Private Information and Insurance Rejections

Nathaniel Hendren*

JOB MARKET PAPER (November, 2011)

Abstract

Across a wide set of non-group insurance markets, applicants are rejected based on observable, often high-risk, characteristics. This paper explores private information as a potential cause. To do so, we develop and test a model in which agents have private information about their risk. We derive a new no-trade result that can theoretically explain how private information could cause rejections. We use the no-trade condition to generate measures of the barrier to trade private information imposes. We develop a new empirical methodology to estimate these measures that uses subjective probability elicitations as noisy measures of agents' beliefs. We apply our approach to three non-group markets: long-term care, disability, and life insurance. Consistent with the predictions of the theory, in all three settings we find larger barriers to trade imposed by private information for those who would be rejected relative to those who are served by the market. For those who would be rejected, private information imposes a barrier to trade equivalent to an implicit tax on insurance premiums of roughly 65-75% in long-term care, 90-130% in disability, and 65-130% in life insurance.

JEL classification numbers: C51, D82

Keywords: Private Information; Adverse Selection; Insurance

1 Introduction

Not everyone can purchase insurance. Across a wide set of non-group insurance markets, companies choose to not sell insurance to potential customers with certain observable, often high-risk, characteristics. In the non-group health insurance market, 1 in 7 applications to the four largest insurance companies in the United States were rejected between 2007 and 2009, a figure that excludes those who would be rejected but were deterred from even applying.[1] In US long-term

*I am very grateful to Daron Acemoglu, Amy Finkelstein, Jon Gruber, and Rob Townsend for their guidance and support. I also thank Victor Chernozhukov, Sarah Miller, Whitney Newey, Ivan Werning, and a extensive list of MIT graduate students for helpful comments and suggestions. Support from NSF Graduate Research Fellowship and the NBER Health and Aging Fellowship, under the National Institute of Aging Grant Number T32-AG000186 is gratefully acknowledged.

[1]Figures obtained through a formal congressional investigation by the Committee on Energy and Commerce, which requested and received this information from Aetna, Humana, UnitedHealth Group, and WellPoint. Congressional report was released on October 12, 2010. The 1 in 7 figure does not subtract duplicate applications if people applied to more than 1 of these 4 firms.

care insurance, 12-23% of 65 year olds have health conditions that would preclude them from being able to purchase insurance (Murtaugh et al. [1995]).[2]

It is surprising that a company would choose to not offer its products to a certain subpopulation. Although the rejected generally have higher expected expenditures, they still face unrealized risk.[3] Regulation does not generally prevent risk-adjusted pricing in these markets, so why not simply offer them a higher price?

In this paper, we explore whether private information can explain rejections. We begin by developing a model of how private information could cause rejections. Our setting is the familiar binary loss environment introduced by Rothschild and Stiglitz [1976], which we generalize to incorporate an arbitrary distribution of privately informed types. We study the set of implementable allocations, which satisfy resource, incentive, and participation constraints - constraints that must hold across market structures such as monopoly or competition.

We derive a "no-trade" condition which characterizes when insurance companies would be unwilling to sell insurance on terms that anyone in the market would accept. This condition has an unraveling intuition similar to the one introduced in Akerlof [1970]. The market unravels when the willingness to pay for a small amount of insurance is less than the pooled cost of providing this insurance to those equal to, or higher than, an individuals' own cost. When this no-trade condition holds, an insurance company cannot offer any contract, or menu of contracts, because they would attract an adversely selected subpopulation that would render them unprofitable. Thus, the theory explains rejections as segments in which the no-trade condition holds.

We use the no-trade condition to generate comparative static predictions for properties of type distributions which are more likely to lead to no trade. In particular, we characterize the barrier to trade in terms of an equivalence to a tax rate levied on insurance premiums in a world with no private information. The comparative statics reveal a qualitative explanation for why it is so often the observably high-risk who are rejected: when distributions can be ordered according to a hazard rate ordering, higher mean risk distributions impose a higher implicit informational tax.

We then develop a new empirical methodology for studying private information to test the predictions of theory. We use information contained in subjective probability elicitations to infer properties of the distribution of private information. At no point do we view these elicitations as true beliefs. Rather, we use information in the joint distribution of elicitations and the realized events corresponding to these elicitations to deal with potential errors in elicitations.[4] We proceed with two complementary approaches. First, we make the weak assumption that agent's elicitations cannot contain more information about the subsequent loss than would the true

[2] Appendix C presents the rejection conditions from Genworth Financial (one of the largest US LTC insurers), gathered from their underwriting guidelines provided to insurance agents for use in screening applicants.

[3] For example, in long-term care we estimate those who would be rejected have an average five-year nursing home entry rate of less than 20%.

[4] In this sense, our approach builds on previous work using subjective probabilities in economics (e.g. Gan et al. [2005], see Hurd [2009] for a review).

beliefs.[5] We estimate the explanatory power of the subjective probabilities on the subsequent realized event, conditional on public information. This allows us to generate nonparametric lower bounds on a measure of the magnitude of private information provided by the theory. With these bounds, we provide a simple test for the presence of private information, along with a test of whether those who would be rejected have larger estimates of this lower bound.

Our second approach moves from a nonparametric lower bound to a semiparametric point estimate of the distribution of private information by making an additional parametric assumption on the distribution of elicitation error which allows elicitations to be noisy and potentially biased measures of agents true beliefs. We then flexibly estimate the distribution of private information. This allows us to quantify the barrier to trade in terms of the implicit informational tax rate imposed by private information. We then test whether this quantity is larger for those who would be rejected relative to those who are served by the market and whether it is large (small) enough to explain (the absence of) rejections for plausible values of agents' willingness to pay for insurance.

We apply our approach to three non-group markets: long-term care (LTC), disability, and life insurance. We combine two sources of data. First, we use data from the Health and Retirement Study, which elicits subjective probabilities corresponding to losses insured in each of these three settings and contains a rich set of demographic and health information commonly used by insurance companies in pricing insurance. We supplement this with a detailed review of underwriting guidelines from major insurance companies to identify those who would be rejected (henceforth "rejectees"[6]) in each market.

Across all three market settings and a wide set of specifications, we find robust support for the hypothesis that private information causes insurance rejections. We find larger nonparametric lower bounds on a measure of the magnitude of private information for rejectees relative to those served by the market. Our semiparametric approach reveals an informational implicit tax rates for rejectees of 68-73% in LTC, 90-128% in Disability, and 64-127% tax in Life; in each setting we estimate smaller barriers to trade for non-rejectees. Finally, not only can we explain rejections in these three non-group markets, but the estimated distribution of private information about mortality (constructed for our life insurance setting) can also explain the *lack of* rejections in annuity markets. While some individuals are informed about being a relatively high mortality risk, very few are exceptionally informed about having low mortality risk. Thus, low mortality risks can obtain annuities without a significant number of even lower mortality risks adversely selecting their contract.

Our paper is related to several distinct literatures. On the theoretical dimension, it is, to our knowledge, the first paper to show that private information can lead to no gains to trade in an

[5] If beliefs are generated through rational expectations given some information set, this assumption is equivalent to assuming the elicitations are a garbling of the agent's true beliefs in the sense of Blackwell [1951, 1953].

[6] Throughout, we focus on those who "would be rejected", which corresponds to those whose choice set excludes insurance, not necessarily the same as those who actually apply and are rejected.

insurance market with an endogenous set of contracts. While no trade can occur in the Akerlof [1970] lemons model, this model exogenously restricts the set of tradeable contracts, which is unappealing in the context of insurance since insurers generally offer a menu of premiums and deductibles. In this sense, our paper is more closely related to the large screening literature using the binary loss environment initially proposed in Rothschild and Stiglitz [1976]. While the Akerlof lemons model restricts the set of tradeable contracts, this literature generally restricts the distribution of types (e.g. "two types" or a bounded support) and generally argues that trade will always occur (Riley [1979]; Chade and Schlee [2011]). But by considering an arbitrary distribution of types, we show this not to be the case. Indeed, the no trade condition we provide can hold under common distributions previously not addressed. For example, with a uniform distribution of types (over $[0, 1]$), trade cannot occur unless individuals are willing to pay more than a 100% tax for insurance.

Empirically, our paper is related to a recent and growing literature on testing for the existence and consequences of private information in insurance markets (Chiappori and Salanié [2000]; Chiappori et al. [2006]; Finkelstein and Poterba [2002, 2004]; see Einav et al. [2010a] and Cohen and Siegelman [2010] for a review). This literature focuses on the revealed preference implications of private information by looking for a correlation between insurance purchase and subsequent claims. This approach can only identify private information amongst those served by the market. In contrast, our approach can study private information for the entire population, including rejectees. Our results suggest significant amounts of private information for the rejectees, but less for those served by the market. Thus, our results provide a new explanation for why previous studies using the revealed preference approach have not found evidence of significant adverse selection in life insurance (Cawley and Philipson [1999]) and LTC insurance (Finkelstein and McGarry [2006]). The absence of adverse selection may be the insurer's selection.

Finally, our paper is related to the broader literature on the workings of markets under uncertainty and private information. While many theories have pointed to potential problems posed by private information, our paper presents, to the best of our knowledge, the first empirical evidence that private information can lead to a complete absence of trade.

The rest of this paper proceeds as follows. Section 2 presents the theory and the no-trade result. Section 3 presents the comparative statics and testable predictions of the model. Section 4 outlines the empirical methodology. Section 5 presents the three market settings and our data. Section 6 presents the empirical specification and results for the nonparametric lower bounds. Section 7 presents the empirical specification and results of the semiparametric estimation of the distribution of private information. Section 8 concludes.

2 Theory

This section develops a model of private information. Our primary result (Theorem 1) is a no-trade condition which provides a theory of how private information can lead insurance companies to not offer any contracts.

2.1 Environment

There exists a unit mass of agents endowed with non-stochastic wealth $w > 0$. All agents face a potential loss of size $l > 0$ that occurs with privately known probability p, which is distributed with c.d.f. $F(p)$ in the population. We impose no restrictions on $F(p)$; it may be a continuous, discrete, or mixed distribution, and have full or partial support, which we denote by $\Psi \subset [0,1]$.[7] Throughout the paper, we let the uppercase P denote the random variable representing a random draw from the population (with c.d.f. $F(p)$) and the lowercase p denote a specific agent's probability (i.e. their realization of P). Agents have observable characteristics, X. For now, one should assume that we have conditioned on observable information (e.g. $F(p) = F(p|X = x)$ where X includes all observable characteristics such as age, gender, and observable health conditions).

Agents have a standard Von-Neumann Morgenstern preferences $u(c)$ with expected utility given by

$$pu(c_L) + (1-p)\,u(c_{NL})$$

where c_L (c_{NL}) is the consumption in the event of a loss (no loss). We assume $u(c)$ is continuously differentiable, with $u'(c) > 0$ and $u''(c) < 0$. An allocation $A = \{c_L(p), c_{NL}(p)\}_{p\in\Psi}$ consists of consumption in the event of a loss, $c_L(p)$, and in the event of no loss, $c_{NL}(p)$ for each type $p \in \Psi$.

While it is common in this environment to now introduce a specific institutional structure, such as a game of competition or monopoly, our approach is different. Instead, we abstract from specific institutional structure and study the set of implementable allocations.

Definition 1. An allocation $A = \{c_L(p), c_{NL}(p)\}_{p\in\Psi}$ is **implementable** if

1. A is resource feasible:

$$\int [w - pl - pc_L(p) - (1-p)\,c_{NL}(p)]\,dF(p) \geq 0$$

[7] By choosing particular distributions $F(p)$, our environment nests many previous models of insurance. For example, $\Psi = \{p_L, p_H\}$ yields the classic two-type model considered initially by Rothschild and Stiglitz [1976] and subsequently analyzed by many others. Assuming $F(p)$ is continuous with $\Psi = [a,b] \subset (0,1)$, one obtains an environment similar to Riley [1979]. Chade and Schlee [2011] provide arguably the most general treatment to-date of this environment in the existing literature by considering a monopolists problem with an arbitrary F with bounded support $\Psi \subset [a,b] \subset (0,1)$.

2. A is incentive compatible:

$$pu\left(c_L\left(p\right)\right)+\left(1-p\right)u\left(c_{NL}\left(p\right)\right)\geq pu\left(c_L\left(\tilde{p}\right)\right)+\left(1-p\right)u\left(c_{NL}\left(\tilde{p}\right)\right)\quad\forall p,\tilde{p}\in\Psi$$

3. A is individually rational:

$$pu\left(c_L\left(p\right)\right)+\left(1-p\right)u\left(c_{NL}\left(p\right)\right)\geq pu\left(w-l\right)+\left(1-p\right)u\left(w\right)\quad\forall p\in\Psi$$

Our focus on the set of implementable allocations makes our results applicable across institutional settings, such as monopoly or competition. Any economy which faces the above information and resource constraints must yield implementable allocations. Moreover, by focusing on implementable allocations we circumvent problems arising from the potential non-existence of competitive Nash equilibriums, as highlighted in Rothschild and Stiglitz [1976].

2.2 The No-Trade condition

Theorem 1 characterizes when the endowment is the only implementable allocation.

Theorem 1. *(No Trade). The endowment, $\{(w-l,w)\}$, is the only implementable allocation if and only if*

$$\frac{p}{1-p}\frac{u'\left(w-l\right)}{u'\left(w\right)}\leq\frac{E\left[P|P\geq p\right]}{1-E\left[P|P\geq p\right]}\quad\forall p\in\Psi\backslash\left\{1\right\}\tag{1}$$

where $\Psi\backslash\{1\}$ denotes the support of $F(p)$ excluding the point $p=1$.

Conversely, if (1) does not hold, then there exists an implementable allocation which strictly satisfies resource feasibility and individual rationality for a positive mass of types.

Proof. See Appendix A.1[8] □

The left-hand side of equation (1), $\frac{p}{1-p}\frac{u'(w-l)}{u'(w)}$ is the marginal rate of substitution between consumption in the event of no loss and consumption in the event of a loss, evaluated at the endowment, $(w-l,w)$. It is a type p agent's willingness to pay for an infinitesimal amount of additional consumption in the event of a loss, in terms of consumption in the event of no loss. The actuarially fair cost of this transfer to the type p agent is $\frac{p}{1-p}$. However, the right hand side of equation (1) is the price of providing such a transfer, not at type p's own cost of $\frac{p}{1-p}$, but rather at the average cost if all higher-risk types $P\geq p$ also obtained this transfer, $\frac{E[P|P\geq p]}{1-E[P|P\geq p]}$. Intuitively, if no other contracts are offered, then a contract preferred by type p will also be preferred by all types $P\geq p$, rendering the cheapest possible provision of insurance to type p to be at a price ratio of $\frac{E[P|P\geq p]}{1-E[P|P\geq p]}$. If no agent is willing to pay this cost, the endowment is the only implementable allocation.

[8] While Theorem 1 is straightforward, its proof is less trivial because one must show that Condition 1 rules out not only single contracts but also any menu of contracts in which different types may receive different allocations.

Conversely, if equation (1) does not hold, there exists an implementable allocation which does not totally exhaust resources and provides strictly higher utility than the endowment for a positive mass of types. So, a monopolist insurer could earn positive profits by facilitating trade.[9] In this sense, the no-trade condition (1) characterizes when one would expect trade to occur.

The no-trade condition can hold for common distributions, such as the uniform distribution.

Example 1. *Suppose that $F(p)$ is uniform, $F(p) = p$. Then, $E[P|P \geq p] = \frac{1+p}{2}$. The no-trade condition 1 is given by*

$$\frac{p}{1-p}\frac{u'(w-l)}{u'(w)} \leq \frac{\frac{1+p}{2}}{1-\frac{1+p}{2}} \quad \forall p \in [0,1)$$

which holds if and only if

$$\frac{u'(w-l)}{u'(w)} \leq 2$$

With a uniform distribution of private information, trade can only occur if agents marginal utility of consumption is twice as large in the state where the loss occurs. So, unless agents are willing to pay a 100% tax for insurance (which moves consumption from the state of no loss to the state of the loss), there will be no trade.[10]

The no-trade condition has an unraveling intuition similar to that of Akerlof [1970]. His model considers a given contract and shows that it will not be traded when its demand curve lies everywhere below its average cost curve, which is in turn a function of those who demand it. Our model is different. While Akerlof [1970] derives conditions under which a *given contract* would unravel and result in no trade, our model provides conditions under which *any contract or menu of contracts* would unravel.

This distinction is important since previous literature has argued that trade must always occur environments similar to ours with no restrictions on the contract space (Riley [1979]; Chade and Schlee [2011]). The key difference in our approach is that we do not assume types are bounded away from 1.[11] In fact, the no-trade condition requires the highest risk type in the economy have a probability of a loss arbitrarily close to $p = 1$. Otherwise the highest risk type, say $\bar{p}$, would be able to obtain an actuarially fair full insurance allocation, $c_L(\bar{p}) = c_{NL}(\bar{p}) = w - \bar{p}l$, which would not violate the incentive constraints of any other type.

Corollary 1. *Suppose condition (1) holds. Then $F(p) < 1 \; \forall p < 1$.*

This corollary highlights the unraveling intuition: no trade occurs when people don't want to subsidize risks worse than themselves; this naturally requires the perpetual existence of worse risks.[12]

[9] Also, one can show that a competitive equilibrium, as defined in Miyazaki [1977] and Spence [1978] can be constructed for an arbitrary type distribution $F(p)$ and would yield trade (result available from the author upon request).

[10] We discuss this tax rate analogy further in Section 3

[11] Both Riley [1979] and Chade and Schlee [2011] assume $\Psi \subset [a,b] \subset (0,1)$, so that $b < 1$.

[12] Note that we do not require any positive mass at $p = 1$, as highlighted in Example 1.

At the same time, the fact that the no-trade condition requires risks arbitrarily close to 1 can be viewed as a technicality. In reality, insurance companies offer a finite set of contracts, presumably because they incur a setup cost for creating each contract. If we require that each allocation other than the endowment must attract a non-trivial fraction of types, then we no longer require risks arbitrarily close to 1, as illustrated in Remark 1.

Remark 1. *Suppose each consumption bundle* (c_L, c_{NL}) *other than the endowment must attract a non-trivial fraction* $\alpha > 0$ *of types. More precisely, suppose allocations* $A = \{c_L(p), c_{NL}(p)\}_p$ *must have the property that for all* $q \in \Psi$*,*

$$\mu\left(\{p | (c_L(p), c_{NL}(p)) = (c_L(q), c_{NL}(q))\}\right) \geq \alpha$$

where μ *is the measure defined by* $F(p)$*. Then, the no-trade condition is given by*

$$\frac{p}{1-p}\frac{u'(w-l)}{u'(w)} \leq \frac{E[P|P \geq p]}{1 - E[P|P \geq p]} \; \forall p \in \hat{\Psi}_{1-\alpha}$$

where $\hat{\Psi}_{1-\alpha} = \left[0, F^{-1}(1-\alpha)\right] \cap (\Psi \backslash \{1\})$*.*[13] *Therefore, the no-trade condition need only hold for values* $p < F^{-1}(1-\alpha)$*.*

In other words, if contracts must attract a nontrivial fraction of types, then no trade can occur even if types are bounded away from $p = 1$*. Going forward, we retain the benchmark assumption of no such frictions or transactions costs, but return to this discussion in our empirical work in Section 7.*

The no-trade condition (1) provides a theory of rejections: they occur in market segments where (1) holds and insurance is offered in segments where (1) does not hold, where market segments are defined by observabe information. In order to derive testable implications of this theory, the next section examines properties of distributions, $F(p)$, which make the no-trade condition more likely to hold.

