is therefore at its lower bound (triangle CDE). In the same vein, consider the other limiting case in which the utility benefit to the individual from the increased costs associated with coverage H is at its theoretical maximum. In this case the welfare loss from adverse selection is at its upper bound (triangle CJK). How high can the utility benefit be from the increased costs associated with coverage H? By the definition of moral hazard, we know that the utility gain from the behavior must be strictly less than the increase in expected costs, otherwise individuals would have undergone this behavior even absent the insurance. Thus the welfare gain must be (weakly) less than the difference between MC_H and MC_L. Moreover, to the right of point E (where demand crosses below the MC_H curve) we know that the welfare gain from the increased cost associated with coverage H can be bound more tightly, as the difference between the demand curve and MC_L. Individuals in this area are not willing to pay their expected cost under H, which must imply that they do not fully value (in utility terms) these incremental costs. For these individuals, the welfare gain from the increased coverage is bound by the demand curve. Thus, at the upper bound social surplus is generated by the entire region between MC_L (rather than MC_H) and the demand curve, creating the upper bound estimate of the triangle CJK. Note that the potential welfare loss arising from individuals to the left of point E represents welfare loss from under-insurance by individuals for whom it is not privately efficient to insure (demand is below MC_H) but it might be socially efficient to insure (since demand is above MC_L).

Estimation This discussion also shows how – with the exact same data elements discussed – we can obtain both lower and upper bounds for the cost of adverse selection in the presence of moral hazard. Identification of the lower bound requires exactly the same data elements as before; in other words, AC_H is the observed average costs for those who buy coverage H; the moral hazard effect is "built in" to the data. The upper bound estimate of the welfare cost of adverse selection requires that we also identify AC_L (and use it and the complement demand function – for coverage L – to back out MC_L). This is traced out from the complementary set of individuals: it is the observed average costs for those with coverage L as we vary p. To see the intuition, recall that c_H is the difference in the cost to the insurance company of insuring the individual with H rather than with L assuming the individual behaves as if she has coverage H; c_L is this same difference in the cost to the insurance company of insuring the individual with H rather than with L, but now measured assuming that the individual behaves as if she has coverage L. Thus, we estimate these two cost curves based on the same dependent variable but in different (complementary) samples; c_H is estimated off of the sample of individuals who choose H (as described and operationalized previously), while c_L is estimated off of the complement sample of individuals who choose L. By the same token, with these estimates in hand we also obtain ("for free") an estimate of the moral hazard effect of insurance – i.e., the (point-by-point) difference between MC_H and MC_L.

It is interesting to discuss whether and when one could get "inside the bounds" and obtain an estimate of the exact welfare cost of adverse selection in the presence of moral hazard. The preceding discussion makes it clear that this will require information about the utility gains (or lack thereof) to individuals from the increased insurance utilization in response to higher coverage. This is not feasible with the type of data considered here; that is, where the only "good" variation in the data is in prices, but not in the comprehensiveness of coverage. However, with "good" variation in coverage (e.g., consumer cost sharing), in addition to the price variation, an exact estimate can be obtained. Variation in the comprehensiveness of the available insurance contracts would allow us to pin down how the MC and the demand curves vary with the level of coverage, which is what is needed to obtain the exact welfare cost of adverse selection in the presence of moral hazard. Intuitively, if the utility cost of the behavioral change associated with lowering expenditures in the face of lower coverage is close to zero (so the exact welfare cost is close to the lower bound), then contracts providing anything less than full coverage will be associated with MC curves that are very close to the no-coverage MC_L curve. At the other extreme, if the utility cost of the behavioral change associated with lowering expenditures in the face of lower coverage is close to its theoretical maximum (so the exact welfare cost is close to the upper bound), contracts providing anything more than no coverage will be associated with MC curves that are very close to the full-coverage MC_H curve. Situations in between will give rise to cases where the exact welfare cost of adverse selection lies somewhere in between these two bounds, and the exact estimate could be identified by the way the MC curve changes in response to different levels of coverage. While it may be feasible in some contexts to obtain data containing variation in relative comprehensiveness as well as relative premiums of contracts (e.g., Adams, Einav, and Levin, 2007), exogenous variation in both comprehensiveness and premiums seems to us to be a sufficiently more stringent data requirement (relative to "merely" observing exogenous variation in premiums) that it seemed prudent to discuss how to bound the welfare cost of adverse selection in insurance markets with moral hazard.

5.2 Empirical application

As just discussed, the preceding estimates – based on the demand curve for higher coverage and the cost curve of those who endogenously chose the High coverage – provide a lower bound on the welfare cost of adverse selection in the presence of moral hazard. An upper bound can be obtained by estimating the demand curve for the Low coverage and the cost curve of those who endogenously chose the Low coverage. In other words, we use the same pricing variation and the same strategy in estimating the demand and cost curves, with two small modifications. First, the dependent variable in the demand equation (equation (10)) is now reversed, obtaining a value of 1 when individual i chose Low coverage (rather than High). Second, the sample for the cost curve analysis is the complement to the one used previously (i.e. those choosing Low coverage rather than those choosing High coverage). Recall that with two partial coverage contracts our dependent variable in the cost analysis is the difference in costs to the insurance company from covering the individual with H rather than with L, taking as given the individual's behavior. We therefore estimate AC_H (as before) by comparing the cost to the insurance company of covering a given individual's behavior under H relative to L for the sample of individuals who have the H policy. Likewise, we estimate AC_L by comparing the cost to the insurance company of covering a given individual's behavior under H relative to L for the sample of individuals who have the L policy (i.e. the complement to the sample used to estimate AC_H).

The bottom panel (Panel C) of Table 4 presents the estimates of AC_L. Like our estimates of AC_H in Panel B, our estimates of AC_L all indicate that average costs are rising with price, and thus indicate adverse selection in the market. Once again the family coverage and all coverage tiers estimates are similar, while the inclusion of state fixed effects in either sample increases the estimate. As mentioned, the difference between the implied MC_H and MC_L curves provides a direct estimate of moral hazard. Using our baseline estimates (column (1) of Table 4), we find that the average distance between the two curves is $**XX**, which is an effect of about **XX**% of insurer average costs.

The bottom part of Table 5 reports some of the welfare implications of these estimates. Note that all of the welfare implications reported in the top part of Table 5 (which used the estimate of demand and the estimate of AC_H from Table 4) are, in the presence of moral hazard, lower bounds for both the welfare cost of adverse selection and for the social cost of the subsidy needed to achieve the efficient allocation. By contrast, the bottom part of Table 5 uses the complement demand estimate and the estimate of AC_L from Table 4; in the presence of moral hazard, this provides an upper bound for both the welfare cost of adverse selection and the social cost of the subsidy needed to achieve the efficient allocation.

The results ... *[to be completed once we have results on this one]*

6 Conclusions

This paper proposed a simple approach to estimating welfare in insurance markets. As indicated by the title, the key to the approach is the existence of "good" price variation. Applied welfare analysis usually relies on pricing variation that allows the researcher to trace out a demand curve. The defining feature of selection markets is that costs vary endogenously as market participants respond to the price of insurance. Welfare analysis of selection markets therefore requires that we also trace out the (endogenous) cost curve, and therefore that we have pricing variation that is exogenous with respect to both the individual's demand for insurance and his insurable costs.

These two curves deliver several bangs for the buck. An estimate of the cost curve provides a direct test for whether selection exists in the market. Specifically, rejection of the null hypothesis of a constant (i.e. horizontal) marginal cost curve allows us to reject the null of no selection, while the slope of the marginal cost curve tell us whether the detected selection is adverse (downward sloping) or advantageous (upward sloping). Crucially, this cost curve test of selection is unaffected by the existence (or lack therefore) of moral hazard. This stands in marked contrast to the current "industry standard" – the widely used "bivariate probit" test of Chiappori and Salanie (2000) – which is a joint test for the presence of adverse selection and moral hazard, but cannot distinguish between them. These very different forms of private information have very different implications for public policy – in particular the government may have a comparative advantage in redressing the inefficiencies created by adverse selection, but tends not to in the case of moral hazard. Being able to detect selection distinctly is therefore of considerable importance. We showed how having pricing variation in addition to the data elements required by the bivariate probit test allows the researcher to do so.

Of course, detecting selection is only a first step. If selection is found, we want to be able to quantify its efficiency costs, and compare welfare in the selection equilibrium to what could be achieve by alternative public policies. We showed that with the demand curve as well as the cost curve in hand, such welfare analysis can proceed in a familiar and straightforward fashion.

We applied this framework in the particular context of the employer-provided health insurance in a large private employer in the United States. The business structure of the company generates plausibly exogenous variation in the price of more comprehensive relative to less comprehensive health insurance for otherwise similar salaried workers. Using this variation, we found evidence of adverse selection in the market. However, we estimated that its efficiency implications are quantitatively small in both absolute and relative terms, and are an order of magnitude lower than our estimate of the social cost of the subsidies that would correct the market failure.

Of course, our estimates rely on a subset of (salaried) workers within a single firm choosing between two very specific coverage options. It is not clear that these results are representative of other populations, other institutional environments, or in other insurance markets. Our empirical findings, however, highlights the importance of moving beyond a simple detection of informational asymmetries in insurance markets, and towards quantifying the efficiency costs of such asymmetries. As our particular findings illustrates, it is empirically possible to find markets where adverse selection is detected (potentially raising an efficiency concern), but then to realize that this concern is not, in fact, quantitatively important, or easily remediable through government policy. Whether the same is true in other markets and in which, is be an important area for future work, which we hope will apply the framework and strategy we have developed to other contexts – other populations, insurance contracts, and insurance markets.

We believe that the approach we propose in this paper is likely to be broadly applicable. It requires three essential data elements: insurance options and choices, subsequent risk realization, and exogenous variation in pricing. Researchers have already demonstrated considerable success in obtaining insurance company data on the options consumers face, as well as their coverage choices and subsequent claims experience in a wide range of different insurance markets.[37] Indeed, a nice

[37] Examples of such insurance company data in property-casualty insurance markets include Chiappori et al. (2006) and Cohen and Einav (2007) for automobile insurance and Sydnor (2006) for home owner's insurance. In health- and life-related insurance markets, examples include Finkelstein adn Poterba (2004) for annuities, Cutler and Reber (1998) and Eichner, Wise, and McCellelan (1998) for health insurance, and Finkelstein and McGarry (2006) for

feature of welfare analysis in insurance markets is that cost data are much easier to obtain than in many other product markets, since they involve information on accident occurrences or insurance claims, rather than insight into the underlying production function of the firm.

Another attractive aspect of the insurance context is that near-ubiquitous regulation provides numerous instances of the exogenous pricing variation that is essential for estimating the demand and cost curves. Changes in state regulations of private insurance markets create variation in the prices charged to different individuals at a point in time as well as over time. Tax policy is another useful source of pricing variation; for example, a large literature has documented (and used) the substantial variation across space and time in the tax subsidy for employer-provided health insurance and hence the price that individual employees face for this benefit. Beyond the myriad opportunities provided by public policy, researchers have also found useful pricing variation stemming from the idiosyncrasies of firm pricing behavior. Examples include firm experimentation with their pricing policy (e.g. Cohen and Einav 2007), discrete pricing policy changes (Adams, Einav, and Levin, 2007), out of equilibrium pricing decisions made by human resource managers for employee insurance benefits (Cutler and Reber, 1998), and the rules firms use to adjust individuals' prices in response to their prior claims experience (Israel, 2004; Abbring, Chiappori, and Pinquet, 2003). More generally, any of the standard instruments used in demand analysis more generally - such as "cost shifters" to production function for insurance (i.e. variation in F in the theoretical setting of Section 3) or exogenous changes in market competition – can serve as the requisite pricing variation for welfare analysis in insurance markets.

Of course, a key issue is the validity of the pricing instruments. This can be evaluated in specific applications. Indeed, we see the transparency of the approach as one of its key attractions. It is also relatively straightforward to implement, and fairly general. As a result, comparisons of welfare estimates obtained by this approach across different settings may be quite informative. For example, they can be used to try to assess in which settings – e.g. particular insurance markets, products, or populations – the welfare gains from government intervention are likely to be substantial (relative to the costs) and in which they may not be.