3 Comparative Statics and Testable Predictions

Qualitatively, Theorem 1 suggests a property of distributions which lead to no trade: thick upper tails of risks. The presence of a thicker upper tail increases the value of $E[P|P \geq p]$ at given values of p. In this section, we formalize this intuition by constructing precise measures of the barrier to trade imposed by private information which will guide our empirical tests of the theory.

[13] *If* $F^{-1}(1-\alpha)$ *is a set, we take* $F^{-1}(1-\alpha)$ *to be the supremum of this set*

3.1 Two Measures of Private Information

We construct two measures of private information. To begin, we multiply the no-trade condition (1) by $\frac{1-p}{p}$ yielding,

$$\frac{u'(w-l)}{u'(w)} \leq \frac{E[P|P \geq p]}{1-E[P|P \geq p]} \frac{1-p}{p} \quad \forall p \in \Psi \backslash \{1\}$$

The left-hand side is the ratio of the agents' marginal utilities in the loss versus no loss state, evaluated at the endowment. The right-hand side independent of the utility function, u, and is the cost of providing an infinitesimal transfer to type p if the pool of types worse than p, $P \geq p$, also were attracted to the contract. We define this term the pooled price ratio.

Definition 2. For any $p \in \Psi \backslash \{1\}$, the **pooled price ratio at** p, $T(p)$, is given by

$$T(p) = \frac{E[P|P \geq p]}{1-E[P|P \geq p]} \frac{1-p}{p} \tag{2}$$

Given $T(p)$, the no-trade condition has a succinct expression.

Corollary 2. *(Quantification of the barrier to trade) The no-trade condition holds if and only if*

$$\frac{u'(w-l)}{u'(w)} \leq \inf_{p \in \Psi \backslash \{1\}} T(p) \tag{3}$$

Whether or not there will be trade depends on only two numbers: the agent's underlying valuation of insurance, $\frac{u'(W-L)}{u'(W)}$, and the cheapest cost of providing an infinitesimal amount of insurance, $\inf_{p \in \Psi \backslash \{1\}} T(p)$. When this cost is above the underlying valuation of insurance, there can be no trade. We call $\inf_{p \in \Psi \backslash \{1\}} T(p)$ the *minimum pooled price ratio.* This number characterizes the barrier to trade imposed by private information.

Equation (3) has a simple tax rate interpretation. Suppose for a moment that there were no private information but instead a government levies a sales tax of rate t on insurance premiums in a competitive insurance market. The value $\frac{u'(w-l)}{u'(w)}-1$ is the highest such tax rate an individual would be willing to pay to purchase any insurance.[14] Thus, $\inf_{p \in \Psi \backslash \{1\}} T(p) - 1$ is the tax rate equivalent of the barrier to trade imposed by private information. In this sense, it quantifies the magnitude of the barrier to trade imposed by private information.

Equation (3) leads to a simple comparative static.

Corollary 3. *(Comparative static in the minimum pooled price ratio) Consider two market segments with pooled price ratios $T_1(p)$ and $T_2(p)$ and common vNM preferences u. Suppose*

$$\inf_{p \in \Psi \backslash \{1\}} T_1(p) \leq \inf_{p \in \Psi \backslash \{1\}} T_2(p)$$

[14] To clarify, the equivalence is to a tax rate paid only in the state of no loss, so that it can be interpreted as a tax on the insurance premium.

then if the no-trade condition holds in segment 1, it must also hold in segment 2.

Higher values of the minimum pooled price ratio are more likely to lead to no trade. Because the minimum pooled price ratio characterizes the barrier to trade imposed by private information, Corollary 3 is the key comparative static on the distribution of private information provided by the theory.

In addition to the minimum pooled price ratio, we also provide another metric which leads to a less precise comparative static but will be useful to guide portions of our empirical analysis.

Definition 3. For any $p \in \Psi$, define the **magnitude of private information at** p by $m(p)$, given by

$$m(p) = E[P|P \geq p] - p \tag{4}$$

The value $m(p)$ is the difference between p and the average probability of everyone worse than p. Note that $m(p) \in [0,1]$ and $m(p) + p = E[P|P \geq p]$. The following comparative static follows directly from the no-trade condition (1).

Corollary 4. *(Comparative static in the magnitude of private information) Consider two market segments with magnitudes of private information $m_1(p)$ and $m_2(p)$ and common support Ψ and common vNM preferences u. Suppose*

$$m_1(p) \leq m_2(p) \;\; \forall p \in \Psi$$

Then if the no-trade condition holds in segment 1, it must also hold in segment 2.

Higher values of the magnitude of private information are more likely to lead to no trade. Notice that the values of $m(p)$ must be ordered for all $p \in \Psi$, and it is thus a less precise statement than the comparative static provided in Corollary 3.

3.2 High-Risk Distributions

Before turning to our empirical methodology, we note that the comparative statics of the model already provide a qualitative explanation of the fact that it is often the high (mean) risks who are rejected. Let P_1 and P_2 be two continuously distributed random variables with common support $\Psi \subset [0,1]$ and hazard rates $h_j(p) = \frac{f_j(p)}{1-F_j(p)}$, where $f_j(p)$ is the p.d.f. and $F_j(p)$ is the c.d.f. of P_j. We say that the two random variables are *ordered according to the hazard rate ordering* if either $h_1(p) \leq h_2(p)$ for all p or $h_1(p) \geq h_2(p)$ for all p.

Proposition 1. *Suppose P_1 and P_2 are ordered according to the hazard rate ordering. Let T_1 and T_2 denote their associated pooled price ratios. Then*

$$E[P_1] \leq E[P_2] \implies \inf_{p \in \hat{\Psi}} T_1(p) \leq \inf_{p \in \hat{\Psi}} T_2(p) \tag{5}$$

for any $\hat{\Psi} \subset \Psi \backslash \{1\}$. In particular, (5) holds for $\hat{\Psi} = \Psi \backslash \{1\}$ or $\hat{\Psi} = \Psi_{1-\alpha}$ as defined in Remark 1.

Proof. Follows immediately from the fact that the hazard rate ordering implies the mean-residual life ordering. See Shaked and Shanthikumar [1994]. □

When distributions can be ordered according to their hazard rates, the higher mean risk distribution has a larger minimum pooled price ratio.[15] Therefore, it satisfies the no-trade condition for a larger set of values of $\frac{u'(w-l)}{u'(w)}$. In this sense, higher risk distributions are more likely to lead to no trade, which can explain why it is so often those with high (mean) risk characteristics who are rejected.

3.3 Moving Towards Data: Testable Hypotheses

Our goal of the rest of the paper is to test the empirical predictions of the theory by estimating properties of the distribution of private information, $F(p|X)$, for rejectees and non-rejectees. Assuming for the moment that $F(p|X)$ is observable to the econometrician, our ideal tests are as follows. Qualitatively, we test whether $F(p|X)$ has a thicker upper tail of high risks for the rejectees. Quantitatively, we estimate the minimum pooled price ratio for each X and conduct two types of tests: first, we test the comparative statics given by Corollaries 3 and 4 of higher values of the minimum pooled price ratio for rejectees versus non-rejectees. Second, we ask whether the minimum pooled price ratio is large (small) enough to explain (the absence of) rejections for plausible values of agents' willingness to pay, as suggested by Corollary 2.[16]

Of course the execution of these tests require estimating properties of the distribution of private information, $F(p|X)$, to which we now turn.

4 Empirical Methodology

We develop an empirical methodology to study private information and operationalize the tests in Section 3.3. The key feature of our approach is that we utilize information contained in subjective probability assessments to infer properties of the distribution of private information. Let L denote an event (e.g. dying in the next 10 years) that is commonly insured in some

[15] Note that the hazard rate ordering is weaker than the likelihood ratio ordering. So if distributions can be ordered according to their likelihood ratios (e.g. they have the monotone likelihood ratio property, "MLRP"), then higher mean risk distributions lead to larger minimum pooled price ratios.

[16] Our tests do not focus on potential demand side variation across values of X (i.e. how willingness-to-pay, $\frac{u'(w-l)}{u'(w)}$ varies with X and potentially differs across rejectees and non-rejectees). Finding empirical support for our comparative static tests would only be inconsistent with the theory if the difference in willingness-to-pay for rejectees versus non-rejectees is larger than our estimated differences in the minimum pooled price ratio. In contrast, if rejectees have lower willingness-to-pay than non-rejectees, our tests are too strict: they may lead us to find evidence inconsistent with the theory when in fact the theory is correct.

insurance market (e.g. life insurance).[17] Let Z denote an individual's subjective probability elicitation about event L (i.e. Z is a response to the question "what do you think is the probability that L will occur?"). A premise of our approach is that these elicitations are non-verifiable to an insurance company. Therefore, they can be excluded from the set of public information, which we will denote by X, and used to infer properties of the distribution of private information. But while these elicitations are non-verifiable to insurance companies, they are arguably noisy and potentially biased measures of true beliefs.

We develop two complementary approaches for dealing with the potential error in subjective probability elicitations. Our first approach provides a nonparametric lower bound on the average magnitude of private information, $E[m(P)]$, and tests whether rejection segments have higher values of $E[m(P)]$. This provides a test in the spirit of the comparative static in $m(p)$ (Corollary 4) while relying on very minimal assumptions on the relationship between agents' beliefs and their probability elicitations. Our second approach adds a parametric structure to the distribution of elicitation error, which allows us to (non-parametrically) identify the distribution of private information (so that the overall approach is semiparametric). We then estimate the pooled price ratio, $T(p)$, and a close analogue to the minimum pooled price ratio, $\inf_{p\in\Psi\setminus\{1\}} T(p)$, where we focus on the minimum over a compact set $\hat{\Psi}$ which excludes points in the upper quantiles of $F(p)$ to avoid problems associated with extreme value estimation. We then test both whether segments facing rejection have larger values of the minimum pooled price ratio (Corollary 3) and whether these estimates are large (small) to explain (the absence of) rejections for plausible values of $\frac{u'(w-l)}{u'(w)}$, as suggested by Corollary 2.

In this section, we introduce these empirical approaches. We defer a discussion of the empirical specification and statistical inference to Sections 6 and 7, after we have discussed our data and settings.

4.1 Nonparametric Lower Bound Approach

To begin, we retain the assumption from the theoretical section that agents act as if they have beliefs about the probability of the loss L.[18] Moreover, as has heretofore been implicit, we assume these beliefs are correct.

Assumption 1. *Beliefs P are correct:* $\Pr\{L|X,P\} = P$

Assumption 1 states that if we hypothetically gathered a large group of individuals with the same observable values X and the same beliefs P and then observed whether or not they

[17] Of course, individuals face more than a single binary event and insurance generally insures a combination of many different events. Our approach is to focus on one commonly insured event and ask whether the pattern of rejections in that market is consistent with the predictions of our theory about whether insurance could be provided for that binary event.

[18] Our approach therefore follows the view of personal probability expressed in the seminal work of Savage [1954]: Although agents may not perfectly express their beliefs through survey elicitations, they would behave consistently in response to gambles over L ("consistently" in the sense of Savage's axioms).

experience the loss L, we would find that, on average, a fraction P of this group experiences the loss. As an empirical assumption, it is relatively strong, but it provides perhaps the simplest link between the realized loss L and beliefs.[19] Note that we have now introduced public information, X. To most closely match the theory, we assume X is the set of information that an insurance company would use to price insurance. We discuss this important data requirement further in Section 5.

Although agents act as if they have beliefs, they may not report these beliefs in probabilistic survey questions. Our lower bound approach assumes only that Z contains no additional information about L than would the true beliefs.

Assumption 2. *Z contains no additional information than P about the loss L, so that* $\Pr\{L|X,P,Z\} = \Pr\{L|X,P\}$

Assumption 2 is very weak; it would be violated only if people could provide elicitations Z which are informative about L even conditional on the true beliefs of those making the reports.[20]

For the empirical tests, we classify segments X into those in which insurance companies do and do not sell insurance, $X \in \Theta^{NoReject}$ and $X \in \Theta^{Reject}$. We then proceed as follows. First, we form the predicted value of L given the observable variables X and Z,

$$P_Z = \Pr\{L|X,Z\}$$

Loosely, our approach asks how much Z explains L, conditional on X. To assess this qualitatively, we plot the predicted values of P_Z separately for rejectees ($X \in \Theta^{Reject}$) and non-rejectees $\left(X \in \Theta^{NoReject}\right)$. If Z is more informative for the rejectees, we would expect to see that the distribution of P_Z given X is more dispersed for the rejectees.

We then measure the extent to which Z explains L conditional on X using a measure of dispersion inspired by the theory. Recall from Definition 3 that $m(p)$ in segment X is given by $m(p) = E[P|P \geq p, X] - p$.[21] We construct an analogue with P_Z,

$$m_Z(p) = E_{Z|X}[P_Z|P_Z \geq p, X] - p$$

[19] This is a common assumption made, either implicitly or explicitly, in existing (revealed preference) approaches to studying private information (e.g. Einav et al. [2010b]). We find some motivation for correct beliefs and our treatment of subjective probability elicitations in existing empirical work in the forecasting literature spanning economics, psychology, and engineering. Broadly, this literature suggests survey elicitations suffer significant limitations as measures of beliefs, but implicit forecasts based on behavior, as in prediction markets, tend to be more accurate (for an overview, see Sunstein [2006] and [Arrow et al. 2008]). Examples of the limitations of survey elicitations of beliefs include Gan et al. [2005] who consider the subjective mortality probabilities we use in this paper. Additional examples in psychology and cognitive engineering shows that simple improvements in elicitation methods can substantially improve forecasts by reducing elicitation biases (Miller et al. [2008], Gigerenzer and Hoffrage [1995]).

[20] Assumptions 1 and 2 are jointly implied by a rational expectations model in which agents know both X and Z in formulating their beliefs P. In this case, our approach views Z as a "garbling" of the agent's true beliefs in the sense of Blackwell ([1951], [1953]).

[21] The expectation is conditional on X but for brevity we omit explicit reference to X in our notation for $m(p)$.

which is difference between p and the average predicted probability, P_Z, of those with predicted probabilities higher than p (note that $m_Z(p)$ is defined for any p).[22] We then construct the average magnitude of private information implied by Z in segment X, $E[m_Z(P_Z)|X]$, which is the average difference in segment X between an individual's predicted loss, and the predicted losses of those with higher predicted probabilities. Intuitively, $E[m_Z(P_Z)|X]$ is a (nonnegative) measure of the dispersion of the distribution of P_Z.

In the spirit of the comparative statics given by Corollary 4, we test whether rejectees have higher values of $E[m_Z(P_Z)|X]$:

$$\Delta_Z = E\left[m_Z(P_Z)|X \in \Theta^{Reject}\right] - E\left[m_Z(P_Z)|X \in \Theta^{NoReject}\right] > 0 \tag{6}$$

which asks whether segments in which insurance companies have chosen to not sell insurance have higher average magnitudes of private information implied by Z than segments in which they sell insurance. Stated more loosely, equation (6) asks whether the subjective probabilities of the rejectees better explain the realized losses than the non-rejectees, where "better explain" is measured using $E[m_Z(P_Z)|X]$. Equation (6) is the key empirical test provided by the lower bound approach.[23]

Lower Bounds Our estimable variable, P_Z, is not equal to the true beliefs, P. Rather we obtain distributional lower bounds, as illustrated in Proposition 2.

Proposition 2. *(Lower bound) Suppose assumptions 1 and 2 hold. Then*

1. *The true beliefs, P, are a mean-preserving spread of P_Z:*

$$P_Z = E[P|X, Z] \tag{7}$$

2. *The average magnitude of private information implied by Z is a lower bound for the true average magnitude of private information:*

$$E[m_Z(P_Z)|X] \le E[m(P)|X] \tag{8}$$

Proof. See Appendix B.1. □

Because Z contains no additional information about L than do the true beliefs P, the true beliefs are a mean preserving spread of P_Z. Correspondingly, the average magnitude of private information implied by Z, $E[m_Z(P_Z)|X]$ is a lower bound for $E[m(P)|X]$.

[22]The subscript Z notes that the variable Z is used in its construction; it does not mean to indicate we are conditioning on a realized value of Z in the construction of $m_Z(p)$.

[23]In addition to aggregating the data across all X in each rejection classification, we also conduct the test for subgroups (e.g. conditional on age or gender).

Statement (2) highlights that testing $E[m_Z(P_Z)|X] = 0$ provides a nonparametric test for the presence of private information (Note $E[m_Z(P_Z)|X] > 0$ implies $E[m(P)|X] > 0$). Since $E[m_Z(P_Z)|X] = 0$ if and only if $\Pr\{L|X,Z\} = \Pr\{L|X\}$, the test for private information is straightforward: do the subjective probabilities explain the realized loss?[24]

Our approach is nonparametric in the sense that we have made no parametric restrictions on how the elicitations Z relate to the true beliefs P.[25] For example, P_Z and $m_Z(p)$ are invariant to monotonic transformations in Z: $P_Z = P_{h(Z)}$ and $m_Z(p) = m_{h(Z)}(p)$ for any monotonic function h. Thus, we do not require that Z be a probability or have any cardinal interpretation. Respondents could all change their elicitations to $1-Z$ or $100Z$; this would not change the value of $E[m_Z(P_Z)|X]$.

But while the benefit of the lower bound approach is that we make only minimal assumptions on how subjective probabilities relate to true beliefs, the resulting empirical test in equation (6) suffers several limitations. First, orderings of lower bounds of $E[m(P)|X]$ across segments do not necessarily imply orderings of its true magnitude.[26] Second, orderings of $E[m(P)|X]$ does not imply orderings of $m(p)$ for all p, which was the statement of the comparative static in $m(p)$ in Corollary (4).[27] Finally, in addition to having limitations as a test of the comparative static, this approach cannot quantify the minimum pooled price ratio. These shortcomings motivate our second approach, which imposes some structure on the relationship between Z and P and allows us to move from lower bounds to point estimates of the distribution of private information.

4.2 Semiparametric Approach: Estimation of the Distribution of Private Information

The goal of the second approach is to estimate the distribution of private information and the minimum pooled price ratio. We then examine the distribution for the presence of thicker upper tails for the rejectees relative to the non-rejectees. With the minimum pooled price ratio, we

---

[24] Our test for the presence of private information is different from the test used by Finkelstein and McGarry [2006] that was initially proposed in Finkelstein and Poterba [2006]. Their approach treats subjective probabilities as "unused observables" that are excluded from the set of variables used by insurance companies for pricing insurance. They infer the presence of asymmetric information if two conditions are satisfied: 1) the subjective probabilities are correlated with the realized loss (conditional on observables) and 2) the subjective probabilities are correlated with insurance purchase (conditional on observables). In contrast, we show that the second requirement is not necessary when using subjective probabilities for identifying private information. Indeed, it would prevent identification of private information amongst rejectees.

[25] In particular, we have not imposed parametric restrictions on the distribution of Z given beliefs P, $f_{Z|P}(Z|P)$.

[26] In Appendix (B.1.3), we provide a stylized example of elicitation error which yields conditions under which orderings of our lower bounds do imply orderings of the true magnitude. Loosely, we require the error in the elicitation to be similar between the two segments under comparison.