A final feature of our empirical approach to welfare analysis in insurance markets that deserves emphasis is that it does not require the researcher to make assumptions about the underlying nature of individuals' information or preferences that gives rise to the estimated demand and cost curves. As long as we are willing to use the individuals' revealed choices for welfare analysis, the precise source of the selection (for example, the role of unobserved preferences for insurance or private information about risk type) is not germane for analyzing the efficiency cost of the resultant selection, or the welfare consequences of public policies that change the equilibrium price (for example, through mandating or subsidizing a particular policy). Since such modeling assumptions are often ad hoc, and may have non trivial effects on the welfare estimates, we view the ability to avoid them as a key feature of our proposed approach. It is not, however, a costless one; it restricts our ability to analyze the welfare consequences of counterfactual policies to those that change the

long-term care insurance.

prices of existing products. Our approach is unable to shed light on the welfare consequences of introducing products that are not observed in the data. Analysis of such questions would require that we model the underlying structural primitives behind the revealed demand and cost curves; this is an important and useful complement to the empirical approach outlined in this paper.

7 References

Abbring, Jaap, Pierre-Andre Chiappori, and Jean Pinquet (2003), "Moral hazard and Dynamic Insurance Data." *Journal of the European Economic Association* 1(4): 767-820.

Acemoglu, Daron, and Robert Shimer (2000), "Productivity Gains from Unemployment Insurance," *European Economic Review* 44(7): 1195-1224.

Adams, William, Liran Einav, and Jonathan Levin (2007), "Liquidity Constraints and Imperfect Information in Subprime Lending," NBER Working Paper No. 13067.

Akerlof, George (1970), "The market for 'lemons': quality uncertainty and the market mechanism," *Quarterly Journal of Economics* 84(3), 488-500.

Berry, Steven, James Levinsohn, and Ariel Pakes (1995), "Automobile Price in Market Equilibrium," *Econometrica* 63, 841-890.

Buchmueller, Thomas C., and Paul J. Feldstein (1997), "The Effect of Price on Switching Among Health Plans," *Journal of Health Economics* 16: 231-247.

Buchmueller, Thomas C., and John DiNardo (2002), "Did Community Rating Induce an Adverse Selection Death Spiral? Evidence from New York, Pennsylvania, and Connecticut," *American Economic Review* 91(1): 280-294.

Bundorf, Kate M., Jonathan Levin, and Neale Mahoney (2008), "Pricing, Matching and Efficiency in Health Plan Choice," mimeo, Stanford University.

Busch, Susan H., Colleen L. Barry, Sally J. Vegso, Jody L. Sindelar and Mark R. Cullen. 2006. "Effects of a Cost-Sharing Exemption on Use of Preventive Services At One Large Employer." *Health Affairs* 25(6): 1529 - 1536.

Chernew, Michael, Gautam Gowrisankaran, and Dennis P. Scanlon (forthcoming), "Learning and the Value of Information: Evidence from Health Plan Reports Cards," *Journal of Econometrics*, available at http://www.u.arizona.edu/~gowrisan/pdf_papers/healthplanreportcard.pdf.

Chetty, Raj (forthcoming), "Moral Hazard vs. Liquidity and Optimal Unemployment Insurance," *Journal of Political Economy*, available at http://elsa.berkeley.edu/~chetty/papers/mh_liq_ui_jpe.pdf.

Chiappori, Pierre-Andre, Bruno Jullien, Bernard Salanie, and Francois Salanie (2006), "Asymmetric Information in Insurance: General Testable Implications," Rand Journal of Economics 37(4).

Chiappori, Pierre-Andre, and Bernard Salanie (2000), "Testing for Asymmetric Information in Insurance Markets," *Journal of Political Economy* 108(1): 56-78.

Cohen, Alma, and Liran Einav (2007), "Estimating risk preferences from deductible choice," *American Economic Review* 97(3), 745-788.

Cutler, David M., and Sarah J. Reber (1998), "Paying for Health Insurance: The Trade-Off between Competition and Adverse Selection," *Quarterly Journal of Economics* 113(2): 433-466.

Cutler, David M., Amy Finkelstein, and Kathleen McGarry (forthcoming), "Preference Heterogeneity and Insurance Markets: Explaining a Puzzle of Insurance," *American Economic Review papers and proceedings*, available at http://econ-www.mit.edu/files/2354.

Cutler, David M., and Richard J. Zeckhauser (2000), "The anatomy of health insurance," in A. J. Culyer and J. P. Newhouse eds., *Handbook of Health Economics*, edition 1, volume 1, chapter 11, 563-643, Elsevier.

De Meza, David and David C. Webb (2001), "Advantageous selection in insurance markets," *Rand Journal of Economics* 32(2), 249-262.

Deaton, Agnus, and John Muellbauer (1980), "An Almost Ideal Demand System," *American Economic Review* 70, 312-326.

Eichner, Matthew J. (1998), "Medical Expenditures and Major Risk Health Insurance," Ph.D. Dissertation, MIT.

Eichner, Matthew J., Mark B. McClellan, and David A. Wise (1998), "Insurance or Self-Insurance?: Variation, Persistence, and Individual Health Accounts," in D. Wise (eds.), *Inquiries in the Economics of Aging*, University of Chicago Press.

Einav, Liran, Amy Finkelstein, and Paul Schrimpf (2007), "The Welfare Cost of Asymmetric Information: Evidence from the U.K. Annuity Market," NBER Working Paper No. 13228.

Fang, Hanming, Michael Keane, and Dan Silverman (forthcoming), "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market," *Journal of Political Economy*, available at http://www.econ.duke.edu/%7Ehf14/WorkingPaper/advant/asrev2-all.pdf.

Finkelstein, Amy, and Kathleen McGarry (2006), "Multiple dimensions of private information: evidence from the long-term care insurance market," *American Economic Review* 96(4), 938-958.

Finkelstein, Amy, and James Poterba (2004), "Adverse Selection in Insurance Markets: Policyholder Evidence from the U.K. Annuity Market," *Journal of Political Economy* 112(1), 193-208.

Fronstin, Paul (2003), "Sources of Health Insurance and Characteristics of the Uninsured: Analysis of the March 2003 Current Population Survey," Employee Benefit Research Institute Issue Brief 264, available at http://www.ebri.org/pdf/briefspdf/EBRI_IB_10a-20061.pdf.

Grabowski, David, and Jonathan Gruber (2007), "Moral Hazard in Nursing Home Use," *Journal of Health Economics* 27: 560-577.

Gruber, Jonathan (1994), "The Incidence of Mandated Maternity Benefits," *American Economic Review* 84(3): 622-641.

Gruber, Jonathan and David Grabowski. 2007. "Moral Hazard in Nursing Home Use," *Journal of Health Economics* 26: 560-577.

Harberger, Arnold C. (1964), "Taxation, Resource Allocation, and Welfare," in J. Due, ed., *The role of direct and indirect taxes in the federal revenue system*, Princeton University Press, Princeton, New Jersey.

Hausman, Jerry (1981), "Exact Consumer's Surplus and Deadweight Loss," *American Economic Review* 71(4): 662-676.

Hausman, Jerry (1997), "Valuation of New Goods Under Perfect and Imperfect Competition," in Bresnahan and Gordon (eds.), *The Economics of New Goods*, NBER Studies in Income and

Wealth 58: 209-237.

Hosseini, Roozbeh (2007), "Adverse Selection in the Annuity Market and the Role for Social Security," mimeo, University of Minnesota.

Israel, Mark (2004), "Do We Drive More Safely When Accidents Are More Expensive? Identifying Moral Hazard from Experience Rating Schemes," unpublished mimeo, available at http://www.kellogg.northwestern.edu/faculty/israel/htm/research.html.

Jin, Ginger Z., and Alan T. Sorensen (2005), "Information and Consumer Choice: The Value of Publicized Health Plan Ratings," *Journal of Health Economics* 25(2): 248-75.

Kowalski, Amanda (2007), "Ouch, That's Expensive! Censored Quantile Instrumental Variables Estimates of the Price Elasticity of Expenditure on Medical Care," mimeo, MIT, available at http://econ-www.mit.edu/files/2142.

Lancaster, Kelvin J.(1966), "A New Approach to Consumer Theory," *Journal of Political Economy* 74(2), 132-157.

Lustig, Joshua D. (2007), "The Welfare Effects of Adverse Selection in Privatized Medicare," mimeo, Yale University.

Madrian, Brigitte C., and Dennis F. Shea (2001), "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior," *Quarterly Journal of Economics* 116(4): 1149-1187.

Poterba, James. 1996. "Government Intervention in the Markets for Education and Health Care: How and Why?" in *Individual and Social Responsibility* Victor Fuchs (ed). University of Chicago Press.

President's Advisory Panel on Federal Tax Reform (2005), "Simple, Fair and Pro-Growth: Proposals to Fix America's Tax System," U.S. Government Printing Office, Washington D.C.

Rothschild, Michael, and Joseph E. Stiglitz. 1976. "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics* 90, 630-649.

Royalty, Anne B., and Neale Solomon (1999), "Health Plan Choice: Price Elasticities in a Managed Competition Setting," *Journal of Human Resources* 34: 1-41.

Sydnor, Justin (2006), "Sweating the Small Stuff: The Demand for Low Deductibles in Homeowners Insurance," mimeo, Case Western University, available at http://www.kellogg.northwestern.edu/finance/fac

Wilson, Charles (1980), "The Nature of Equilibrium in Markets with Adverse Selection," *Bell Journal of Economics* 11(1): 108-130.

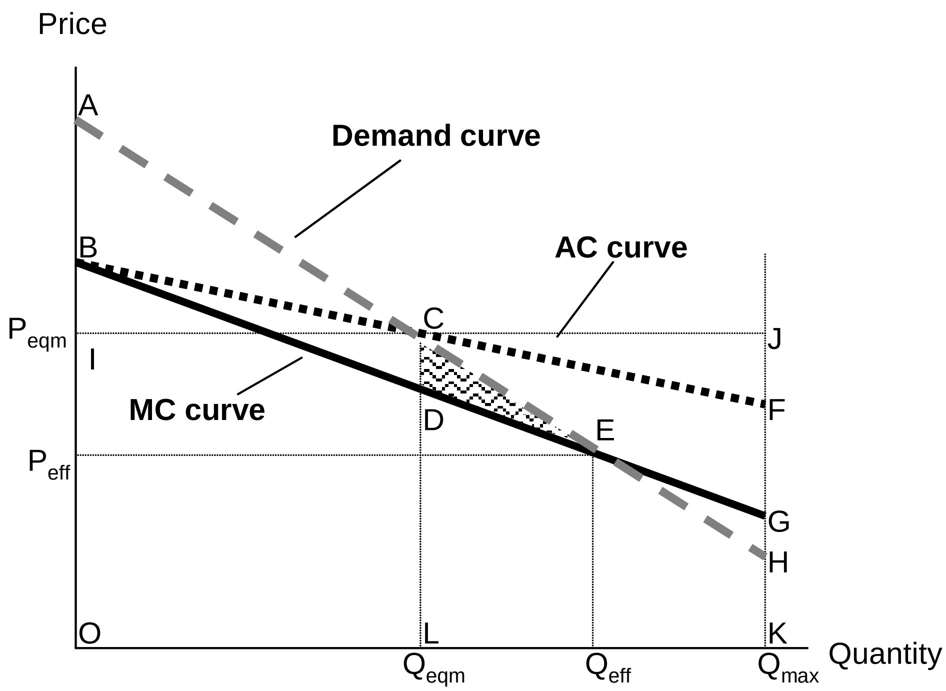
Appendix A

[*Still incomplete; Will contain examples of various behavioral models that could give rise to the same demand function*]