[27] Because $E[m(P)]$ is a measure of dispersion, it is invariant to location shifts in the distribution of P (i.e. if $\tilde{P} = P + \eta$, then $E[m(P)] = E\left[m\left(\tilde{P}\right)\right]$). So, testing equation (6) is distinct from analyzing whether rejectees have higher mean risk, as suggested by Proposition 1. Since rejectees have almost universally higher mean risks, testing for higher values of $E[m(P)]$ for the rejectees may a priori be an overly restrictive test of the theory. Since this biases us against finding results consistent with the theory, we do not discuss this interaction in detail. We discuss this further in Appendix B.1, where we show that the minimum pooled price ratio is bounded above (using a Holder inequality) by a term increasing in both the mean, $\Pr\{L|X\}$, and $E[m(P)|X]$.

test whether it is larger for rejectees versus non-rejectees (Corollary 3) and whether it is large (or small) enough to explain (the absence of) rejections for plausible values of the willingness to pay for insurance (Corollary 2). Whereas our nonparametric lower bound approach allowed for an arbitrary relationship between Z and P, we now restrict the way in which elicitations relate to beliefs.

Assumption 3. *Z is distributed with p.d.f./p.m.f. $f_{Z|P}(Z|P;\theta)$ of a known parametric family with unknown parameters θ of finite dimension.*

This restriction limits the extent to which the distribution of Z can vary with P. In our particular specification discussed futher in Section 7.1.1, we will allow $f(Z|P;\theta)$ to capture noise and bias. In addition to Assumption 3, we retain Assumptions 1 and 2 which ensure that $\Pr\{L=1|X,Z,P\}=P$.

With Assumptions 1-3, the joint p.d.f./p.m.f. of the observed variables L[28] and Z (conditional on $X=x$), denoted $f_{L,Z}(L,Z)$, is given by

$$\begin{aligned} f_{L,Z}(L,Z) &= \int_0^1 f_{L,Z}(L,Z|P=p)\, f_P(p)\, dp \\ &= \int_0^1 \left(\Pr\{L=1|Z,P=p\}\right)^L \left(1-\Pr\{L=1|Z,P=p\}\right)^{1-L} f_{Z|P}(Z|p;\theta)\, f_P(p)\, dp \\ &= \int_0^1 p^L (1-p)^{1-L} f_{Z|P}(Z|p;\theta)\, f_P(p)\, dp \end{aligned}$$

where $f_P(p)$ is the unobserved density of the distribution of private information (assumed to be continuous for ease of exposition). The first equality follows by taking the conditional expectation with respect to P. The second equality follows by expanding the joint density of L and Z given P and Assumption 3. The third equality follows from Assumptions 1 and 2.

Assumption 3 allows us to estimate θ, as opposed to an arbitrary two-dimensional continuous function, $f_{Z|P}$. This allows us to estimate both θ and f_P using the observed joint distribution of the data, $f_{L,Z}(L,Z)$. We discuss identification in general and for our particular functional form choice in Appendix B.2. But the order condition is straightforward. The observed joint distribution, $f_{L,Z}$, contains two continuous functions of Z (one for $L=1$ and another for $L=0$). We use one of these functions to identify f_P and another to identify θ. While we have imposed a functional form on θ, we do not impose a functional form on the distribution of private information, f_P. In practice, we flexibly approximate f_P and estimate all parameters (both θ and the approximating parameters for f_P) using maximum likelihood.

Given estimates of the distribution of private information, we translate these into measures of the barrier to trade imposed by private information. Recall from Corollary 2 that this magnitude is fully characterized by the minimum pooled price ratio, $\inf_{p\in\Psi\setminus\{1\}} T(p)$, where $T(p)$ can be calculated at each p using estimates of $E[P|P\geq p]$ derived from the estimated distribution of

[28] For notational brevity, we let L also denote the binary indicator that the event L occurs, $1\{L\}$.

private information. One remaining limitation is that for values of p in the upper quantiles of $F(p)$, $E[P|P \geq p]$ is an extreme value that is not well-identified, since the expectation is taken with respect to a smaller and smaller effective sample as p increases. However, for a fixed quantile τ, estimates of the minimum pooled price ratio over $\hat{\Psi}_\tau = \left[0, F^{-1}(\tau)\right] \cap (\Psi \backslash \{1\})$ are continuously differentiable functions of the MLE parameter estimates of $F(p)$ for $p \leq F^{-1}(\tau)$.[29] So, derived MLE estimates of $\inf_{p \in \hat{\Psi}_\tau} T(p)$ are consistent and asymptotically normal. Thus, our approach is to construct the minimum pooled price ratio over $\hat{\Psi}_\tau$ for a fixed $\tau < 1$. We then assess robustness to the choice of τ.

While our motivation for restricting attention to $\hat{\Psi}_\tau$ as opposed to Ψ is primarily because of statistical limitations, Remark 1 in Section 2.2 provides an economic rationale for why $\inf_{p \in \hat{\Psi}_\tau} T(p)$ may not only be a suitable substitute for $\inf_{p \in \Psi \backslash \{1\}} T(p)$ but also may actually be more relevant if firms face frictions to setting up contracts. If contracts must attract a non-trivial fraction $1 - \tau$ of the market in order to be viable, then $\inf_{p \in \hat{\Psi}_\tau} T(p)$ characterizes the barrier to trade imposed by private information.

In short, the semiparametric approach makes an additional parametric assumption on the statistical relationship between elicitations Z and beliefs P which allows us to estimate the distribution of private information and implement the tests outlined in Section (3.3).

5 Setting and Data

We employ our empirical approach to ask whether private information can explain rejections in three non-group insurance market settings: long-term care, disability, and life insurance.

5.1 The Three Non-Group Market Settings

Long-term care (LTC) insurance insures against the financial costs of nursing home use and professional home care. Expenditures on LTC represent one of the largest uninsured financial burdens facing the elderly. LTC expenditures in the US totaled over \$135B in 2004 (CBO [2004]), and expenditures are heavily skewed: less than half of the population will ever enter a nursing home in their life. Despite this, the LTC insurance market is small, with roughly 4% of all nursing home expenses paid by private insurance, compared to 31% paid out-of-pocket (CBO [2004]).[30]

The private disability insurance protects against the lost income resulting from a work-limiting disability. It is primarily sold through group settings, such as one's employer; more than 30% of private workers have group-based disability policies. In contrast, the non-group market is quite small. Only 3% of non-government workers own a private non-group disability

[29] Non-differentiability could hypothetically occur at points where the infimum is attained at distinct values of p.

[30] Medicaid pays for nursing home stays provided one's assets are sufficiently low and is a substantial payer of long-term stays.

policy, most of whom are self-employed or professionals who do not have access to employer-based group policies (ACLI [2010]).[31]

Life insurance provides payments to ones' heirs or estate upon death, insuring lost income or other expenses. Policies either expire after a fixed length of time (term life) or cover one's entire life (whole life). In contrast to the non-group disability and LTC markets, the private non-group life insurance market is quite big. More than half of the adult US population owns life insurance. 54% of these policies are sold in the non-group market. 43% of these are term policies, while the remaining 57% are whole life policies (ACLI [2010]).

Not everyone can purchase insurance in these three non-group markets. As mentioned in the introduction, Murtaugh et al. [1995] estimates that 12-23% of 65 year olds have a health condition which would cause them to be rejected by LTC insurers. In life and disability insurance, we know of no formal studies documenting the prevalence of rejections, but our review of underwriting guidelines and conversations with underwriters in these markets establish a prevalence of rejections based on certain pre-existing conditions that we discuss in more detail in Section 5.2.2.

Insurance companies in these markets are not legally prevented from charging higher prices to reflect actuarial differences in risk.[32] They do face some regulation. Capital levels must be maintained to prevent policy default. Also, they are limited in the extent to which policy prices can be raised over time after purchase, which is intended to prevent exploitative price increases on those who have already sunk payments into a policy. But no regulation prevents insurance companies from offering risk-adjusted prices to those who are currently rejected in these three market settings.[33]

Previous research has found minimal or no evidence of private information using the revealed preference approach in these settings. In life insurance, Cawley and Philipson [1999] find no evidence of adverse selection. He [2009] revisits this with a different sample focusing on new purchasers and does find evidence of small amounts of adverse selection. In long-term care, Finkelstein and McGarry [2006] find direct evidence of private information by showing subjective probabilities are correlated with subsequent nursing home use. However, they find no evidence that this private information leads to adverse selection in the form of a correlation between insurance purchase and subsequent losses in the LTC insurance market.[34] To our knowledge, there is no previous study of private information in the non-group disability market.

[31] In contrast to health insurance where the group market faces significant tax advantages, group disability policies are taxed. Either the premiums are paid with after-tax income, or the benefits are taxed upon receipt.

[32] The Civil Rights Act does prevent purely racial discrimination in pricing.

[33] Interviews with underwriters in these markets also suggest that fear of regulation is not an issue in preventing charging a higher price to those currently rejected.

[34] They suggest heterogeneous preferences, in which good risks also have a higher valuation of insurance, can explain why private information doesn't lead to adverse selection.

5.2 Data

Both of our approaches have the same data requirements. The ideal dataset would contain, for each setting, four pieces of information:

1. Loss indicator, L, corresponding to a commonly insured loss
2. Agents' subjective probability elicitation, Z, about this loss
3. The set of public information, X, which would be observed by insurance companies in setting contract terms
4. The classification, Θ^{Reject} and $\Theta^{NoReject}$, of who would be rejected if they applied for insurance

Our data source for the loss, L, subjective probabilities, Z, and public information X, come from years 1993-2008 of the Health and Retirement Study (HRS). The HRS is an individual-level panel survey of individuals over 55 and their spouses (included regardless of age). It contains a rich set of health and demographic information, along with subjective probability elicitations about future events.

To construct the rejection classification, we primarily rely on insurance company underwriting guidelines which are used by underwriters and often provided to insurance agents with the purpose of preventing those with rejection conditions from applying. We supplement this information with interviews with insurance underwriters. We discuss each piece of our data in further detail.

5.2.1 Loss Variables and Subjective Probability Elicitations

The HRS contains three subjective probability elicitations about future events which correspond to a commonly insured loss in each of our settings:

> Long-Term Care: *"What is the percent chance (0-100) that you will move to a nursing home in the next five years?"*
>
> Disability: *"[What is the percent chance] that your health will limit your work activity during the next 10 years?"*
>
> Life: *"What is the percent chance that you will live to be AGE or more?"* (where AGE$\in$ {75,80,85,90,95,100} is respondent-specific and chosen to be 10-15 years from the date of the interview)

Figures 1(a,b,c) display histograms of these responses (divided by 100 to translate into probabilities).[35] As has been noted in previous literature using these subjective probabilities (Gan et al.

[35] We use the sample selection described in Subsection (5.2.4)

[2005]; Finkelstein and McGarry [2006]), these histograms highlight why it would be problematic to view these as true beliefs. Many respondents report 0, 50, or 100. Taken literally, responses of 0 or 100 imply an infinite degree of certainty, which is difficult to believe. We find it more likely that respondents who report focal point values are responding on more of an ordinal scale (e.g. high, medium, low) as opposed to having a literal probabilistic interpretation. Our lower bound approach remains agnostic on the way in which focal point responses relate to true beliefs.[36] Our parametric approach will take explicit account of this focal point response bias, discussed further in Section 7.1.1.

Corresponding to each subjective probability elicitation, we construct binary indicators of the loss, L. In long-term care, L denotes the event that the respondent enters a nursing home in the subsequent 5 years. In disability, L denotes the event that the respondent reports that their health limits their work activity in the subsequent 10-11 years.[37] In life, L denotes the event that the respondent dies before AGE, where AGE$\in$ {75,80,85,90,95,100} corresponds to the subjective probability elicitation, which is 10-15 years from the survey date.[38]

5.2.2 Rejection Classification

Not everyone can purchase insurance in these three non-group markets. An ideal dataset would classify our entire samples into rejectees and non-rejectees. Practically, this requires knowing the conditions that cause rejection and matching these conditions to those reported in the HRS. As we discuss below, this match faces limitations which lead us to construct a third group, "Uncertain", which allows us to be relatively confident in our classification of rejectees and non-rejectees.

To identify conditions that lead to rejection, we obtain underwriting guidelines used by underwriters and provided to insurance agents for use in screening applicants. An insurance company's underwriting guidelines provide a list of conditions for which underwriters are instructed to not offer insurance at any price. These guidelines are not publicly available, which limits our ability to obtain this information. The extent of our access varies by market: In long-term care, we obtain a set of guidelines used by an insurance broker from 18 of the 27 largest long-term care insurance companies collectively representing over 95% of the US market.[39] In disability and life, we obtain several underwriting guidelines and supplement this information with interviews with underwriters at several major insurance companies.

To match these conditions to our dataset, we use the detailed health and demographic information available in the HRS to identify individuals with conditions which would lead them

[36] For our empirical specification, we will include indicators for focal point responses

[37] Our loss variable is necessarily defined as 11 years for those in the AHEAD 1993 wave 2 group because the panel does not provide responses exactly 10 years from 1993. Our results are robust to the exclusion of this group.

[38] We construct the corresponding elicitation to be $100\% - Z^{live}$ where Z^{live} is the survey elicitation for the probability of living to AGE.

[39] These guidelines display broad consistency in the rejection practices across firms. We thank Amy Finkelstein for making this broker-collected data available.

to be rejected. While the HRS contains a relatively comprehensive picture of respondents' health, sometimes the conditions which would lead to rejection are too precise too be accurately matched in the HRS. For example, individuals with advanced stages of lung disease would be unable to purchase life insurance; however, the HRS only provides information for the presence of a lung disease.

We exercise caution in performing this match by constructing a third classification, "Uncertain", to which we classify those who may be rejected, but for whom data limitations prevent a solid assessment. This allows us to be relatively confident in our classification of rejectees and non-rejectees. We present our lower bound estimates for all three classifications.[40]

Table 1 presents the list of conditions for the rejection and uncertain classification, along with the frequency of each condition in our sample (using the sample selection outlined below in Section 5.2.4). In long-term care, activity of daily living (ADL) restrictions (e.g. needs assistance walking, dressing, using toilet, etc.), any previous stroke, any previous home care, and anyone over the age of 80 would be rejected. In disability, a back condition, obesity (40+ BMI), and doctor-diagnosed psychological conditions such as depression or bi-polar would lead to rejection. In life, individuals with a past stroke or current cancer would be rejected. We classify individuals with these conditions as rejected in their respective markets.

Table 1 also lists the conditions leading to an uncertain classification in each market. In addition to specific conditions for which the HRS data is too coarse, we also attempt to capture the presence of rarer conditions not asked in the HRS (e.g. Lupus would lead to rejection in LTC, but is not explicitly reported in the HRS). To do so, we take advantage of a question in the HRS which asks respondents if they have any additional major health problems which were not asked about in the survey. We classify individuals reporting yes to this question as Uncertain.

5.2.3 Public Information

Our ideal dataset would contain all information that insurance companies would use in pricing contracts. For non-rejectees, this is a straightforward requirement which involves analyzing existing contracts. But for rejectees, we must make an assumption about how insurance companies would price contracts to these people if they were to offer them. Our preferred approach is to assume insurance companies price rejectees separately from those to whom they currently offer contracts, but use a relatively similar set of public information. Thus, our primary data requirement is the public information currently used by insurance companies in pricing insurance.

The HRS contains an extensive set of health, demographic, and occupation information which allows us to approximate the set of information which insurance companies use in pricing insurance.[41] The quality of this approximation varies by market. For long-term care, we replicate the information set of the insurance company quite well. For example, perhaps the most obscure

[40]For brevity, we do not present results from our semiparametric approach for the uncertain group.

[41]We are not the first to note the ability of the HRS to replicate the information used by insurance companies in pricing; for LTC, see Finkelstein and McGarry [2006] and for Life, see He [2009].

piece of information that is acquired by some LTC insurance companies is an interview in which applicants are asked to perform word recall tasks to assess memory capabilities; the HRS conducts precisely this test with survey respondents. In disability and life, we replicate most of the information used by insurance companies in pricing. One caveat is that insurance companies will sometimes perform tests, such as blood and urine tests, which we will not observe in the HRS. Conversations with underwriters in these markets suggest these tests are primarily to confirm application information, which we can approximate quite well with the HRS. But, we cannot rule out the potential that there is additional information which can be gathered by insurance companies in the disability and life settings.[42]

In addition to our preferred specification which includes variables used in pricing, we also assess the robustness of our estimates to alternative sets of controls.[43] This is for two reasons. First, although we have been careful in constructing the pricing controls, it may not be a perfect representation of the set of information used in pricing. Second, we do not want our conclusions for the amount of private information for the rejectees to depend on an assumption of how insurance companies would hypothetically use information to price their contracts. We therefore perform our analysis for three increasing sets of public information:

1. "Age and Gender": A baseline specification with fully saturated age-by-gender dummies
2. "Pricing Controls": Includes all variables currently used in pricing
3. "Extended Controls": Includes all Pricing Controls plus a large set of additional variables not currently used in pricing but potentially related to the outcome

The age and gender specification provides a baseline. The pricing controls assumes insurance companies would price similarly for those facing rejection. This is our preferred specification. The extended controls specification adds a rich set of interactions between health conditions and demographic variables that could be, but are not currently, used in pricing insurance.

We conduct the lower bound approach for all three sets of controls. For brevity, we focus exclusively on our preferred specification of pricing controls for our semiparametric approach.[44]

The variables used in the pricing and full controls specifications for each market are presented in Table 2. In LTC, our preferred specification includes age, age squared, and gender interactions; indicators for various health conditions; ADL restrictions; and performance on a word recall

[42] In LTC, insurance companies are legally able to conduct tests, but it is not common industry practice.

[43] While it might seem intuitive that including more controls would reduce the amount of private information, this need not be the case. To see why, consider the following example of a regression of quantity on price. Absent controls, there may not exist any significant relationship. But, controlling for supply (demand) factors, price may have predictive power for quantity as it traces out the demand (supply) curve. Thus, adding controls can increase the predictive power of another variable (price, in this case). Of course, conditioning on additional variables X' which are uncorrelated with L or Z has no effect on the population value of $E\left[m\left(P\right)|X \in \Theta\right]$.

[44] Because the extended controls specification includes a lot of variables, we risk over-fitting the data. As we discuss in Section 6.1, this does not pose an insurmountable problem for the lower bound approach. However, it would pose a problem for the semiparametric approach, and thus provides another reason for our exclusive focus on the preferred pricing specification.

test. Our extended controls specification adds full interactions for age and gender, along with interactions of 5 year age bins with measures of health conditions, indicators for the number of living relatives (up to 3), census region, and income deciles. For disability, our preferred specification includes age, age squared, and gender interactions; indicators for self employment and various health conditions; BMI; and wage decile. Our extended controls specification adds full interactions of age and gender; full interactions of wage decile, part time status indicator, job tenure quartile, and self-employment indicator; interactions between 5 year age bins and various health conditions and BMI; full interactions of job characteristics (e.g. "job requires heavy lifting"); and full interactions of 5 year age bins and census region. For life, our preferred specification includes age, age squared, and gender interactions, smoking status, indicators for the death of a parent before age 60, BMI, income decile, and indicators for a set of health conditions. We also include a set of indicators for the years between the survey date and the AGE corresponding to the loss.[45] Our extended controls specification adds full interactions of age and gender; full interactions between age and the AGE in the subjective probability question; interactions between 5 year age bins and smoking status, income decile, census region, and various health conditions;[46] BMI; and an indicator for death of a parent before age 60.

5.2.4 Sample Selection

For each sample, we begin with years 1993-2008 of the HRS. Our selection process varies across each of the three market settings due to data constraints. Table 3 presents the summary statistics for each sample.

LTC For LTC, we exclude individuals for whom we cannot follow for a subsequent five years to construct our loss indicator variable; years 2004-2008 are used but only for construction of the loss indicator. Also, we exclude individuals who currently reside in a nursing home. Our primary sample consists of 9,051 observations from 4,418 individuals for our no reject sample, 10,108 observations from 3,215 individuals for the reject sample, and 10,690 observations from 5,190 individuals for the uncertain sample. In each of our samples, we include multiple observations for a given individual (which are spaced roughly two years apart) to increase power. All standard errors will be clustered at the household level.

In addition to our primary sample, we will report results for our nonparametric lower bounds using a sample that excludes individuals who own long-term care insurance (roughly 13% of remaining sample) to ensure we estimate private information inherently held by the individual which is not the effect of insurance contract choice on subsequent utilization (a.k.a. "moral

[45] We also include this in our age & gender and extended control specifications for life.

[46] Although the HRS asks whether respondents have (non-basal cell) cancer, it only asks which organ the cancer occurs in the 2nd wave (1993/1994) of the survey. In the robustness section, we will consider an additional extended controls specification for life insurance which uses data only from these years and includes a full set of cancer organ indicators (50+ indicators).

hazard").[47]

Rejectees differ from non-rejectees on many dimensions. They are older (average age of 79 versus 71), more likely to have health conditions such as arthritis, diabetes, and high blood pressure, and have a 17% entry rate to a nursing home in the subsequent 5 years, compared to an entry rate of only 4% for those not facing rejection. But while they are higher risk, on average they still have less than a 20% chance of going to a nursing home in the next five years. This suggests they still face significant of un-realized risk.

Disability For disability, we begin with the set of individuals up to age 60 who are currently working and report no presence of work-limiting disabilities. To construct the corresponding loss realization, we limit the sample to individuals who we can observe for a subsequent 10 years (years 2000-2008 are used solely for the construction of the loss indicator). Our final sample consists of 2,540 observations from 1,480 individuals for our no reject classification, 2,216 observations from 1,280 individuals for our reject classification, and 3,757 observations from 1,929 individuals for our uncertain classification.[48]

Rejectees differ from non-rejectees on many dimensions. They are more likely to have high blood pressure, diabetes, and arthritis,[49] and have a higher risk of experiencing a work-limiting disability (44.1% versus 15.6%). But, similar to LTC, not everyone with a rejection condition will experience a work-limiting disability in the subsequent 10 years, which again suggests they face unrealized risk.

Life For our life sample, we restrict to individuals who we are able to follow through the age corresponding to the subjective probability elicitation 10-15 years in the future, so that years 2000-2008 are used solely for the construction of the loss indicator. For example, if a 63 year old is asked about the probability they will live to age 75, we require being able to see this person for a subsequent 12 years in the survey. Our final sample consists of 2,689 observations from 1,720 individuals for our no reject classification, 2,362 observations from 1,371 individuals for our reject classification, and 6,800 observations from 4,270 individuals for our uncertain classification.

Similar to LTC, we include those who own life insurance in our primary sample (64% of the sample) but present results excluding this group for robustness. Similar to our other settings, the rejectees are older, sicker, and more likely to experience the loss than non-rejectees.

[47]While one might be tempted to control for the purchase of insurance or the contract characteristics, this would be misguided. If agents with different beliefs sort into different contracts, controlling for contract choice could lead to a finding of no private information. Insurance purchase is a potentially endogenous response to the presence of private information and thus should not be included as a control variable.

[48]Ideally, we would also test the robustness of our results using a sample of those who do not own disability insurance, but unfortunately the HRS does not ask about disability insurance ownership.

[49]Diabetes and arthritis may lead to rejection, so those without a rejection condition but with one of these two conditions are placed in the uncertain classification

Discussion There are several broad patterns across our three samples. First, a sizable fraction of the sample would be rejected in each setting. Because the HRS primarily surveys older individuals, our sample is older (and therefore sicker) than the average purchaser in each market. This is a primary benefit of the HRS; it allows us to obtain a significant sample size of rejectees. But, it is important to understand that this fraction of rejectees is not a measure of the fraction of the applicants in each market that would be rejected.

Second, many rejectees own insurance. These individuals could (and perhaps should) have purchased insurance prior to being stricken with their rejection condition. Also, they may have been able to purchase insurance in group markets through their employer, union, or other group which has less stringent underwriting requirements.

Third, rejectees differ from non-rejectees on many dimensions; their older, sicker, and have a higher probability of experiencing the loss. This is consistent with Proposition 1 which showed that higher risk distributions are more likely to satisfy the no-trade condition.

5.2.5 Relation to Ideal Data

The extent to which our data resembles an ideal dataset varies by market. In general, we approximate the ideal dataset quite well, aside from our necessity to classify a relatively large fraction of our sample as uncertain. In disability and in life, we are able to classify a smaller fraction of the sample as rejected or not rejected as compared with LTC. Also, for disability and life we rely on a smaller set of underwriting guidelines (along with underwriter interviews) to obtainn rejection conditions, as opposed to LTC where we obtain an fairly large fraction of the underwriting guidelines used in the market. In disability and life we also do not observe medical tests which may be used by insurance companies to price insurance (although our conversations with underwriters suggest this is primarily to verify application information, which we approximate quite well using the HRS). In contrast, in LTC we are able to classify a relatively large fraction of the sample, are able to closely approximate the set of public information, and are able to assess the robustness of our results to the exclusion of those who own insurance to remove the potential impact of a moral hazard channel driving any findings of private information. While re-iterating that all three of our samples approximate our ideal dataset quite well, our LTC sample is arguably the best of our three samples.

6 Lower Bound Estimation

We now turn to the estimation of lower bounds of the average magnitude of private information, $E\left[m_Z\left(P_Z\right)|X\right]$, outlined in Section 4.1.

6.1 Empirical Estimation

We estimate $E[m_Z(P_Z)|X]$ separately for each setting (e.g. LTC), sample (e.g. Reject), and specification (e.g. Price Controls). Implementation involves two steps. First, while the approach is theoretically nonparametric, in practice we choose a flexible parametric approximation for $\Pr\{L|X,Z\}$. Second, we must make an assumption that allows us to reduce the dimensionality of the way in which the distribution P_Z varies with X to enable estimation of the distribution of P_Z and $m_Z(p)$ for each X.

To approximate $P_Z = \Pr\{L|X,Z\}$ for our age/gender and price controls specifications, we use a probit specification,

$$\Pr\{L|X,Z\} = \Phi(\beta X + \Gamma(age, Z))$$

where X are our control variables and $\Gamma(age, Z)$ captures the effect of Z on $\Pr\{L|X,Z\}$.[50] This form allows the affect of Z to vary with age (note that age is already included in X).[51] We approximate $\Gamma(age, Z)$ using full interactions of functions of Z and functions of age. For Z, we use second-order Chebyshev polynomials plus separate indicators for focal point responses at $Z = 0$, 50, and 100. We use a linear function of age. Our approximation of $\Gamma(age, Z)$ is then given by the full set of these interactions (whose coefficients are to be estimated). All results are robust to the inclusion of additional or fewer polynomials in Z or age, or the use of a linear or logit specification, as opposed to the probit. For our extended controls specification, the high dimensionality of X leads us to use a linear specification, $\Pr\{L|X,Z\} = \beta X + \Gamma(age, Z)$. We use the same approximation for Γ.[52]

Estimating $m_Z(p) = E[P_Z|P_Z \geq p, X] - p$ requires estimating the entire distribution of P_Z at each possible value of X. To make this feasible, we adopt an assumption for how the distribution of P_Z varies with X: conditional on ones age and rejection classification, the distribution of residual private information implied by Z, $P_Z - E[P_Z|X]$, does not vary with X. This allows observable variables affect the mean but not the shape of the distribution of P_Z (conditional on age and rejection classification).[53] We then estimate the conditional expectation, $m_Z(p) = E[P_Z|P_Z \geq p, X] - p$ using the empirical distribution of $P_Z - E[P_Z|X]$ within each age grouping.

After estimating $m_Z(p)$, we construct its average using the empirical distribution of P_Z,

[50] One could allow the the effect of Z to vary with other covariates. Our results are robust to much simpler specifications (e.g. assuming $\Gamma(age, Z) = \gamma Z$) and richer specifications, such as including gender in Γ. Note that although the coefficients for the effect of Z on $\Pr\{L|X,Z\}$ is restricted via functional form, we are not necessarily restricting the estimated distribution of $\Pr\{L|X,Z\}$, since the distribution of Z can (and does) vary with X.

[51] In our LTC Reject Sample, we also include full interactions between Γ and an indicator for having a rejection health condition; this allows Γ to vary differentially for those over age 80 with no other rejection conditions besides age>80.

[52] Of course, the switch from the probit to linear specification leads Γ to have a different interpretation.

[53] This assumption is only required to arrive at a point estimate for $E[m_Z(P_Z)|X \in \Theta]$, and is not required to test for the presence of private information (i.e. whether $\Gamma = 0$). Also, our results for $E[m_Z(P_Z)|X \in \Theta]$ are robust to alternative assumptions, such as assuming the residual distribution does not vary with X within 5 year age bins.

yielding $E\left[m_Z\left(P_Z\right)|X \in \Theta\right]$, where Θ is a given sample (e.g. LTC rejectees). For each market, we then construct the difference between the reject and no reject estimates,

$$\Delta_Z = E\left[m_Z\left(P_Z\right)|X \in \Theta^{Reject}\right] - E\left[m_Z\left(P_Z\right)|X \in \Theta^{NoReject}\right]$$

and test whether we can reject a null hypothesis that $\Delta_Z \leq 0$. While choosing Θ to be an entire sample (e.g. all LTC rejectees) increases power, we will also construct estimates for subgroups (e.g. age groupings) of the rejectees and non-rejectees.

6.2 Statistical Inference

Statistical inference for $E\left[m_Z\left(P_Z\right)|X \in \Theta\right]$ for a given sample Θ and for Δ_Z is straightforward, but requires a bit of care to cover the possibility of no private information. In any finite sample, our estimates of $E\left[m_Z\left(P_Z\right)|X \in \Theta\right]$ will be positive (Z will always have some predictive power in finite samples). Provided the true value of $E\left[m_Z\left(P_Z\right)|X \in \Theta\right]$ is positive, the bootstrap provides consistent, asymptotically normal, standard errors for $E\left[m_Z\left(P_Z\right)|X \in \Theta\right]$ (Newey [1997]). But, if the true value of $E\left[m_Z\left(P_Z\right)|X \in \Theta\right]$ is zero (as would occur if there were no private information amongst those with $X \in \Theta$), then the bootstrap distribution is not asymptotically normal and does not provide adequate finite-sample inference.[54] We therefore supplement the bootstrap with a Wald test which restricts $\Gamma\left(age, Z\right) = 0$.[55] This tests for the presence of private information. We report results from both the Wald test and the bootstrap.

We conduct inference on Δ_Z in a similar manner. To test the null hypothesis that $\Delta_Z \leq 0$, we construct conservative p-values by taking the maximum p-value from two tests: 1) a Wald test of no private information held by the rejectees, $E\left[m_Z\left(P_Z\right)|X \in \Theta^{Reject}\right] = 0$, and 2) the p-value from the bootstrapped event of less private information held by the rejectees, $\Delta \leq 0$.[56]

6.3 Graphical Results for $P_Z - E\left[P_Z|X\right]$

We begin with graphical evidence of the predictive power of subjective probability elicitations. Figures 2(a,b,c) plot the estimated distribution of $P_Z - E\left[P_Z|X\right]$ aggregated by rejection classification for the rejectees and non-rejectees, using our preferred pricing control specification.[57]

Consistent with the hypothesis that rejectees are better informed about whether or not they would experience the loss, the distribution of $P_Z - E\left[P_Z|X\right]$ is more dispersed for the rejectees relative to those served by the market in all three market settings we consider. As we now show,

[54] In this case, $\hat{\Gamma} \to 0$ in probability, so that estimates of the distribution of $P_Z - E\left[P_Z|X\right]$ converge to zero in probability (so that the bootstrap distribution converges to a point mass at zero).

[55] The event $\Gamma\left(age, Z\right) = 0$ in sample Θ is equivalent to both the event $\Pr\left\{L|X, Z\right\} = \Pr\left\{L|X\right\}$ for all $X \in \Theta$ and the event $E\left[m_Z\left(P_Z\right)|X \in \Theta\right] = 0$.

[56] More precise p-values would be a weighted average of these two p-values, where the weight on the Wald test is given by the unknown quantity $\Pr\left\{E\left[m_Z\left(P_Z\right)|X \in \Theta^{Reject}\right] = 0|\Delta \leq 0\right\}$. Since this weight is unknown, we construct conservative p-values robust to any weight in $[0, 1]$.

[57] Subtracting $E\left[P_Z|X\right] = \Pr\left\{L|X\right\}$ allows for simple aggregation across X within each sample.

this translates into higher estimates of our lower bounds on the average magnitude of private information.

6.4 Lower Bound Results

The top row of Table 4 provides the estimates of Δ_Z. Across all specifications and all market settings, we estimate larger lower bounds on the average magnitude of private information for the rejectees relative to those served by the market. These differences are all statistically significant at the 1% level (third row of Table 4). Consistent with the theory, this suggests private information imposes a greater barrier to trade for rejectees relative to those served by the market.

The lower sets of rows in Table 4 report the estimated magnitude of private information for each classification (No Reject, Reject, and Uncertain), along with their standard errors and p-values for the presence of private information. We discuss these details by market.

LTC In long-term care, we find significant evidence of private information amongst the rejectees, with estimated magnitudes of 0.0286 ($p < 0.001$) in our preferred specification.[58] In contrast, we find no statistically significant evidence of private information held by the non-rejectees (0.0041, $p = 0.433$ for our preferred specification). The estimates are quite similar for different control specifications: all estimates lie within an estimated standard error (0.004).

Our finding that the subjective probabilities are significant predictors of subsequent nursing home use is consistent with the empirical results of Finkelstein and McGarry [2006]. However, splitting the sample by the rejection classification, our results reveal that this private information is primarily held by those who would be rejected. This provides a new explanation for the absence of a positive correlation found in Finkelstein and McGarry [2006] between insurance purchase and realized claims in the long-term care market: insurance companies choose to not sell insurance to those whose observable characteristics indicate they may have significant amounts of private information.

Disability In disability, we find significant evidence of private information held by both the rejectees and non-rejectees. In our preferred specification, we estimate magnitudes of 0.0512 ($p < 0.01$) for the rejectees and 0.0257 ($p < 0.01$) for non-rejectees, leading to an estimated difference of $\Delta_Z = .0255$ ($p = 0.006$). The results are robust to the inclusion of additional controls: the extended controls specification yields statistically indistinguishable results from our preferred specification (.0234, $p = .018$). Our age and gender specification leads to slightly

[58]Because the estimated magnitudes are lower bounds, we do not focus our discussion on their absolute magnitudes. But the interpretation is straightforward: the estimated magnitude of 0.0286 implies that $E\left[P_Z|P_Z \geq p\right]$ differs from p by 0.0286 on average, which implies that, on average, the average predicted probability of a loss (given Z) for worse risks differs from ones' own risk by 2.86pp.

higher estimated magnitudes for the rejectees of 0.0737, but not significantly different from our price controls specification.

To the best of our knowledge, these estimates provide the first evidence of private information in the non-group private disability insurance market. Although many factors could be driving this market's small size (only 3% of private employees own a non-group private disability policy (ACLI [2010])), private information may be a contributing factor.

Life In life, we find significant evidence of private information amongst the rejectees with magnitudes of 0.0587 ($p < 0.001$) in our preferred specification. The magnitudes are quite similar with the extended controls (0.0604, $p < 0.001$). In contrast, we find smaller magnitudes for the non-rejectees (0.0250, $p = 0.119$) and cannot reject the null hypothesis of no private information. Yet our point estimate of 0.025 remains similar to the statistically significant estimate conditional on age and gender alone (0.0310, $p = 0.01$). So, we also cannot rule out the presence of some private information for the non-rejectees.

Our finding of minimal evidence of private information for those served by the market is consistent with existing empirical work in life insurance using the revealed preference approach (Cawley and Philipson [1999], He [2009]). But, while Cawley and Philipson [1999] suggest their results imply that there is no evidence of asymmetric information afflicts the life insurance market, our results suggest much of agents' private information is held by those who would be rejected by insurance companies. So although private information may not significantly affect the adverse selection of observed contracts, it may simply pose a barrier to the existence of the market itself.

Uncertain Classification The estimated magnitudes for the uncertain classification generally fall between the estimates for the rejection and no rejection groups, as indicated by the bottom set of rows in Table 3. In general, our theory does not have a prediction for the uncertain group. However, if $E\left[m_Z\left(P_Z\right)|X\right]$ takes on similar values for all rejectees (e.g. $E\left[m_Z\left(P_Z\right)|X\right] \approx m^R$) and non-rejectees (e.g. $E\left[m_Z\left(P_Z\right)|X\right] \approx m^{NR}$), then linearity of the expectation implies

$$E\left[m_Z\left(P_Z\right)|X \in \Theta^{\text{Uncertain}}\right] = \lambda m^R + (1-\lambda)\, m^{NR} \tag{9}$$

where λ is the fraction in the uncertain group who would be rejected. Thus, it is perhaps not unreasonable to have expected $E\left[m_Z\left(P_Z\right)|X \in \Theta^{\text{Uncertain}}\right]$ to lie in between our estimates for the rejectees and non-rejectees, as we find. Nevertheless, we have no theoretical reason to suppose the average magnitude of private information is constant within rejection classification; thus this should be viewed only as a potential method for interpreting the results, not as a robust prediction of the theory.

6.5 Subsample Analysis

The results in Table 4 aggregate across all observables, X, within each rejection classification. While this aggregation improves statistical power, it is important to also examine the results within subgroups to test whether the rejectees have higher magnitudes of private information conditional on observable variables.[59] In this section, we examine age-based subgroups.

Figures 3(a,bi,bii,c) plot the estimates of $E\left[m_Z\left(P_Z\right)|X\in\Theta^{age,rejectclass}\right]$ separately for each age and rejection classification.[60] In all three settings, we find larger estimates for the rejectees versus non-rejectees conditional on age.

In long-term care, we can also more closely examine the specific rejection practices based on age. LTC insurers reject applicants above age 80 regardless of health conditions (such as ADL restrictions or a past stroke). Figure 3a plots the lower bound estimates at each age, separately reporting estimates for those with and without rejection health conditions above age 80. The results show that the estimates for those without rejection conditions increases at ages nearing 80. Indeed, an individual at age 81 with no other rejection conditions (but who would be rejected based on age) has a very similar magnitude of private information to a 65 year old who would be rejected. In short, the results provide a picture of why insurance companies automatically reject individuals beginning at age 80 as opposed to other age cutoffs.

In life (Figure 3c), we find larger estimates for the rejectees across the age spectrum. For disability, we also generally find larger estimates (Figures 3bi and 3bii), although the difference between the reject and no reject estimates appears to be increasing in age.[61] In short, we find larger estimates for rejectees conditional on age.

6.6 Robustness

6.6.1 Insurance Ownership Sample Selection

Our primary results in Table 4 do not exclude individuals who own insurance. If insurance choice affects the risk of experiencing the loss, then differential insurance ownership could cause a finding of private information. We test the robustness of our results to this potential bias by

[59] In addition to analyzing subgroups as a finer test of the theory, one might also worry that aggregation masks other potential drivers of the magnitude of private information aside from the presence of rejection conditions. In particular, the rejectees are generally older than the non-rejectees. If older people naturally, for some reason, have more private information, irrespective of whether or not they would be rejected, then we would estimate $\Delta_Z>0$ in aggregate, even though it may not be the case that $\Delta_Z>0$ conditional on age.