8 Appendix B

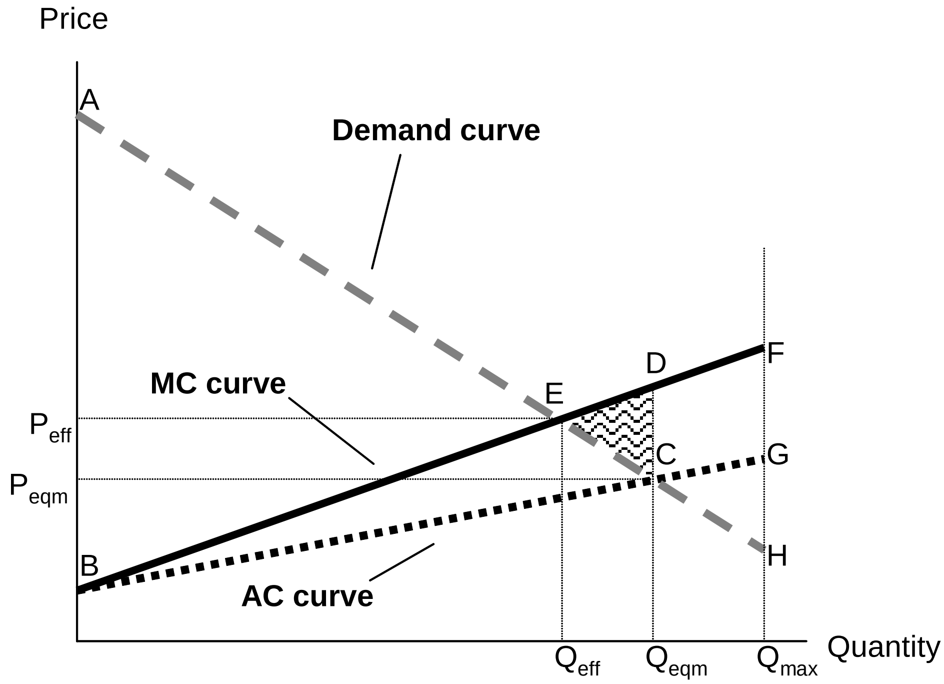
[*Still incomplete; Will contain the algebra that gives rise to the various welfare calculations*]

Figure 1: Efficiency cost of adverse selection - theory



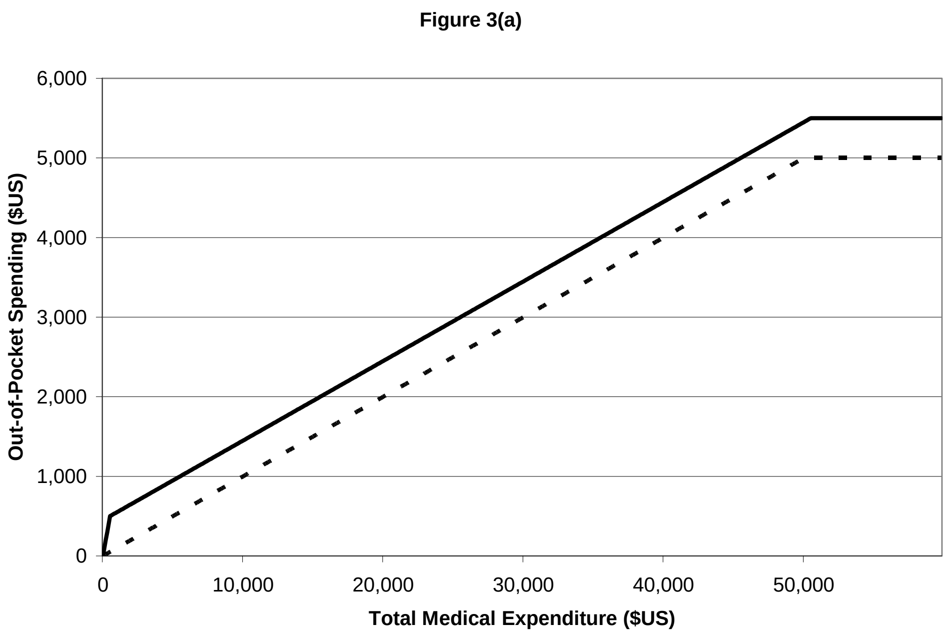
This figure presents the theoretical effect of adverse selection. This is adverse selection because the marginal cost curve is downward sloping, so that the people who have the highest willingness to pay also have the highest expected cost to the insurer. Competitive equilibrium is given by point C (where the demand crosses the average cost curve), while the efficient allocation is given by point E (where the demand crosses the marginal cost curve). The (shaded) triangle CDE represents the welfare cost from under-insurance due to adverse selection.

Figure 2: Efficiency cost of advantageous selection - theory



This figure presents the theoretical effect of advantageous selection. This is advantageous selection because the marginal cost curve is upward sloping, so that the people who have the highest willingness to pay have the lowest expected cost to the insurer. Competitive equilibrium is given by point E (where the demand crosses the average cost curve), while the efficient allocation is given by point C (where the demand crosses the marginal cost curve). The (shaded) triangle CDE represents the welfare cost from over-insurance due to advantageous selection.