[60] We use our preferred specification, which is quite flexible in age and allows Γ to vary with age. Figures 3(a,bi,bii,c) provide bootstrapped standard errors, which are consistent as long as $\Gamma\neq 0$. In general, we cannot reject the null hypothesis that $\Gamma=0$ on a subsample consisting of one specific age, but our results in Table 4 do reject $\Gamma=0$ at all ages for all but the LTC and life no reject samples.

[61] We present separate results for males and females in Figures 3bi and 3bii because of the changing gender composition of the sample over time. Individuals below age 55 are included in the HRS only if they have a spouse above age 55. Thus, we have relatively more females below age 55. But, as shown in these figures, we generally find larger estimates for rejectees conditional on age and gender. We have also examined LTC and Life by age & gender and the results again show larger magnitudes for the rejectees conditional on age and gender.

restricting to those who do not own insurance in our LTC and Life samples. Table 5 presents these results.

For LTC, our estimates of Δ_Z with the restricted sample are almost identical to the preferred specification estimates (0.0245 versus 0.0257). Across each group (reject, no reject uncertain), our estimated results for $E\left[m_Z\left(P_Z\right)|X \in \Theta\right]$ excluding those who own insurance are also nearly identical. In particular, we still cannot reject the null hypothesis of no private information for the non-rejectees ($p = 0.828$).

For Life, our estimate of Δ_Z with the restricted sample is smaller (0.011 versus 0.0328), and no longer statistically significant. But closer inspection reveals that the drop in magnitude is primarily driven by a larger, yet still statistically insignificant estimate for the non-rejectees (0.0377, $p = 0.233$).[62] For the rejectees, the estimates lie within an estimated standard error of our preferred estimate, 0.0491 versus 0.0587, when we exclude those who own insurance. Thus in both LTC and life, we find our results are robust to the inclusion of those with insurance.

6.6.2 Organ Controls for Life Specification

Our specifications for life insurance did not include controls for the affected organ of cancer sufferers in the reject sample. Although later years of the survey do not specify the organ of the cancer, it is provided in the 1993/4 wave of the survey. In Table 6, we report the results from our primary specification (all years) and the results from a specification restricted to years 1993/1994 which includes a full set of 54 indicators for the affected organ added to our set of extended controls. Our finding of significant private information amongst rejectees is robust to including these additional controls. We estimate a value for $E\left[m_Z\left(P_Z\right)|X \in \Theta^{Reject}\right]$ of 0.0308 ($p = .018$) including these controls, as compared to 0.0338 ($p < 0.001$) for our primary specification.

6.7 Summary

We estimate significantly larger lower bounds for the average magnitude of private information for the rejectees versus non-rejectees. Our estimates are robust to a wide set of controls for public information, are robust to excluding those who own insurance in LTC and life, and are also consistent within age-based subsamples. Consistent with the theory in Section (2), our results suggest private information imposes a greater barrier to trade for the rejectees.

7 Estimation of Distribution of Private Information

While the lower bound approach provides evidence that private information imposes larger barriers to trade for the rejectees, it suffers several limitations. First, we made comparisons

[62] This is consistent with the much smaller sample size leading to a greater (spurious) predictive power of the subjective probabilities

using lower bounds, not the levels of $E[m(P)|X \in \Theta]$. Second, we made comparisons using the average magnitude of private information, $E[m(P)|X]$, not $m(p)\ \forall p$ or $\inf T(p)$ as suggested by Corollaries 3 and 4. Third, we could not quantify the minimum pooled price ratio.

To overcome these limitations, we introduce additional structure to the statistical relationship between elicitations and beliefs, as outlined in Section 4.2.

7.1 Empirical Specification

7.1.1 Elicitation Error Structure

Elicitations Z may differ from true beliefs P in many ways. They may be systematically biased, with values either higher or lower than true beliefs. They may be noisy, so that two individuals with the same beliefs may have different elicitations. Moreover, as shown in Figures 1(a,b,c) and recognized in previous literature (e.g. Gan et al. [2005]) people may have a tendency to report focal point values at 0, 50, and 100%.

Our model of elicitations will capture these three forms of elicitation error. To do so, we assume that the elicitation Z is drawn from a mixture of a censored normal and an ordered probit distribution. With probability $1-\lambda$, agents with belief P report Z from a censored normal distribution (censored on $[0,1]$) with mean $P+\alpha(X)$ and variance σ^2.[63] With probability λ, agents report $Z \in \{0, .5, 1\}$ according to an ordered probit with mean $P+\alpha(X)$, variance σ^2, and ordered probit cutoffs of κ and $1-\kappa$, where $\kappa \in [0, .5]$. The ordered probit allows a fraction λ of agents to report their beliefs not on a scale of 0-100%, but rather on a scale of "low, medium, and high", corresponding to elicitations of 0%, 50%, and 100%. Letting $f(Z|P,X)$ denote the p.d.f./p.m.f. of the distribution of elicitations, we have

$$f(Z|P,X) = \begin{cases} (1-\lambda)\Phi\left(\frac{-P-\alpha(X)}{\sigma}\right) + \lambda\Phi\left(\frac{\kappa-P-\alpha(X)}{\sigma}\right) & \text{if} \quad Z=0 \\ \lambda\left(\Phi\left(\frac{1-\kappa-P-\alpha(X)}{\sigma}\right) - \Phi\left(\frac{\kappa-P-\alpha(X)}{\sigma}\right)\right) & \text{if} \quad Z=0.5 \\ (1-\lambda)\Phi\left(\frac{1-P-\alpha(X)}{\sigma}\right) + \lambda\left(1-\Phi\left(\frac{1-\kappa-P-\alpha(X)}{\sigma}\right)\right) & \text{if} \quad Z=1 \\ \frac{1}{\sigma}\phi\left(\frac{Z-P-\alpha(X)}{\sigma}\right) & \text{if} \quad o.w. \end{cases}$$

where ϕ denotes the standard normal p.d.f. and Φ the standard normal c.d.f. We estimate four elicitation error parameters: $(\sigma, \lambda, \kappa, \alpha(X))$. σ captures the dispersion in the elicitation error, λ is the fraction of focal point respondents, κ is the focal point window. We allow the elicitation bias term, $\alpha(X)$, to vary with the observable variables, X.[64] Throughout this section we use our preferred pricing controls.

By modeling focal point responses as an independent ordered probit, we are assuming that

[63] In Appendix B.2.2, we provide Monte Carlo evidence that our estimation of the distribution of private information, $F(p)$, is reasonably robust to relaxing normality by introducing skewness and kurtosis.

[64] This allows elicitations to be biased, conditional on X; but we maintain the assumption that true beliefs are unbiased.

those who respond with focal point responses at 0, 50%, and 100% are drawn from the same distribution for P as those who report non-focal point values. Ideally, one would allow this distribution to differ; yet the focal point bias inherently limits the extent of information that can be extracted from their responses. In practice, this independence assumption means that most of our identification for the distribution of P will come from those reporting non-focal point values.

7.1.2 Flexible Approximation for the Distribution of Private Information

Although we impose a restrictive parametric structure on the distribution of elicitations given beliefs, we will flexibly estimate the distribution of private information.

Ideally, we would flexibly estimate $F(p|X)$ separately for every possible value of X. Unfortunately, the dimensionality of X prevents this in practice. Instead, we adopt an index assumption:

$$F(p|X) = \tilde{F}(p|\Pr\{L|X\}) \tag{10}$$

where we assume $\tilde{F}(p|q)$ is continuous in q. This assumes that the distribution of private information is the same for two segments, X and X', that have the same observable loss probability, $\Pr\{L|X\} = \Pr\{L|X'\}$. We will refer to q as an index. This assumption provides empirical tractability while still allowing the shape of F to vary with observables. Also, recall we conduct estimation separately for the rejectees and non-rejectees, so that we only impose the index assumption conditional on rejection classification.

We approximate $\tilde{F}(p|q)$ for $q = \Pr\{L|X\}$ using mixtures of beta distributions,

$$\tilde{F}(p|q) = \Sigma_i \eta_i Beta(\mu_i(q), \psi_i)$$

where η_i is the weight on each beta distribution, $\mu_i(q)$ is the mean of the i^{th} beta distribution at q, and ψ_i is the shape parameter of the i^{th} beta distribution.[65] We allow the shape parameter to vary for each beta distribution and we allow the mean of each beta distribution to vary as a linear function of q, $\mu_i(q) = \mu_i^0 + \mu_i^1 q$. Consistent beliefs (Assumption 1) imposes the restriction $\Sigma_i \eta_i \mu_i(q) = q$ which provide constraints on $\{\mu_i^0, \mu_i^1\}$, reducing the number of estimated parameters.

Beta distributions are quite flexible and well-suited for approximating arbitrary distributions. In practice, they fit our data quite well with a small mixture; we use two beta distributions for our preferred results in all settings except the no reject sample for LTC where we include an

[65]The p.d.f. of a beta distribution with parameters α and β is given by

$$beta(x;\alpha,\beta) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1} x^{\beta-1}$$

The mean of a beta distribution with parameters α and β is given by $\mu = \frac{\alpha}{\alpha+\beta}$ and the shape parameter is given by $\psi = \alpha + \beta$.

additional term to capture a point-mass at q.[66] All of our results are robust to including a 3rd beta distribution.

Bootstrap delivers consistent standard errors provided the true distribution, $\tilde{F}(p|q)$, is continuous. This assumption is violated in the event of no private information (in which case $\tilde{F}(p|q) = 1\{p \leq q\}$ for all p, q). However, the Wald tests for the presence of private information (constructed using our lower bound approach) provide a simple test of this event.[67]

7.1.3 Pooled Price Ratio (and its Minimum)

In principle, we can construct an estimate of the minimum pooled price ratio for any value of X; given our index assumption, this amounts to constructing the pooled price ratio for various values of the index q. We will often focus on results for the mean loss conditional on rejection classification, $q = \Pr\{L|X \in \Theta]\}$, as these estimates are based on the most in-sample information. But, we also present results for the 20th, 50th, and 80th percentiles of the distribution of q within each sample. This allows us to assess the minimum pooled price ratio varies with observables, X, within each sample.

As described in Section 4.2, we estimate the analogue to the minimum pooled price ratio, $\inf_{p \in \hat{\Psi}_\tau} T(p)$ for $\hat{\Psi}_\tau = \left[0, F^{-1}(\tau)\right]$. Our preferred choice for τ is 0.8, as this ensures at least 20% of the sample (conditional on q) is used to estimate $E[P|P \geq p]$ and produces estimates that are quite robust to changes in the number of approximating beta distributions. For robustness, we also present results for $\tau = 0.7$ and $\tau = 0.9$ along with plots of the pooled price ratio for all p below the estimated 90th quantile, $F^{-1}(0.9)$. We construct 5/95% confidence intervals for $\inf_{p \in \hat{\Psi}_\tau} T(p)$ by combining bootstrapped confidence intervals and extending the 5% boundary to 1 in the event that we cannot reject a null hypothesis of no private information.

[66] Although the beta distributions can theoretically approximate uninformed (point-mass) distributions quite well, convergence is slow in practice. We speed up our estimation for the no reject sample for LTC by mixing a truncated normal distribution which converts to a point mass distribution for a variance below 0.000025. This allows the estimation to more easily capture uninformed distributions. Including this point mass term in the other samples does not affect our results.

[67] Notice that $F(p|q) = 1\{p \leq q\}$ for all p and q if and only if $\Pr\{L|X,Z\} = Pr\{L|X\}$ for all X, so that our Wald test for the latter equality (from our lower bound approach) continues to provide a valid test for the presence of private information in our semiparametric approach. One could also construct a test for no private information for various sets of q values, however this suffers problems of limited power (and potential multiple testing issues), so we choose to focus on one aggregate test for the presence of private information in each rejection classification. In principle, imposing our restrictions on $f_{Z|P}(Z|P)$ could produce a more powerful test for the presence of private information. However, such a test faces technical hurdles since it involves testing whether $F(p|q)$ lies along a boundary of the set of possible distributions and must account for sample clustering (which makes a likelihood ratio test inappropriate). Andrews [2001] provides a potential method for constructing an appropriate test, but we leave this for future work.

7.2 Estimation Results

7.2.1 Graphs of the Distribution of Private Information

Qualitatively, no trade requires the existence of a "thick upper tail" of high risks who prevent the provision of insurance to lesser risks. We now assess this prediction. Figure 4(a,b,c) and Figure 5(a,b,c) present the estimated p.d.f.s and c.d.f.s of private information $\tilde{F}(p|q)$, plotted for the index equal to the mean loss in each sample, $q = \Pr\{L|X \in \Theta\}$.[68] In all three market settings, we find evidence of this pronounced upper tail for the rejectees. In contrast, we do not find such a significant upper tail for the non-rejectees.

For the LTC non-rejectees, we estimate a relatively condensed distribution around the mean probability of a loss of approximately 0.04 (as indicated in Figure 4a), consistent with our results from the lower bound test of no significant evidence of private information. However, for the rejectees, Figure 5a shows a more dispersed distribution and illustrates the presence of a relatively thick upper tail for the rejectees: 10% of the distribution lies above 0.3.

For Disability, we estimate a fairly dispersed distribution for the rejectees, along with a thick upper tail of risks: roughly 40% of the distribution is scattered between 0.4 and 1. In contrast, we find a smaller upper tail for the non-rejectees with 10% of the distribution above 0.3 and minimal mass above 0.7.

For Life, we estimate a fairly dispersed distribution for the rejectees, along with an upper tail of risks: we find significant mass ranging from 0.2 to 1. In contrast, for the non-rejectees we find a smaller upper tail of risks. Overall, these results are consistent with the qualitative prediction of the theory that rejectees have a thicker upper tail of risks than non-rejectees.

7.2.2 Minimum Pooled Price Ratio

We now turn to our quantitative estimates of the minimum pooled price ratio. Table 7 presents the estimates of the minimum pooled price ratio evaluated at several values of the index, q: the sample mean ($q = \Pr\{L|X \in \Theta\}$), the 20th, 50th, and 80th quantiles of the distribution of q within each sample. We let $\tau = 0.8$ and assess robustness to this choice in Table 8, discussed below.

LTC For the rejectees, the pooled price ratio reaches a minimum of 1.715 (5/95% CI of [1.575,1.779]) at the mean value of the index, $q = 0.175$, as reported in Table 7. This implies private information imposes an implicit tax of 71.5%. The estimates are similar for rejectees with other values of the index, q, ranging from 1.681 to 1.730. Together, the results are consistent with Corollary 2 provided the rejectees are unwilling to pay a 70% tax for insurance.

Our empirical approach does not attempt to estimate willingness-to-pay directly. To help understand whether it is reasonable to expect individuals to be unwilling to pay a 70% tax for

[68]The results are similar for other values of the index, such as the 20th, 50th, and 80th quantiles of q.

insurance, Table 9 presents calibrated values of $\frac{u'(w-l)}{u'(w)} - 1$ for values of the coefficient of relative risk aversion (henceforth CRRA) of 1, 2 or 3, and the size of the uninsured drop in consumption of 10%, 15%, and 20%.[69] For example, if CRRA is 2 and the nursing home entry is equivalent to a drop in consumption of 15%, then agents are would be willing to pay a 38.4% tax on insurance, much less than the needed 68-72% tax to sustain trade according to our estimates.[70]

For those served by the market, we cannot reject the null hypothesis of no private information, $\tilde{F}(p|q) = 1\{p \leq q\}\ \forall q$. We estimate a minimum pooled price ratio of 1.206 (5/95% CI of [1.00-1.484]) at the mean value of the index q ($q = 0.175$); estimates range from 1.337 to 1.147 as we vary the index between the 20th and 80th percentile of its distribution ($q = 0.017$ to $q = 0.057$). Our point estimates are consistent with the presence of trade as long as non-rejectees are willing to pay a 14-34% implicit tax.[71] Thus, a willingness-to-pay of 38.4% generated by our example above (with CRRA of 2 and a 15% consumption drop) would be sufficient to sustain trade.

Finally, the bottom rows of Table 7 report the estimated difference between the pooled price ratio for the rejectees relative to the non-rejectees, suggesting we can reject a null hypothesis of smaller minimum pooled price ratios for the rejectees relative to the non-rejectees.[72]

Disability For the rejectees, we estimate a minimum pooled price ratio of 1.954 (5/95% CI of [1.884,2.032]) at the mean index ($q = 0.441$), which implies a tax rate equivalence of 95.4%. The estimates are similar at the 20th and 50th quantile of q (1.900 and 1.937), and higher at the 80th quantile (2.282) of the index. Our estimates are fairly precise, with the 5/95% confidence intervals generally yielding a range of 0.1-0.2. The results are consistent with Corollary 2 provided the rejectees are unwilling to pay a 90-130% tax for insurance. We find these magnitudes to be quite large; indeed, if CRRA is 3 and disability leads to a 15% consumption drop, then the stylized calculation suggests individuals would be willing to pay a 62.8% tax for insurance, which would be insufficient to sustain trade amongst the rejectees.

For the non-rejectees, we estimate a minimum pooled price ratio of 1.611 (5/95% CI of [1.272,2.391]), implying that the barrier to trade faced by a person with an average observable loss probability among the rejectees is equivalent to a 61% tax on insurance premiums. This ranges from 1.703 to 1.572 as we vary the index q from its 20th to 80th quantile of its distribution

[69] These calculations are of course highly stylized since we do not estimate the CRRA nor do we take a stand on the consumption impact on the losses we study–indeed, the factors determining willingness-to-pay in these settings may be quite complicated. We only provide these numbers to aid in interpreting the magnitude of the results.

[70] Our results are also consistent with existing estimates of willingness to pay for LTC insurance provided in Brown and Finkelstein [2008] whose calibrated model suggests individuals are willing to pay roughly a 27-62% tax for existing LTC insurance policies (these numbers are not provided directly, but can be inferred from Figure 1 and Table 2. Figure 2 suggests the break-even point for insurance purchase is at the 60-70th percentile of the wealth distribution. Table 2 shows this corresponds to individuals being willing to pay a tax of 27-62%).

[71] Note that our inability to reject a tax rate of 0% at the 5% level suggests our results are consistent with the presence of trade for any loss size or risk aversion parameter.

[72] These comparisons are conditional on a given value of the percentile of q; although not reported, results are similar for other comparisons (e.g. 80th percentile of q for the rejectees compared to the 20th percentile of q for the non-rejectees).

($q = 0.109$ to $q = 0.197$). This suggests individuals must be willing to pay a 55-70% tax on insurance premiums in order to facilitate trade in this market. Thus, our results suggest an economically significant amount of private information in the private disability insurance market, even for those served by the market. The fact that the market exists despite this private information suggests a significant demand for insurance against becoming disabled.[73] Using the calibration in Table 9, this suggests that risk aversion must be above 3 if the loss size is 15% or above 2 if the loss size is 20%. However, we should note that the estimates are fairly imprecise for the non-rejectees–they include a 25% tax rate equivalence at the lower end of the 5/95% confidence interval.

Finally, the estimated differences between the rejectees and non-rejectees are all positive, yet statistically insignificantly different from zero, arguably a result of the imprecise estimation for the non-rejectees.

Life For the rejectees, we estimate a minimum pooled price ratio of 1.727 (5/95% CI of [1.527,2.193]) at the mean index, $q = 0.572$, indicating a tax rate equivalence of 72.7%. The estimates for other values of the index range from 1.642 at the 20th quantile of q ($q = 0.572$) to 2.269 at the 80th quantile of q ($q = 0.791$). The results are consistent with Corollary 2 as long as the rejectees are unwilling to pay a 65-130% tax for insurance.[74]

For those served by the market, we cannot reject the null hypothesis of no private information, as in LTC. We estimate a minimum pooled price ratio of 1.361 (5/95% CI of [1.00,1.421]), which implies a tax rate of 36.1% for a rejectee with an average observable loss probability. Our estimates for other values of the index, q, range from 1.640 at the 20th quantile ($q = 0.273$) to 1.345 at the 80th quantile ($q = 0.458$). Although some of these point estimates are large, we cannot reject the null hypothesis of a zero tax rate faced by the non-rejectees. The estimated differences between the rejectees and non-rejectees are generally significant at the 5% level, aside from the comparisons involving the point estimate of 1.640 for the 20th percentile of the index for the non-rejectees.

Choice of τ Table 8 presents results for $\tau = 0.7,\ 0.8,\ 0.9$ in each sample using the mean index value, q, in each sample.. Also, Figures 6a,b,c graph the estimated pooled price ratios, $T(p)$, for values of p less than the estimated 90th quantile of the distribution of private information for the mean value of the index, q, in each sample.

For LTC, the minimum of the pooled price ratio occurs at an interior point of the distribution,

[73] Although welfare and government policy is beyond the scope of this paper, our findings of significant amounts of private information in the private disability market are perhaps suggestive of a rationale for government disability insurance programs such as SSDI.

[74] The mapping to a willingness to pay in terms of CRRA preferences and a consumption drop is a bit more abstract and perhaps less useful in our life insurance setting; but if death is equivalent to a 15% consumption drop and CRRA is 3, then individuals would be willing to pay a 62.8% tax, insufficient to sustain trade and consistent with Corollary 2.

both for the rejectees and non-rejectees. Thus, our results are not sensitive to the choice of τ (in the range where $\tau \leq 0.9$) For disability, the minimum of the pooled price ratio occurs at an interior point for the non-rejectees, but is on the boundary for the rejectees, so the minimum pooled price ratio for the rejectees drops as we increase τ, as reported in Table 8 and shown in Figure 6b. Although the estimates for the rejectees fall to 1.727 for $\tau = 0.9$ and rise to 2.350 for τ =0.7, they remain quite large across these choices of τ. For life, the minimum of the pooled price ratio lies at the boundary for both the rejectees and non-rejectees. For the rejectees, the minimum pooled price ratio falls from 1.727 to 1.572 at $\tau = 0.9$ and rises to 1.865 at $\tau = 0.7$. For the non-rejectees, our estimates rise to 1.444 at $\tau = 0.7$ and fall to 1.281 at $\tau = 0.9$. As indicated by the bottom rows of Table 8, we can still reject the null hypothesis of a lower minimum pooled price ratio for rejectees relative to non-rejectees at each value of τ.

In short, the values of the minimum pooled price ratio and the comparisons between rejectees and non-rejectees are quite robust to the choice of τ.

7.2.3 Results for Elicitation Error Distribution

Table 10 presents our estimated results for the elicitation error distribution. In general, we find considerable support for the maintained hypothesis that subjective probabilities are only imperfect measures of agents beliefs. Estimates of the standard deviation of the elicitation error are primarily around 0.3-0.4, with the exception of an estimate of 0.1 for the non-rejectees in disability. Also, we estimate a sizable fraction of focal point respondents in each sample (35-50%).[75]

7.2.4 Annuities

Finally, we consider one additional test of our theory that private information leads to insurance rejections. There are no rejections in annuity markets. At first glance, it may seem odd that we find evidence of private information about mortality that, we argue, leads to rejections in life insurance. Yet annuities, which provide a fixed income stream regardless of one's length of life, insure the same (yet opposing) risk of living too long.

Our estimated distribution of private information about mortality reveals that, although some agents know that they have a relatively higher than average mortality risk, *few agents know that they have an exceptionally lower than average mortality risk.* As shown in Figure 4c, there are relatively few people, rejected or otherwise, with probabilities below the large mass

[75] We omit a discussion of bias and the focal point window. We do find significant evidence of bias, $\alpha(X)$, which varies with X. The mean bias by sample is already given by the difference of the first two rows of Table 3. Also, as shown in Table 10, we estimate focal windows around 0.2 in LTC (both rejectees and non-rejectees) and the non-rejectees in Life. This suggests focal responses of 0 correspond to non-focal response ranges of $[0, 0.2]$, responses of 50 correspond to non-focal responses of $[0.2, 0.8]$ and responses of 1 correspond to non-focal responses of $[0.8, 1]$. For rejectees in life and for both rejectees and non-rejectees in disability, we find estimates of the focal window close to zero. Estimates of a focal window of 0 have the simple interpretation that the focal point responders report 50% regardless of their private information.

around 0.15-0.2. Repeating our estimation of the pooled price ratio for $1 - P$ (probability of living 10-15 years), Table 11 reports a minimum of 1.177 for the life non-rejectees sample (for $\tau = 0.8$ and mean index q), which occurs around 0.2. As long as individuals are willing to pay a 18% tax for an annuity, our results are consistent with the presence of trade in annuity markets (indeed, we cannot reject the null hypothesis of a zero tax). Because there are few people with rejection conditions that have probabilities below 0.2, providing an annuity to the relatively healthy with probabilities around 0.2 does not require preventing the sick from being able to purchase it. By reversing the direction of the incentive constraints, rejections no longer occur.

8 Conclusion

This paper finds evidence consistent with the hypothesis that private information leads insurance companies to choose to not sell insurance to a subset of the population. We provide a new "no-trade" theorem which shows why insurance companies may choose to not offer insurance at any price acceptable to anyone in the market. We use the model to develop metrics to measure the barrier to trade imposed by private information. And, we develop a new empirical methodology to study private information which allows us to test whether a) those who would be rejected have larger barriers to trade imposed by private information and b) whether this barrier, measured as an implicit tax rate on insurance premiums, is sufficiently large to explain an absence of trade. We apply our approach to three markets: long-term care, disability, and life insurance, each of which have segments to whom insurance companies choose to not offer insurance. Across all of our settings, we find evidence of more private information for the rejectees, and we find its magnitude large enough to plausibly explain an absence of trade. In short, our results suggest that if insurance companies were to offer any contract or set of contracts to those currently rejected, they would be too adversely selected to yield a positive profit.

Our finding of no significant amounts private information for those who are served by the market in LTC and life is consistent with previous literature finding no evidence of adverse selection in these markets Finkelstein and McGarry [2006]; Cawley and Philipson [1999]). But our results suggest a new interpretation of the role of private information in insurance markets. The most salient impact of private information may not be the adverse selection of existing contracts, but rather the existence of the market itself.

References

G Akerlof. The market for lemons: Qualitative uncertainty and the market mechanism. *Quarterly journal of economics*, 84(3):488–500, 1970. 1, 2.2

D Andrews. Testing when a parameter is on the boundary of the maintained hypothesis. *Econometrica*, 69:683–734, 2001. 67

R. Forsythe M. Gorham R. Hahn R. Hanson J. Ledyard S. Levmore R. Litan P. Milgrom F. Nelson G. Newmann M. Ottaviani T. Schelling R. Shiller V. Smith E. Snowberg C. Sunstein P. C. Tetlock P. E. Tetlock H. Varian J. Wolfers E. Zitzewitz Arrow, K. The promise of prediction markets. *Science*, 320:877–878, 2008. 19

D Blackwell. Comparison of experiments. *Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 93–102, Jan 1951. 5, 20

D Blackwell. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, Jan 1953. 5, 20

J Brown and A Finkelstein. The interaction of public and private insurance: Medicaid and the long-term care insurance market. *The American Economic Review*, 98:1083–1102, 2008. 70

J Cawley and T Philipson. An empirical examination of information barriers to trade in insurance. *The American economic review*, 89(4):827–846, 1999. 1, 5.1, 6.4, 8

H Chade and E Schlee. Optimal insurance with adverse selection. *Theoretical Economics*, Forthcoming, May 2011. 1, 7, 2.2, 11

P Chiappori and B Salanié. Testing for asymmetric information in insurance markets. *Journal of Political Economy*, pages 56–78, 2000. 1

P Chiappori, B Jullien, B Salanié, and F Salanié. Asymmetric information in insurance: General testable implications. *RAND Journal of Economics*, pages 783–798, 2006. 1

A Cohen and P Siegelman. Testing for adverse selection in insurance markets. *The Journal of Risk and Insurance*, 77:39–84, 2010. 1

L Einav, A Finkelstein, and J Levin. Beyond testing: Empirical models of insurance markets. *Annual Review of Economics*, 2(1):311–336, Sep 2010a. 1

L Einav, A Finkelstein, and P Schrimpf. Optimal mandates and the welfare cost of asymmetric information: evidence from the uk annuity market. *Econometrica*, 78(3):1031–1092, 2010b. 19

A Finkelstein and K McGarry. Multiple dimensions of private information: Evidence from the long-term care insurance market. *American Economic Review*, 96(4):938–958, 2006. 1, 24, 5.1, 5.2.1, 41, 6.4, 8

A Finkelstein and J Poterba. Selection effects in the united kingdom individual annuities market. *The Economic Journal*, Jan 2002. 1

A Finkelstein and J Poterba. Adverse selection in insurance markets: Policyholder evidence from the uk annuity market. *Journal of Political Economy*, Jan 2004. 1

A Finkelstein and J Poterba. Testing for adverse selection with "unused observables". *NBER Working Paper No. 12112*, 2006. 24

L Gan, M Hurd, and D McFadden. Individual subjective survival curves. *Analyses in the Economics of Aging, ed. D Wise*, Aug 2005. 4, 19, 5.2.1, 7.1.1

G Gigerenzer and U Hoffrage. How to improve bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 102:684–704, 1995. 19

D He. The life insurance market: Asymmetric information revisited. *Journal of Public Economics*, Jan 2009. 5.1, 41, 6.4

Y Hu and S.M Schennach. Instrumental variable treatment of nonclassical measurement error models. *Econometrica*, 76(1):195–216, 2008. B.2

M Hurd. Subjective probabilities in household surveys. *Annual Review of Economics*, pages 543–562, 2009. 4

S. Miller, A. Kirlik, A. Kosorukoff, and J. Tsai. Supporting joint human-computer judgement under uncertainty. *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting*, pages 408–412, 2008. 19

H Miyazaki. The rat race and internal labor markets. *The Bell Journal of Economics*, 8(2): 394–418, 1977. 9

C Murtaugh, P Kemper, and B Spillman. Risky business: Long-term care insurance underwriting. *Inquiry*, 35(3):204–218, Jan 1995. 1, 5.1

W K Newey. Convergence rates and asymptotic normality for series estimators. *Journal of Econometrics*, 79(1):147–168, 1997. 6.2

W.K Newey and J.L Powell. Instrumental variable estimation of nonparametric models. *Econometrica*, 71(5):1565–1578, 2003. B.2

American Council of Life Insurers. *Life Insurers Fact Book 2010*. American Council of Life Insurers, Nov 2010. 5.1, 6.4

Congressional Budget Office. *Financing Long-Term Care for the Elderly*. Congressional Budget Office, Apr 2004. 5.1

JG Riley. Informational equilibrium. *Econometrica*, 47(2):331–359, 1979. 1, 7, 2.2, 11

M Rothschild and J Stiglitz. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *The Quarterly Journal of Economics*, pages 629–649, 1976. 1, 7, 2.1

L. J. Savage. *The Foundations of Statistics*. John Wiley & Sons Inc., New York, 1954. 18

M. Shaked and J. G. Shanthikumar. *Stochastic Orders and Their Applications*. Academic Press, Boston, 1994. 3.2

A.M Spence. Product differentiation and performance in insurance markets. *Journal of Public Economics*, 10(3):427–447, 1978. 9

C. Sunstein. *Infotopia: How Many Minds Produce Knowledge*. Oxford University Press, USA, 2006. 19

A Theory Appendix

A.1 Proof of No-Trade Theorem

We prove the no-trade theorem in several steps. First, we translate the problem to a maximization problem in utility space. Second, we prove the converse of the theorem directly by constructing an implementable allocation other than the endowment when Condition 1 does not hold. Third, we prove the no trade theorem for a finite type distribution. Fourth, we show finite type distributions can approximate solutions to arbitrary distributions, proving the no trade theorem for a general type distribution.

Most of these steps are straightforward. In our opinion, the key theoretical contribution comes in step 3 (Lemma (A.5)), where we show that Condition 1 implies that a separating allocation cannot improve over a full pooling allocation. Indeed, the ability for insurance companies to offer separating contracts is an important ingredient in previous models of this environment (Spence, 1979; Riley, 1979; Chade and Schlee, 2011).

A.1.1 Utility Space

First, we translate the problem to utility space. With this translation, the incentive and individual rationality constraints are linear in utility. Let $c(u) = u^{-1}(u)$ denote the inverse of the utility function $u(c)$, which is strictly increasing, continuously differentiable, and strictly convex. We denote the endowment allocation by $E = \{(c_L(p), c_{NL}(p))\}_p = \{(w-l, l)\}_p$. Let us denote the endowment allocation in utility space by $E^U = \{u(w-l), u(w)\}_p$. For allocations in utility space, we normalize $u_{NL}(1) = u(w)$.

Given a utility allocation $A^U = \{u_L(p), u_{NL}(p)\}_{p\in\Psi}$, let us denote the slack in the resource constraint by

$$\Pi\left(A^U\right) = \int \left[w - pl - pc\left(u_L(p)\right) - (1-p)\, c_{NL}(p)\right] dF(p)$$

We begin with a useful lemma that allows us to characterize when the endowment is the only implementable allocation.

Lemma A.1 (Characterization)**.** *The endowment is the only implementable allocation if and only if E^U is the unique solution to the following constrained maximization program, $P1$*

$$\begin{aligned}
P1 \quad &: \quad \max_{\{u_L(p), u_{NL}(p)\}_p} \int \left[w - pl - pc\left(u_L(p)\right) - (1-p)\, c\left(u_{NL}(p)\right)\right] dF(p) \\
s.t. \quad & \quad pu_L(p) + (1-p)\, u_{NL}(p) \geq pu_L(\hat{p}) + (1-p)\, u_L(\hat{p}) \quad \forall p, \hat{p} \in \Psi \\
& \quad pu_L(p) + (1-p)\, u_{NL}(p) \geq pu(w-l) + (1-p)\, u(w) \quad \forall p \in \Psi
\end{aligned}$$

Proof. Note that the constraint set is linear and the objective function is strictly concave. The first constraint is the incentive constraint in utility space. The second constraint is the individual rationality constraint in utility space. The linearity of the constraints combined with strict concavity of the objective function guarantees that the solutions are unique. Suppose that the endowment is the only implementable allocation and suppose, for contradiction, that the solution to the above program is not the endowment. Then, there exists an allocation $A^U = \{u_L(p), u_{NL}(p)\}$ such that $\int \left[w - pl - pc\left(u_L(p)\right) - (1-p)\, c\left(u_{NL}(p)\right)\right] dF(p) > 0$ which also satisfies the IC and IR constraints. Therefore, A^U is implementable, which yields a contradiction.

Conversely, suppose that there exists an implementable allocation B such that $B \neq E$. Let B^U denote the associated utility allocations to the consumption allocations in B. Then, B^U satisfies the incentive and individual rationality constraints. Since the constraints are linear, we know that the allocations $C^U(t) = tB^U + (1-t)E^U$ lie in the constraint set. By strict concavity of the objective function, $\Pi\left(C^U(t)\right) > 0$ for all $t \in (0,1)$. Since $\Pi\left(E^U\right) = 0$, E^U cannot be the solution to the constrained maximization program.

□

The lemma allows us to focus our attention on solutions to P1, a simple concave maximization program with linear constraints.

A.1.2 Converse

We begin the proof with the converse portion of the theorem: if the no-trade condition does not hold, then there exists an implementable allocation $A \neq E$ which does not utilize all resources and provides a strict utility improvement to a positive measure of types.

Lemma A.2 (Converse)**.** *Suppose Condition 1 does not hold so that there exists $\hat{p} \in \Psi \backslash \{1\}$ such that $\frac{\hat{p}}{1-\hat{p}} \frac{u'(w-l)}{u'(w)} > \frac{E[P|P \geq \hat{p}]}{1-E[P|P \geq \hat{p}]}$. Then, there exists an allocation $\hat{A}^U = \{(\hat{u}_L(p), \hat{u}_{NL}(p))\}_p$ and a positive measure of types, $\hat{\Psi} \subset \Psi$, such that*

$$p\hat{u}_L^A(p) + (1-p)\hat{u}_{NL}(p) > pu(w-l) + (1-p)u(w) \quad \forall p \in \hat{\Psi}$$

and

$$\int [W - pL - pc(\hat{u}_L(p)) - (1-p)c(\hat{u}_{NL}(p))] \, dF(p)$$

Proof. The proof follows by constructing an allocation which is preferable to all types $p \geq \hat{p}$ and showing that the violation of Condition 1 at $\hat{p}$ ensures its profitability. Given $\hat{p} \in \Psi$, either $P = \hat{p}$ occurs with positive probability, or any open set containing $\hat{p}$ has positive probability. In the case that $\hat{p}$ occurs with positive probability, let $\hat{\Psi} = \{\hat{p}\}$. In the latter case, note that the function $E[P|P \geq p]$ is locally continuous in p at $\hat{p}$ so that WLOG the no-trade condition does not hold for a positive mass of types. WLOG, we assume $\hat{p}$ has been chosen so that there exists a positive mass of types $\hat{\Psi}$ such that $p \in \hat{\Psi}$ implies $p \geq \hat{p}$. Then, for all $p \in \hat{\Psi}$, we have $\hat{\Psi} \subset \Psi$ such that

$$\frac{p}{1-p} \frac{u'(w-l)}{u'(w)} > \frac{E[P|P \geq p]}{1 - E[P|P \geq p]} \quad \forall p \in \hat{\Psi}$$

Now, for $\varepsilon, \eta > 0$, consider the augmented allocation to types $p \in \hat{\Psi}$:

$$\begin{aligned} u_L(\varepsilon, \eta) &= u(w-l) + \varepsilon + \eta \\ u_{NL}(\varepsilon, \eta) &= u(w) - \frac{1-\hat{p}}{\hat{p}}\varepsilon \end{aligned}$$

Note that if $\eta = 0$, ε traces out the indifference curve of individual $\hat{p}$. Construct the utility allocation $A^U(\varepsilon, \eta)$ defined by

$$(\hat{u}_L(p), \hat{u}_{NL}(p)) = \begin{cases} \left(u(w-l) + \varepsilon + \eta, u(w) - \frac{\hat{p}}{1-\hat{p}}\varepsilon\right) & if \quad p \geq \hat{p} \\ (u(w-l), u(w)) & if \quad p < \hat{p} \end{cases}$$

Note that for $\varepsilon > 0$ and $\eta > 0$ the utility allocation $(\hat{u}_L(p), \hat{u}_{NL}(p))$ is strictly preferred by all types $p \geq \hat{p}$ relative to the endowment utility allocation. Therefore, A_ε^U is individually rational and incentive compatible. We now only need to verify that there exists an allocation with $\varepsilon > 0$ and $\eta > 0$ which does not exhaust resources. We have

$$\Pi(\varepsilon, \eta) = \int [w - pl - pc(\hat{u}_L(p)) - (1-p)c(\hat{u}_{NL}(p))] \, dF(p)$$

Notice that this is continuously differentiable in ε and η. Differentiating with respect to ε and evaluating at $\varepsilon = 0$ yields

$$\frac{\partial \Pi}{\partial \varepsilon}|_{\varepsilon=0} = \int \left[-pc'(u(w-l+\eta)) + \frac{\hat{p}}{1-\hat{p}}(1-p)c'(u(w))\right] 1\{p \geq \hat{p}\} \, dF(p)$$

which is strictly positive if and only if

$$E[P|P \geq \hat{p}] \, c'(u(w-l+\eta)) < \frac{\hat{p}}{1-\hat{p}}(1 - E[P|P \geq \hat{p}]) \, c'(u(w))$$

Notice that this is continuous in η. So, at $\eta = 0$, we have

$$\frac{\partial \Pi}{\partial \varepsilon}|_{\varepsilon=0,\eta=0} > 0 \quad \Longleftrightarrow \quad \frac{\hat{p}}{1-\hat{p}} \frac{u'(w-l)}{u'(w)} > \frac{E[P|P \geq \hat{p}]}{1 - E[P|P \geq \hat{p}]}$$

and thus by continuity, the above condition holds for sufficiently small $\eta > 0$, proving the existence of an allocation which both delivers strictly positive utility for a positive fraction of types and does not exhaust all resources.