Figure 3: Description of the High and Low coverages



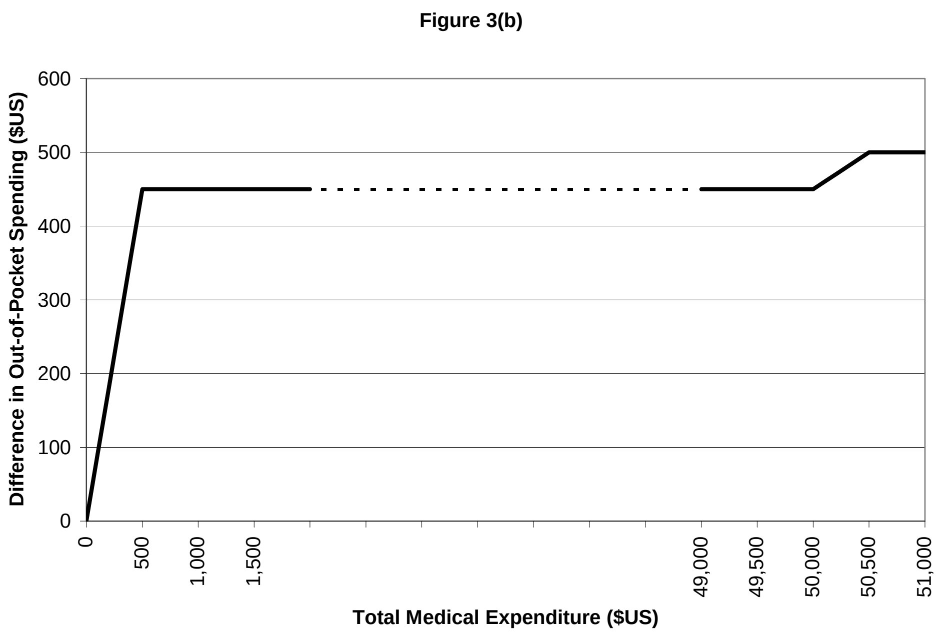
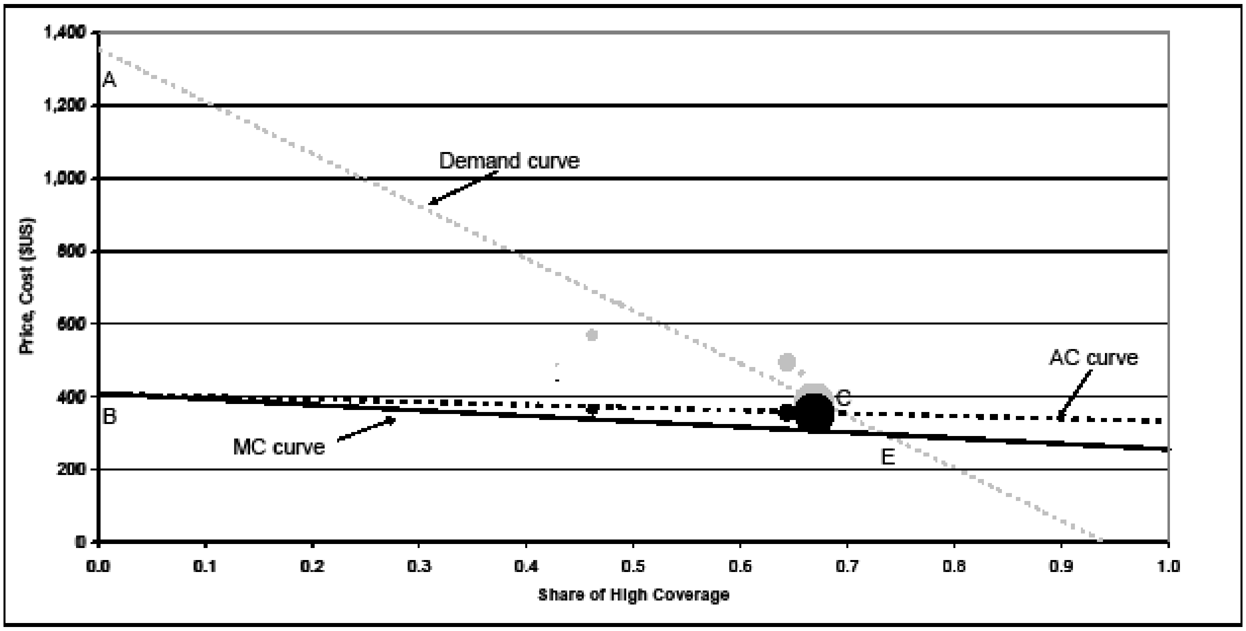


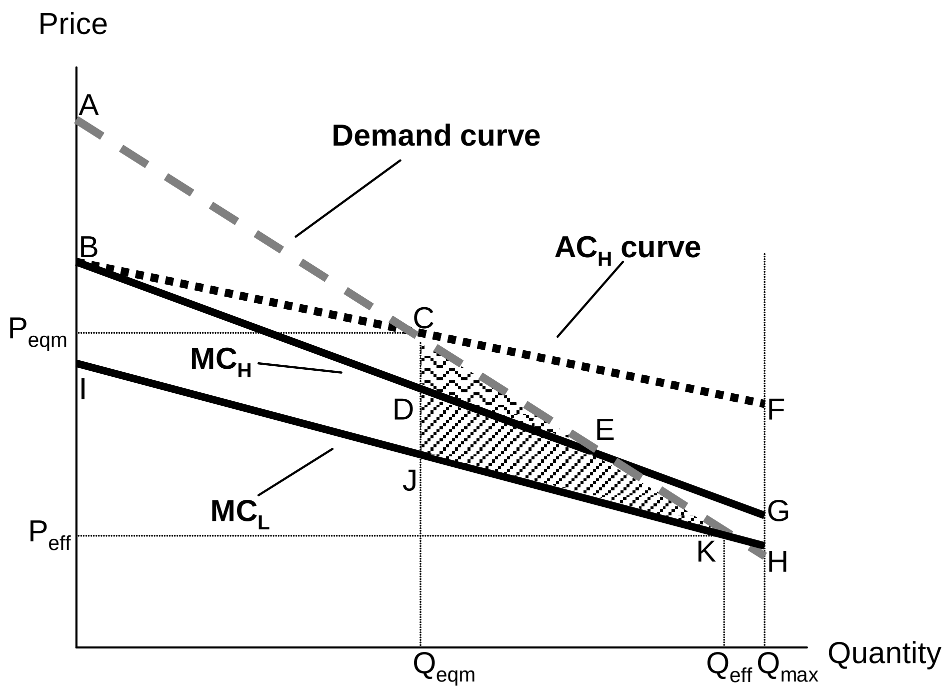
Figure 3(a) presents the main features of the High (dashed) and Low (solid) coverages offered by the company, which is based on a deductible (0 and $500, respectively) and an out-of-pocket maximum ($5,000 and $5,500, respectively). The text provides additional coverage details that the figure above abstracts from. Figure 3(b) presents the corresponding cost difference to the insurer by providing the High coverage instead of the Low coverage, for a given level of medical expenditure. With only few individuals in the data spending close to or more than the out-of-pocket maximum, the relative insurer costs from High coverage are flat (at $450) for almost everyone who spends more than the $500 deductible of the Low coverage.