This shows that Condition 1 is necessary for the endowment to be the only implementable allocation.

□

A.1.3 Lemmas

Here, we prove two useful lemmas. First, we show that if Condition 1 holds, then the MRS is bounded by the pooled price ratio in the relevant quadrant of allocations.

Lemma A.3. *Suppose Condition 1 holds. Then for all $c_L, c_{NL} \in [w-l, l]$, we have*

$$\frac{p}{1-p} \frac{u'(c_L)}{u'(c_{NL})} \leq \frac{E[P|P \geq p]}{1 - E[P|P \geq p]} \quad \forall p \in \Psi \backslash \{1\}$$

and if $c_L, c_{NL} \in (w-l, l)$, we have

$$\frac{p}{1-p} \frac{u'(c_L)}{u'(c_{NL})} < \frac{E[P|P \geq p]}{1 - E[P|P \geq p]} \quad \forall p \in \Psi \backslash \{0, 1\}$$

Proof. Since $u'(c)$ is decreasing in c, we have $\frac{u'(c_L)}{u'(c_{NL})} \leq \frac{u'(w-l)}{u'(w)}$. Therefore, the result follows immediately from Condition 1. The strict inequality follows from strict concavity of $u(c)$.

□

Lemma A.4. *In any solution to $P1$, we have $c_L(p) \geq w - l$ and $c_{NL}(p) \leq w$.*

Proof. Suppose $A = \{c_L(p), c_{NL}(p)\}_p$ is a solution to $P1$. First, suppose that $c_L(\hat{p}) < w - l$. For this contract to be individually rational, we must have $c_{NL}(\hat{p}) > w$. Incentive compatibility requires $c_L(p) \leq c_L(p) < w - l \ \forall p < \hat{p}$ and $c_{NL}(p) \geq c_{NL}(\hat{p}) > w \ \forall p < \hat{p}$. Consider the new allocation $\tilde{A} = \{\tilde{c}_L(p), \tilde{c}_{NL}(p)\}$ defined by

$$\tilde{c}_L(p) = \begin{cases} c_L(p) & if\ p > \hat{p} \\ w - l & if\ p \leq \hat{p} \end{cases}$$

$$\tilde{c}_{NL}(p) = \begin{cases} c_{NL}(p) & if\ p > \hat{p} \\ w & if\ p \leq \hat{p} \end{cases}$$

Then $\tilde{A}$ is implementable (IC holds because of single crossing of the utility function). It only remains to show that $\Pi(A) < \Pi\left(\tilde{A}\right)$. But this follows trivially. Notice that the IR constraint and concavity of the utility function requires that points $(c_L(p), c_{NL}(p))$ lie above the zero profit line $p(w - l - c_L) + (1-p)(w - c_{NL})$. Thus, each point $(c_L(p), c_{NL}(p))$ must earn negative profits at each $p \leq \hat{p}$.

Now, suppose $c_{NL}(\hat{p}) > w$. Then, the incentive compatibility constraint requires $c_{NL}(p) > w \ \forall p \leq \hat{p}$. Construct $\tilde{A}$ as above, yielding the same contradiction.

□

We now prove the theorem in two steps. First, we prove the result for a finite type distribution. We then pass to the limit to cover the case of arbitrary distributions.

A.1.4 Finite Types

To begin, suppose that $\Psi = \{p_1, ...p_N\}$. We first show that condition 1 implies that the solution to $P1$ is a pooling allocation which provides the same allocation to all types.

Lemma A.5. *Suppose $\Psi = \{p_1, ..., p_N\}$ and that condition 1 holds (note that this requires $p_N = 1$). Then, the solution to $P1$ is a full pooling allocation: there exists $\bar{u}_L, \bar{u}_{NL}$ such that $(u_L(p), u_{NL}(p)) = (\bar{u}_L, \bar{u}_{NL})$ for all $p \in \Psi \backslash \{0,1\}$, $u_L(1) = \bar{u}_L$, $u_{NL}(0) = \bar{u}_{NL}$.*

Proof. Let $A^U = \{u_L^*(p), u_{NL}^*(p)\}_p$ denote the solution to P and suppose for contradiction that the solution to P is not a full pooling allocation. Let $\hat{p} = \min\{p | u_L^*(p) = u_L^*(1)\}$, let $\hat{p}_- = \max\{p | u_L^*(p) \neq u_L^*(1)\}$. The assumption that Ψ is finite implies that $\hat{p} > \hat{p}_-$. Let us define the pooling sets $J = \{p | u_L^*(p) = u_L^*(1)\}$ and $K = \{p | u_L^*(p) = u_L^*(\hat{p}_-)\}$. We will show that a profitable deviation exists which pools groups J and K into the same allocation. First, notice that if $\hat{p} = 1$, then clearly it is optimal to provide group J with the same amount of consumption in the event of a loss as group K, since otherwise the IC constraint of the type $\hat{p} = 1$ type would be slack. So, we need only consider the case $\hat{p} < 1$.

Notice that if the IR constraint of any member of group J binds (i.e. if the IR constraint for $\hat{p}$ binds), then their IC constraint implies that the only possible allocation for the lower risk types $p < \hat{p}$ is the endowment. This standard result follows from single crossing of the utility function. Therefore, we have two cases. Either all types $\tilde{p} \in \Psi \backslash J$ receive their endowment, $(c_L, c_{NL}) = (w - l, w)$, or the IR constraint cannot bind for any member of J. We consider these two cases in turn.

Suppose $u_L^*(p) = u(w - l)$ and $u_{NL}^*(p) = u(w)$ for all types $\tilde{p} \in \Psi \backslash J$. Clearly, we must then have that the IR constraint must bind for type $\hat{p}$, since otherwise profitability could be improved by lowering the utility provided to types $\tilde{p} \in \Psi \backslash J$. We now show that the profitability of the allocation violates the no-trade condition. The profitability of A^U is

$$\Pi\left(A^U\right) = \int_{p \in J} \left[w - pl - pc\left(u_L^*\left(\hat{p}\right)\right) - (1-p)\, c\left(u_{NL}^*\left(\hat{p}\right)\right)\right] dF\left(p\right)$$

Now, we construct the utility allocation A_t^U by

$$\left(u_L^t\left(p\right), u_{NL}^t\left(p\right)\right) = \begin{cases} \left(u\left(w-l\right)+t, u\left(w\right) - \frac{\hat{p}}{1-\hat{p}} t\right) & if \quad p \in J \\ \left(u\left(w-l\right), u\left(w\right)\right) & if \quad p \notin J \end{cases}$$

Since the IR constraint binds for type $\hat{p}$, we know that there exists $\hat{t}$ such that $A_{\hat{t}}^U = A^U$. By Lemma A.4, $\hat{t} > 0$ and A_t^U satisfies IC and IR for any $t \in \left[0, \hat{t}+\eta\right]$ for some $\eta > 0$. Since profits are maximized at $t = \hat{t}$ and since the objective function is strictly concave, it must be the case that

$$\frac{d\Pi\left(A_t^U\right)}{dt}|_{t=\hat{t}} = 0$$

where

$$\frac{d\Pi\left(A_t^U\right)}{dt}|_{t=\hat{t}} = \int_{p \in J} \left[pc'\left(u_L^*\left(p\right)\right) - (1-p)\, c'\left(u_{NL}^*\left(p\right)\right) \frac{\hat{p}}{1-\hat{p}}\right] dF\left(p\right)$$

Re-arranging and combining these two equations, we have

$$\frac{\hat{p}}{1-\hat{p}} \frac{u'\left(c\left(u_L^*\left(\hat{p}\right)\right)\right)}{u'\left(c\left(u_{NL}^*\left(\hat{p}\right)\right)\right)} = \frac{E\left[P|P \geq \hat{p}\right]}{1 - E\left[P|P \geq \hat{p}\right]}$$

which, by strict concavity of u, implies

$$\frac{\hat{p}}{1-\hat{p}} \frac{u'\left(w-l\right)}{u'\left(w\right)} > \frac{E\left[P|P \geq \hat{p}\right]}{1 - E\left[P|P \geq \hat{p}\right]}$$

which contradicts Condition 1.

Now, suppose that the IR constraint does not bind for any member of J. Then, clearly the IC constraint for type $\hat{p}$ must bind, otherwise profit could be increased by lowering the utility provided to members of J. So, construct the utility allocation B_ε^U to be

$$\left(u_L^\varepsilon\left(p\right), u_{NL}^\varepsilon\left(p\right)\right) = \begin{cases} \left(u_L^*\left(\hat{p}\right) - \varepsilon, u_{NL}^*\left(\hat{p}\right) + \frac{\hat{p}}{1-\hat{p}}\varepsilon\right) & if \quad p \geq \hat{p} \\ \left(u_L^*\left(p\right), u_{NL}^*\left(p\right)\right) & if \quad p < \hat{p} \end{cases}$$

so that B_ε^U consists of allocations equivalent to A^U except for $p \in J$. By construction, B_ε^U, is IR for any ε. Moreover, because of single crossing and because types are separated (finite types), B_ε^U continues to be IC and IR for $\varepsilon \in (-\eta, \eta)$ for some $\eta > 0$ sufficiently small. Therefore, we must have $\frac{d\Pi\left(B_\varepsilon^U\right)}{d\varepsilon}|_{\varepsilon=0} = 0$, which implies

$$\begin{aligned}\frac{d\Pi\left(B_\varepsilon^U\right)}{d\varepsilon}|_{\varepsilon=0} &= \int_{p\in J}\left[pc'\left(u_L^*\left(\hat{p}\right)\right)-\left(1-p\right)c'\left(u_{NL}^*\left(\hat{p}\right)\right)\frac{\hat{p}}{1-\hat{p}}\right]dF\left(p\right)\\ &= \Pr\left\{p\in J\right\}\left[E\left[P|P\geq\hat{p}\right]\frac{1}{u'\left(c\left(u_L^*\left(\hat{p}\right)\right)\right)}-\left(1-E\left[P|P\geq\hat{p}\right]\right)\frac{1}{u'\left(c\left(u_{NL}^*\left(\hat{p}\right)\right)\right)}\frac{\hat{p}}{1-\hat{p}}\right]\\ &= \Pr\left\{p\in J\right\}\frac{\left(1-E\left[P|P\geq\hat{p}\right]\right)}{u'\left(c\left(u_L^*\left(\hat{p}\right)\right)\right)}\left[\frac{E\left[P|P\geq\hat{p}\right]}{\left(1-E\left[P|P\geq\hat{p}\right]\right)}-\frac{u'\left(c\left(u_L^*\left(\hat{p}\right)\right)\right)}{u'\left(c\left(u_{NL}^*\left(\hat{p}\right)\right)\right)}\frac{\hat{p}}{1-\hat{p}}\right]\\ &= 0\end{aligned}$$

which implies

$$\frac{\hat{p}}{1-\hat{p}}\frac{u'\left(c\left(u_L^*\left(\hat{p}\right)\right)\right)}{u'\left(c\left(u_{NL}^*\left(\hat{p}\right)\right)\right)}=\frac{E\left[P|P\geq\hat{p}\right]}{1-E\left[P|P\geq\hat{p}\right]}$$

which, by strict concavity of u, implies

$$\frac{\hat{p}}{1-\hat{p}}\frac{u'\left(w-l\right)}{u'\left(w\right)}>\frac{E\left[P|P\geq\hat{p}\right]}{1-E\left[P|P\geq\hat{p}\right]}$$

which contradicts Condition 1. Therefore, if Condition 1 holds, the only possible solution to $P1$ is a full pooling allocation.

□

This lemma proves the vast majority of the proof for the finite support case. All that remains to show is that a full pooling allocation cannot be a solution to $P1$.

Lemma A.6. *Suppose Condition 1 holds. Then, the only possible full-pooling solution to $P1$ is E^U.*

Proof. Suppose for contradiction that $A^U \neq E^U$ is a full-pooling solution to $P1$. Let u_L^*, u_{NL}^* denote the full pooling allocations A^U. Recall $p_1 = \min\Psi$ is the lowest risk type. Note that the IR constraint for the $p_1 = \min\Psi$ type must bind in any solution to $P1$. Otherwise, profits could be increased by providing all types with less consumption, without any consequences on the incentive constraints of types $p > p_1$. Consider the allocations C_t^U defined by

$$\left(u_L^t, u_{NL}^t\right)=\left(u_L^*+\left(1-t\right)\left(u\left(w-l\right)-u_L^*\right), u_{NL}^*+\left(1-t\right)\left(u\left(w\right)-u_{NL}^*\right)\right)$$

so that when $t=1$ these allocations correspond to A^U and $t=0$ corresponds to the endowment. Because the IR constraint of the p_1 type must hold, we know that these allocations must follow the iso-utility curve of the p_1 type which runs through the endowment. Differentiating with respect to t and evaluating at $t=0$ yields

$$\frac{d\Pi\left(C_t^U\right)}{dt}|_{t=0}=E\left[P|P\geq p_1\right]c'\left(u\left(w-l\right)\right)-\left(1-E\left[P|P\geq p_1\right]\right)c'\left(u\left(w\right)\right)\frac{p_1}{1-p_1}$$

where $\frac{p_1}{1-p_1}$ comes from the fact that we can parameterize the iso-utility curve of the p_1 type by $u_L -$

$\tau, u_{NL} + \frac{p_1}{1-p_1}\tau$. But re-arranging the equation, we have

$$\begin{aligned} \frac{d\Pi\left(C_t^U\right)}{dt}|_{t=0} &= -E\left[P|P \geq p_1\right]\frac{1}{u'\left(w-l\right)} + \left(1 - E\left[P|P \geq p_1\right]\right)\frac{1}{u'\left(w\right)}\frac{p_1}{1-p_1} \\ &= \frac{1 - E\left[P|P \geq p_1\right]}{u'\left(W-L\right)}\left(-\frac{E\left[P|P \geq p_1\right]}{1 - E\left[P|P \geq p_1\right]} + \frac{u'\left(w-l\right)}{u'\left(w\right)}\frac{p_1}{1-p_1}\right) < 0 \end{aligned}$$

which yields a contradiction of Condition 1 at $p = p_1$.

$\square$

Therefore, we have shown that if Ψ is finite, then if Condition 1 holds, the only possible allocation is the endowment. It only remains to show that this property holds when Ψ is not finite.

A.1.5 Arbitrary Distribution

If $F(p)$ is continuous or mixed and satisfies the no-trade condition, we first show that F can be approximated uniformly by a sequence F_n of finite support distributions on $[0,1]$, each of which satisfy the no-trade condition.

Lemma A.7. *Let P be any random variable on $[0,1]$ with c.d.f. $F(p)$. Then, there exists a sequence of random variables, P_N, with c.d.f. $F^N(p)$, such that $F^N \to F$ uniformly and*

$$E\left[P_N|P_N \geq p\right] \geq E\left[P|P \geq p\right] \quad \forall p, \forall N$$

Proof. Since F is increasing, it has at most a countable number of discontinuities on $[0,1]$. Let $D = \{\delta_i\}$ denote the set of discontinuities and WLOG order these points so that $\lim_{\varepsilon \to +0} F(\delta_i) - \lim_{\varepsilon \to -0} F(\delta_i)$ is decreasing in i (so that δ_1 is the point of largest discontinuity). Then, the distribution F is continuous on $\Psi \backslash D$. For any N, let ω_N denote a partition of $[0,1]$ given by $2^N + \min\{N, |D|\} + 1$ points equal to $\frac{j}{2^N}$ for $j = 0, ..., 2^N$ and $\{\delta_i | i \leq N\}$. We write $\omega_N = \left\{p_j^N\right\}_{j=1}^{2^N + \min\{N,|D|\}+1}$. Now, define $\hat{F}^N : \omega_N \to [0,1]$ by

$$\hat{F}^N(p) = F\left(\max\left\{p_j^N | p_j^N \leq p\right\}\right)$$

so that $\hat{F}^N$ converges to F uniformly as $N \to \infty$.

Unfortunately, we cannot be assured that $\hat{F}^N$ satisfies the no-trade condition. But, we can perform the following simple modification to $\hat{F}^N$ to arrive at a distribution that does satisfy the no-trade condition for all N. We first describe the modification in the abstract and then apply it to our $\hat{F}^N$ distribution. For any $\lambda \in [0,1]$ and for any random variable X distributed $G(x)$ on $[0,1]$ define the random variable X_λ to be the random variable with c.d.f. $\lambda G(x)$. In other words, with probability λ the variable is distributed according to X and with probability $1-\lambda$ the variable takes on a value of 1 with certainty. Notice that $E\left[X_\lambda | X_\lambda \geq x\right]$ is continuously decreasing in λ and $E\left[X_0|X_0 \geq x\right] = 1 \; \forall x$.

Now, given $\hat{F}^N$ with associated random variable $\hat{P}^N$, we define P_λ^N to be the random variable with c.d.f. $\lambda \hat{F}^N(p)$. We now define a sequence $\{\lambda_N\}_N$ by

$$\lambda_N = \max\left\{\lambda | E\left[P_\lambda^N | P_\lambda^N \geq p\right] \geq E\left[P|P \geq p\right] \quad \forall p\right\}$$

Note that for each N fixed, the set $\left\{\lambda | E\left[P_\lambda^N | P_\lambda^N \geq p\right] \geq E\left[P|P \geq p\right] \quad \forall p\right\}$ is a compact subset of $[0,1]$, so that the maximum exists. Given λ_N, we define our new approximating distribution, $F^N(p)$, by

$$F^N(p) = \lambda_N F^N(p)$$

which satisfies the no-trade condition for all N. The only thing that remains to show is that $\lambda_N \to 1$ as $N \to \infty$.

By definition of λ_N, for each N there exists $\tilde{p}_N$ such that

$$E\left[P^N_{\lambda_N}|P^N_{\lambda_N} \geq \tilde{p}_N\right] = E\left[P|P \geq \tilde{p}_N\right]$$

Moreover, because λ_N is bounded, it has a convergent subsequence, $\lambda_{N_k} \to \lambda^*$. Therefore,

$$E\left[P^{N_k}_{\lambda^*}|P^{N_k}_{\lambda^*} \geq q\right] \to E\left[P_{\lambda^*}|P_{\lambda^*} \geq q\right]$$

uniformly (over q) as $k \to 0$, where P_{λ^*} is the random variable with c.d.f. $\lambda^* F(p)$. Moreover,

$$E\left[P^{N_k}_{\lambda_{N_k}}|P^{N_k}_{\lambda_{N_k}} \geq q\right] \to E\left[P_{\lambda^*}|P_{\lambda^*} \geq q\right]$$

uniformly (over q) as $k \to 0$. Therefore,

$$E\left[P^{N_k}_{\lambda^*}|P^{N_k}_{\lambda^*} \geq \tilde{p}_N\right] \to E\left[P|P \geq \tilde{p}_N\right]$$

so that we must have $\lambda^* = 1$.