Figure 4: Efficiency cost of adverse selection - empirical analog



This figure is the empirical analog of the theoretical Figure 1. We graph the estimates of the demand and average cost curves from our baseline specification of workers with family coverage (column (1) of Table 4). We also graph the marginal cost curve derived from these estimates using equation (**??**). We superimpose on these estimates the actual data points (from Table 3), with the size of each data point (circle) being proportional to the number of individuals who faced that price. We label points A, B, C, and E, that correspond to the theoretical analog in Figure 1.

Figure 5: Efficiency cost of adverse selection in the presence of moral hazard - theory



This figure extends Figure 1 by accounting for moral hazard. In the presence of moral hazard, the MC_L curve lies below the MC_H curve, and the efficiency cost of adverse selection is bounded between the area of triangle CDE and the triangle CJK. In that sense, the point K, which we term as the efficient allocation, is only efficient at the upper bound. At the lower bound case, the efficient allocation is, as in Figure 1, point E.

Table 1: Summary statistics

		2004 Company Data					March 2005 CPS	
		All Employees	Only Salaried Workers	Only Salaried Workers with New Benefit Design	Col. (3) Limited to only Workers who Chose High or Low	Col. (4) Limited to Workers with Family Coverage	All Full Time Workers	Only in Manufacturing
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
Number of Individuals		36,829	11,966	11,261	7,263	3,779	83,118	11,178
Fraction Male		0.78	0.73	0.73	0.77	0.86	0.58	0.70
Fraction White		0.77	0.87	0.86	0.86	0.86	0.82	0.82
Fraction unionized		0.33	0.00	0.00	0.00	0.00	0.13	0.14
Age	Mean	44.24	44.51	44.45	45.17	42.66	41.39	42.13
	Std. Deviation	9.86	9.23	9.20	9.12	7.22	12.33	11.45
	Median	45	45	45	46	43	41	42
Tenure with company (years)	Mean	13.23	13.26	13.25	13.69	12.70	n/a	n/a
	Std. Deviation	10.28	9.95	9.95	10.01	8.93	n/a	n/a
	Median	11	12	12	13	12	n/a	n/a
Annual Salary (current $US)	Mean	53,097	71,617	72,906	74,017	80,999	41,869	46,195
	Std. Deviation	47,633	77,930	79,510	91,530	112,790	47,955	45,435
	Median	47,271	60,484	61,504	61,822	66,335	32,000	35,000

Columns (1) through (5) present summary statistics for different cuts of the 2004 employees at the company. Column (1) presents statistics for the all active employees in our sample, column (2) for salaried workers only. Column (3) looks at a slightly smaller group of salaried employees who faces the new benefit design, and column (4) further restricts attention to salaried employees who chose Low or High coverage (and who are the primary focus of our analysis). Column (5) further limits the analysis to those who chose family coverage, who are used to generate our baseline estimates. For comparison, we also present these summary statistics for workers employed full time (defined as those who on average worked 35 or more hours per week in the previous year) in the March 2005 CPS. Column (6) presents the results for all full time workers and column (7) presents results for full time workers in manufacturing industries. In both these columns we use CPS sampling weights ("earning weights" for the union variable, and "person weights" for all others).

Table 2: Assessing the exogeneity of the price variation

	Faced lowest relative price (2,939 workers) (1)	Faced higher relative prices (840 workers) (2)	Difference (3)	Coefficient[b] (4)	p-value[b] (5)
Age (Mean)	42.74	42.40	0.33	-0.002	0.31
Tenure (Mean)	13.02	11.63	1.39	-0.006	0.08
Fraction Male	0.862	0.852	0.009	0.013	0.79
Fraction White	0.874	0.825	0.049	-0.070	0.40
Log(Annula Salary)	11.16	11.05	0.11	-0.086	0.17
Log(2003 Medical Spending + 1)[a]					
All	8.13	7.79	0.32	-0.021	0.15
In most common 2003 plan	8.21	8.08	0.13	-0.017	0.08

The table reports average differences in covariates (shown in the left column) across workers who face different relative prices for the higher coverage option. Sample is limited to salaried workers with family coverage who choose High or Low coverage (i.e. column 5 of Table 1). Columns (1) and (2) present, respectively, average characteristics for the approximately three-quarters of employees who face the lowest relative price ($384; see Table 3) and the remaining one quarter who face one of the five higher relative prices ($466 to $659; see Table 3). Column (3) shows the difference between columns (1) and (2). The worker characteristics in the left column represent contemporaneous 2004 characteristics (age, job tenure, male, white, and salary).

[a] In the bottom two rows we look at 2003 medical spending for all the workers in the sample who were in the data in 2003 (2,602 and 658 workers in columns (1) and (2), respectively), and for all the workers who were in the data in 2003 in the most common 2003 health insurance plan (2,284 and 523 workers in columns (1) and (2), respectively). The latter attempts to avoid potential differences in spending arising from moral hazard effects of different 2003 coverages.

[b] Columns (4) and (5) report, respectively, the coefficient and p-value from a regression of the (continuous) relative price variable (in hundred $US) on the characteristic given in the left column; we adjust the standard errors for an arbitrary variance covariance matrix within each state.

Table 3: The effect of price on demand and costs

Relative Price	Number of Obs.	Fraction chose High Coverage	Average Relative Cost	
			High Coverage	Low Coverage
(1)	(2)	(3)	(4)	(5)
$384	2,939	0.67	$354.53	$340.84
$466	67	0.66	$365.28	$341.15
$489	7	0.43	$446.12	$338.60
$495	526	0.64	$357.24	$356.35
$570	199	0.46	$366.62	$348.13
$659	41	0.49	$373.59	$370.67

The table presents the raw data underlying our estimates of the demand and cost curves in the baseline specification. Sample is limited to salaried workers with family coverage who choose High or Low coverage (i.e. column 5 of Table 1). All individuals face one of six different relative prices, each represented by a row in the table. Column (2) gives the number of employees facing each price, and column (3) reports the fraction who chose High coverage. Columns (4) and (5) report (for High coverage and Low coverage individuals, respectively) the average difference in costs to the insurer of a given family's medical expenditures if these expenditures are covered under High coverage relative to if these same expenditures are covered under Low coverage. The graphical analog to this table is presented by the circles shown in Figure 4.

Table 4: Main results

	Family Coverage		All Coverage Tiers	
	No Controls	w/ State Fixed-Effects	No Controls	w/ State Fixed-Effects
	(1)	(2)	(3)	(4)
Panel A: Demand (Dep. Variable: 1 if Chose High Coverage, 0 if choose Low coverage; Estimated using Full Sample)				
Constant ("alpha")	0.940	0.919	0.848	0.892
	(0.123)	(0.167)	(0.109)	(0.145)
	[0.000]	[0.000]	[0.000]	[0.000]
Relative Price ("beta")	-0.00070	-0.00065	-0.00059	-0.00071
	(0.00032)	(0.00040)	(0.00032)	(0.00040)
	[0.034]	[0.119]	[0.077]	[0.086]
Mean Dependent Variable	0.65	0.65	0.64	0.64
Number of Obs.	3,779	3,779	7,263	7,263
R-Squared	0.010	0.037	0.006	0.026
Panel B: Cost High (Dep. Variable: Additional Insurer Cost of High Coverage; Estimated using High coverage Individuals only)				
Constant ("gamma")	334.0	321.7	295.4	283.8
	(12.0)	(17.0)	(8.82)	(12.0)
	[0.000]	[0.000]	[0.000]	[0.000]
Relative Price ("delta")	0.053	0.086	0.045	0.077
	(0.028)	(0.041)	(0.025)	(0.034)
	[0.066]	[0.047]	[0.078]	[0.028]
Mean Dependent Variable	355.8	355.8	313.5	313.5
Number of Obs.	2,465	2,465	4,662	4,662
R-Squared	0.001	0.020	0.308	0.315
Panel C: Cost Low (Dep. Variable: Additional Insurer Cost of High Coverage; Estimated using Low coverage Individuals only)				
Constant ("gamma")	310.4	295.7	269.3	258.7
	(14.70)	(17.06)	(10.05)	(9.91)
	[0.000]	[0.000]	[0.000]	[0.000]
Relative Price ("delta")	0.080	0.115	0.057	0.086
	(0.033)	(0.041)	(0.028)	(0.027)
	[0.022]	[0.008]	[0.050]	[0.003]
Mean Dependent Variable	344.1	344.1	285.9	285.9
Number of Obs.	1,314	1,314	2,641	2,641
R-Squared	0.003	0.027	0.341	0.348

The table reports results from estimation of the demand equation: $D = \alpha + \beta p$ (equation (10)) and the cost equation $c = \gamma + \delta p$ (equation (11)). For the demand equation, the dependent variable is an indicator variable for whether the individual chose High coverage (as opposed to Low coverage). For the cost equation, the dependent variable is the difference in costs to the insurer of a given employee (and coveraged dependents') medical expenditures if these expenditures are covered under High coverage relative to if these same expenditures are covered under Low coverage. In Panel B, the cost equation is estimated on the sub-sample who chose High coverage; in Panel C the cost equation is estimated on the sub-sample who chose Low coverage. Each column reports the results from a different specification. Column (1) reports results for the sample with family coverage; column (2) replicates this analysis with the addition of state fixed effects. In columns (3) and (4) we repeat this analysis (without and with state fixed effects, respectively) using the sample in all four coverage tiers. In both columns 3 and 4 we include (de-meaned) indicator variables for the coverage tier (not shown); we also multiply the price of "employee only" coverage by two to make the analysis comparable to that of the other coverage tiers, which have double the deductible and out-of-pocket maximums of the single coverage. Standard errors (in parentheses) are adjusted for an arbitrary variance covariance matrix within each state; corresponding p-values are in square brackets.

Table 5: Welfare implications

			Family Coverage		All Coverage Tiers	
			No Controls	w/ State Fixed-Effects	No Controls	w/ State Fixed-Effects
			(1)	(2)	(3)	(4)
	Panel A: Lower Bound Estimates					
1	Efficiency Cost of Adverse Selection	Area CDE	$1.09	$3.30	$0.83	$2.24
2	Efficient Price	point E	$294	$240	$254	$221
3	Efficient Quantitiy	point E	0.74	0.76	0.70	0.74
4	Equilibrium Price	point C	$353	$351	$309	$308
5	Equilibrium Quantity	point C	0.69	0.69	0.67	0.67
6	Social Cost of Optimal Subsidy	0.3 * row3 * (row4 - row2)	$13.10	$25.31	$11.55	$19.31
7	Total Surplus from Efficient Pricing	Area ABE	$347.54	$375.79	$377.75	$321.82
8	Relative Efficiency Cost	row 1 / row 7	0.31%	0.88%	0.22%	0.70%
	Panel B: Upper Bound Estimates					
9	Efficiency Cost of Adverse Selection	Area CJK				
10	Efficient Price	point K				
11	Efficient Quantitiy	point K				
12	Equilibrium Price	point C	$353	$351	$309	$308
13	Equilibrium Quantity	point C	0.69	0.69	0.67	0.67
14	Social Cost of Optimal Subsidy	0.3 * row11 * (row12 - row10)				
15	Total Surplus from Efficient Pricing	Area AIK				
16	Relative Efficiency Cost	row 9 / row 15				

Each row uses the demand and cost estimates from the corresponding column in Table 4 to compute the measure described in the left hand column. The labels in the second column refer to Figure 1 (or, equivalently, Figure 4). Appendix B provides the algebra behind each calculation. Panel A uses the cost estimates from Table 4 limited to those who chose High coverage (Panel B of Table 4). Panel B uses the cost estimates from Table 4 limited to those who chose Low coverage (Panel C of Table 4). The estimates in Panel A are exact welfare estimates if there is no moral hazard, and a lower bound on the welfare cost of adverse selection and of the social cost of the optimal subsidy if there is moral hazard. The estimates in Panel B give an upper bound on the welfare cost of adverse selection and of the social cost of the optimal subsidy if there is moral hazard. The social cost of the optimal subsidy (rows 6 and 14) is calculated as the difference between the equilibrium price and the efficient price, multiplied by the marginal cost of public funds, which we assume to be 0.3.