Therefore, the distribution P^{N_k} with c.d.f. $F^{N_k}(p) = \lambda_{N_k} F^{N_k}(p)$ for $k \geq 1$ has the property

$$E\left[P^{N_k}|P^{N_k}_{\lambda} \geq p\right] \geq E\left[P|P \geq p\right] \quad \forall p$$

and $F^{N_k}(p)$ converges uniformly to $F(p)$.

□

Now, returning to problem $P1$ for an arbitrary distribution $F(p)$ which satisfies the no-trade condition. Let $\Pi(A|F)$ denote the value of the objective function for allocation A under distribution F. Suppose for contradiction that an allocation $\hat{A} = (\hat{u}_L(p), \hat{u}_{NL}(p)) \neq (W - L, W)$ is the solution to $P1$ under distribution F, so that $\Pi(A|F) > 0$. Let $F^N(p)$ be a sequence of finite approximating distributions which satisfy the no-trade condition and converge uniformly to F. Let $\omega_N = \left\{p^N_j\right\}$ denote the support of each approximating distribution. For any N, define the augmented allocation $\hat{A}_N = \left(\hat{u}^N_L(p), \hat{u}^N_{NL}(p)\right)$ by choosing $(\hat{u}_L(p), \hat{u}_{NL}(p))$ to be the most preferred bundle from the set $\left\{u_L\left(p^N_j\right), u_{NL}\left(p^N_j\right)\right\}_j$. Since $\hat{A}$ is incentive compatible, clearly we will have $\left(\hat{u}^N_L\left(p^N_j\right), \hat{u}^N_{NL}\left(p^N_j\right)\right) = \left(\hat{u}_L\left(p^N_j\right), \hat{u}_{NL}\left(p^N_j\right)\right)$. By single crossing, for $p \neq p^N_j$ agents with $p \in \left(p^N_{j-1}, p^N_j\right)$ will prefer either allocation for type p^N_{j-1} or p^N_j.

Clearly, $\hat{A}_N$ converges uniformly to $\hat{A}$. Since $\hat{A}_N$ satisfies IC and IR by construction, the no-trade condition implies that the allocation $\hat{A}_N$ cannot be as profitable as the endowment, so that we have

$$\Pi\left(\hat{A}_N|F_N\right) \leq \Pi(E|F_N) = 0 \quad \forall N$$

By the Lebesgue dominated convergence theorem ($\Pi\left(\hat{A}_N|F_N\right)$ is also bounded below by $-(W + L)$), have

$$\Pi\left(\hat{A}|F\right) \leq 0$$

Which yields a contradiction that $\hat{A}$ was the optimal solution (which required $\Pi\left(\hat{A}|F\right) > 0$) and concludes the proof.

B Empirical Methodology Appendix

B.1 Properties of the Lower Bound Estimator

This section further examines properties of the nonparametric lower bound approach. To derive these properties of $E\left[m_Z\left(P_Z\right)\right]$, we first show P is a mean-preserving spread of P_Z. We have

$$\begin{aligned} E\left[P|X,Z\right] &= E\left[\Pr\left\{L|X,P\right\}|X,Z\right] \\ &= E\left[\Pr\left\{L|X,Z,P\right\}|X,Z\right] \\ &= \Pr\left\{L|X,Z\right\} \\ &= P_Z \end{aligned}$$

where the first equality follows from assumption 1, the second equality follows from assumption 2, the third equality follows from the law of iterated expectations (averaging over realizations of P given X and Z), and the fourth equality is simply the definition of P_Z.

We now define the quantiles of P and P_Z which will help describe how $E\left[m_Z\left(P_Z\right)\right]$ relates to $E\left[m\left(P\right)\right]$. Let $Q_P\left(\alpha\right)$ to be the α-quantile of P,

$$Q_P\left(\alpha\right) = \inf_q \left\{q|\Pr\left\{P \leq q\right\} \geq \alpha\right\}$$

and $Q_\alpha\left(P_Z\right)$ to be the α-quantile of our analogue,

$$Q_{P_Z}\left(a\right) = \inf_q \left\{q|\Pr\left\{P_Z \leq q\right\} \geq \alpha\right\}$$

Given these two quantiles, let $e\left(\alpha\right)$ denote the difference between them,

$$e\left(\alpha\right) = Q_P\left(\alpha\right) - Q_{P_Z}\left(\alpha\right) \tag{11}$$

This function parameterizes the effect of the "noise" in Z. If $e\left(\alpha\right) > 0$ (< 0), then the α-quantile of P_Z falls below (above) the true α-quantile of the distribution of private information, P. On average, the effect of the noise is zero, $\int e\left(\alpha\right) d\alpha = 0$, since P is a mean-preserving spread of P_Z.

We now have defined the required variables to characterize the properties of $E\left[m_Z\left(P_Z\right)\right]$. Where applicable, we let the integers 1 and 2 denote two market segments (e.g. $X = x_1$ and $X = x_2$). Subscripted 1 and 2 will denote each segment (e.g. P_1 and P_2 denote the distributions of private information in segments 1 and 2).

Proposition. *The following conditions hold*

1. (Characterization of $E\left[m\left(P\right)\right]$ and $E\left[m_Z\left(P_Z\right)\right]$) $E\left[m\left(P\right)\right]$ and $E\left[m_Z\left(P_Z\right)\right]$ can be written as

$$E\left[m\left(P\right)\right] = \int_0^1 \left(Q_P\left(\alpha\right) - \Pr\left\{L\right\}\right)\log\left(\frac{1}{1-\alpha}\right) d\alpha$$

and

$$E\left[m_Z\left(P_Z\right)\right] = \int_0^1 \left(Q_{P_Z}\left(P_Z\right) - \Pr\left\{L\right\}\right)\log\left(\frac{1}{1-\alpha}\right) d\alpha$$

2. (No private information) If P is a constant, then $E\left[m\left(P\right)\right]=E\left[m_Z\left(P_Z\right)\right]=0$

3. (Lower bound - Re-statement of Proposition 2) $E\left[m_Z\left(P_Z\right)\right]\leq E\left[m_Z\left(P_Z\right)\right]$ so that $E\left[m_Z\left(P_Z\right)\right]$ is a lower bound for $E\left[m\left(P\right)\right]$

4. (Comparisons across segments) $E\left[m_1\left(P_1\right)\right]-E\left[m_2\left(P_2\right)\right]=E\left[m_{Z,1}\left(P_{Z,1}\right)\right]-E\left[m_{Z,2}\left(P_{Z,2}\right)\right]+\int\left[e_1\left(\alpha\right)-e_2\left(\alpha\right)\right]\log\left(\frac{1}{1-\alpha}\right)d\alpha$

5. (Relation to $\inf_p T\left(p\right)$) $\inf_p T\left(p\right)\leq 1+\frac{E[m(P)]}{E[P(1-P)]-E[m(P)]\Pr\{L\}-E[(P-\Pr\{L\})m(P)]}$, with equality if $T\left(p\right)$ is equal to a constant for all p

The first condition shows that $E\left[m\left(P\right)\right]$ and $E\left[m_Z\left(P_Z\right)\right]$ are weighted averages of the quantiles, $Q_P\left(\alpha\right)-\Pr\left\{L\right\}$ and $Q_\alpha\left(P_Z\right)-\Pr\left\{L\right\}$. The term $\log\left(\frac{1}{1-\alpha}\right)$ weights upper quantiles (near $\alpha=1$) more heavily than lower quantiles and implies that $E\left[m\left(P\right)\right]$ and $E\left[m_Z\left(P_Z\right)\right]$ are positive. This weighting has an intuitive meaning: high risks (high values of P) are included in the calculation for the magnitude of private information for more of the population. Therefore, the probabilities for the high risks are weighted more heavily in $E\left[m\left(P\right)\right]$. In this sense, $E\left[m\left(P\right)\right]$ is a measure of the thickness of the upper tail of P.

The second condition shows that testing $E\left[m_Z\left(P_Z\right)\right]=0$ provides a test for the existence of private information in a given segment. The third condition states that $E\left[m_Z\left(P_Z\right)\right]$ is a lower bound for $E\left[m\left(P\right)\right]$, which is a re-statement of Proposition 2 (but for which we will now provide the proof). The fourth condition shows that one can infer comparisons of $E\left[m\left(P\right)\right]$ across market segments using $E\left[m_Z\left(P_Z\right)\right]$ provided the error, $\int\left[e_1\left(\alpha\right)-e_2\left(\alpha\right)\right]\log\left(\frac{1}{1-\alpha}\right)d\alpha$ is small. In Section B.1.3 we use this result to provide an example which illustrates when inference using $E\left[m_Z\left(P_Z\right)\right]$ is valid for inference about $E\left[m\left(P\right)\right]$.

The fifth condition relates $E\left[m\left(P\right)\right]$ to the quantity that characterizes the barrier to trade, $\inf_p T\left(p\right)$. Using a Holder inequality, this condition shows that $E\left[m\left(P\right)\right]$ is monotonically related to an upper bound on $\inf_p T\left(p\right)$. Notice that the RHS of the expression in the fifth condition is increasing in both $E\left[m\left(P\right)\right]$ and $\Pr\left\{L\right\}$, provided $E\left[P\left(1-P\right)\right]$, and $E\left[\left(P-\Pr\left\{L\right\}\right)m\left(P\right)\right]$ remain roughly constant. Thus, smaller values of $E\left[m\left(P\right)\right]$ lead to smaller upper bounds on the minimum pooled price ratio. But also, smaller values of the mean risk, $\Pr\left\{L\right\}$, lead to smaller upper bounds on the minimum pooled price ratio. Thus, since rejectees have larger values of $\Pr\left\{L\right\}$, it could very well be the case that this lower bound is smaller for non-rejectees even if $E\left[m\left(P\right)\right]$ is larger for non-rejectees. In this sense, our lower bound test is a potentially overly restrictive test of the implications of the theory. Since we nonetheless find larger values of $E\left[m\left(P\right)\right]$ for rejectees, we do not discuss this potential bias in detail in the text; it only renders our empirical findings to be even greater support for the theory that private information leads to rejections.

B.1.1 Proof of Proposition

For part (1), let $Q_P\left(\alpha\right)$ denote the α-quantile of P. $\hat{P}$ denote an independent copy of P. We can write $E\left[m\left(P\right)\right]$ by integrating across the quantiles of P,

$$E\left[m\left(P\right)\right]=E_\alpha\left[E_P\left[P|P\geq Q_P\left(\alpha\right)\right]\right]$$

so that we have the expansion

$$\begin{aligned}
E\left[m\left(P\right)\right] &= \int_0^1 \left[E_{\tilde{\alpha}}\left[Q_P\left(\tilde{\alpha}\right) - Q_P\left(\alpha\right)|\tilde{\alpha} \geq \alpha\right]\right] d\alpha \\
&= \int_0^1 \frac{1}{1-\alpha}\left[\int_{\tilde{a}\geq\alpha}\left[Q_P\left(\tilde{\alpha}\right) - Q_P\left(\alpha\right) d\tilde{\alpha}\right]\right] d\alpha \\
&= \int_0^1 \int_{\tilde{a}\geq\alpha} \frac{Q_P\left(\alpha\right)}{1-\alpha} d\tilde{\alpha} d\alpha - E\left[P\right] \\
&= \int_0^1 Q_P\left(\tilde{\alpha}\right) \int_0^{\tilde{\alpha}} \frac{1}{1-\alpha} d\alpha d\tilde{\alpha} - E\left[P\right] \\
&= \int_0^1 \left[Q_P\left(\alpha\right) - E\left[P\right]\right] \log\left(\frac{1}{1-\alpha}\right) d\alpha
\end{aligned}$$

where $\int_0^1 \log\left(\frac{1}{1-\alpha}\right) d\alpha = 1$.

Parts (2) follows from the fact that P is a mean preserving spread of P_Z

Part (3) can be seen as follows. Because P is a mean-preserving spread of P_Z, we know that

$$\int_x^1 Q_{P_Z}\left(\alpha\right) d\alpha \leq \int_x^1 Q_P\left(\alpha\right) d\alpha \quad \forall x \in [0,1]$$

Now, using part (1), we can write

$$\begin{aligned}
E\left[m\left(P\right)\right] - E\left[m_Z\left(P_Z\right)\right] &= \int_0^1 \left[Q_P\left(\alpha\right) - Q_{P_Z}\left(\alpha\right)\right] log\left(\frac{1}{1-\alpha}\right) d\alpha \\
&= \int_0^1 \left[Q_P\left(\alpha\right) - Q_{P_Z}\left(\alpha\right)\right] \int_0^\alpha \frac{1}{1-\tilde{\alpha}} d\tilde{\alpha} d\alpha \\
&= \int_0^1 \int_0^\alpha \left[Q_P\left(\alpha\right) - Q_{P_Z}\left(\alpha\right)\right] \frac{1}{1-\tilde{\alpha}} d\tilde{\alpha} d\alpha \\
&= \int_0^1 \int_{\tilde{\alpha}}^1 \left[Q_P\left(\alpha\right) - Q_{P_Z}\left(\alpha\right)\right] \frac{1}{1-\tilde{\alpha}} d\alpha d\tilde{\alpha} \\
&= \int_0^1 \left(\int_{\tilde{\alpha}}^1 \left[Q_P\left(\alpha\right) - Q_{P_Z}\left(\alpha\right)\right] d\alpha\right) \frac{1}{1-\tilde{\alpha}} d\tilde{\alpha} \\
&\geq 0
\end{aligned}$$

where the last inequality follows from the fact that $\int_{\tilde{\alpha}}^1 \left[Q_P\left(\alpha\right) - Q_{P_Z}\left(\alpha\right)\right] d\alpha \geq 0$ for all $\tilde{\alpha}$ because P is a mean-preserving spread of P_Z.

Part (4) follows from part (1) and the definition of $e\left(\alpha\right)$.

Part (5) can be seen as follows. Let $T\left(p\right)$ be given by

$$T\left(p\right) = \frac{p + m\left(p\right)}{1 - p - m\left(p\right)} \frac{1-p}{p}$$

which can be re-written as

$$m\left(p\right) \frac{1}{t\left(p\right)} = p\left(1-p\right) - pm\left(p\right)$$

where $t(p) = T(p) - 1$. Taking expectations, we have

$$E\left[\frac{1}{t(P)}m(P)\right] = E[P(1-P)] - E[(P - E[P])m(P)] - M * E[P]$$

where $M = E[m(P)]$ is the magnitude of private information. Now using Holder's inequality ($p = 1$, $q = \infty$),

$$E\left[\frac{1}{t(P)}m(P)\right] \leq \left(\sup\frac{1}{t(P)}\right) E[m(P)]$$

So that

$$E[P(1-P)] - E[(P - E[P])m(P)] - M * E[P] \leq \left(\sup\frac{1}{t(P)}\right) * E[m(P)]$$

so that

$$\inf_p t(p) \leq \frac{E[m(P)]}{E[P(1-P)] - E[(P - E[P])m(P)] - E[m(P)] * E[P]}$$

and thus

$$\inf_p T(p) \leq 1 + \frac{E[m(P)]}{E[P(1-P)] - E[(P - E[P])m(P)] - E[m(P)] * E[P]}$$

Equality when T is constant follows from the fact that the holder inequality would hold with equality (We do not claim there exists a distribution for which $T(p)$ is constant; we only state the fact that equality would hold if T is constant result to give a sense of the extent to which the inequality is potentially violated).

B.1.2 "Tight" lower bound

Here, we show that the lower bound is "tight" in the sense that there exists a joint distribution of L, P, and Z satisfying assumptions 1 and 2 such that $P = P_Z$. This follows relatively trivially. For any elicitation Z, assume that $P = \Pr\{L|X, Z\}$. Then let $e = Z - \Pr\{L|X, Z\}$. Then agents' report $Z = P + e$ but have beliefs given by $\Pr\{L|X, Z\}$.

B.1.3 Measurement Error Example

When do differences in our lower bounds $E[m_Z(P_Z)]$ imply differences in the actual average magnitude of private information, $E[m(P)]$? Here, we consider a stylized form of elicitation error which leads to conditions under which our lower bounds are valid for inferring comparisons for the true values. Intuitively, as long as there is not substantial differential measurement error between rejectees and non-rejectees, such inferences are valid.

Suppose that with probability λ agents report their true beliefs, $Z = P$, but with probability $1 - \lambda$ they report a value Z which is independent of their true beliefs (i.e. random noise). It is straightforward to show that this implies

$$E[m_Z(P_Z)] = \lambda E[m(P)]$$

so that our lower bounds are a fraction λ of the true value. In this case, a finding of $\Delta_Z > 0$ implies $\Delta = E\left[m(P)|X \in \Theta^{Reject}\right] - E\left[m(P)|X \in \Theta^{NoReject}\right] > 0$ as long as $\lambda_{NoReject} \geq \lambda_{Reject}$. Moreover, the event that our lower bounds would be misleading (i.e. $\Delta_Z > 0$ and $\Delta < 0$) requires $\frac{\lambda_{Reject}}{\lambda_{NoReject}} > \frac{E\left[m_Z(P_Z)|X\in\Theta^{Reject}\right]}{E\left[m_Z(P_Z)|X\in\Theta^{NoReject}\right]} > 1$. Thus, the difference in the measurement error must be larger to overturn inference using lower bounds if we estimate much larger values of $E\left[m_Z(P_Z)\,|X\in\Theta^{Reject}\right]$ relative to $E\left[m_Z(P_Z)\,|X\in\Theta^{Reject}\right]$.

B.2 Semiparametric Identification

In this section, we discuss identification of the distribution of private information, f_P, and the distribution of elicitation error parameters, θ. For simplicity, we condition on $X=x$ and drop notation with respect to X.

Our approach assumes the econometrician observes data on Z and L. We make the following assumptions:

- L is a binary random variable (realizations in $\{0,1\}$, indicating the event of experiencing a loss
- The joint density of Z and L is observed and given by the p.d.f./p.m.f. $f_{L,Z}(l,z)$ with conditional distributions $f_{L|Z}$, $f_{Z|L}$, and marginal distribution $f_Z\in D_Z$. for some domain D_Z We assume Z is continuously distributed over $[0,1]$.
- The variable P is unobserved and continuously distributed with density $f_P(p)\in D_P$ for some domain D_P, where we denote the true value by $f_P^*(p)$. We assume D_p is closed under multiplication by p, so that $f_P(p)\in D_p$ implies $pf_P(p)\in D_p$.

Recall we have made several assumptions. First, we have assumed agents have correct beliefs and that Z contains no additional information about L than do agents' true beliefs, P, which together imply

$$Pr\{L|Z,P\}=Pr\{L|P\}=P$$

and second, we have assumed that Z is distributed with p.d.f. $f_{Z|P}(z|P;\theta)$ where $\theta\in\Sigma$ is an unknown parameter from a known set Σ. We denote the true θ by θ^*. Given these assumptions, the density of L and Z can be expressed as

$$\begin{aligned}
f_{L,Z}(L,Z) &= \int_0^1 f_{L,Z|P}(L,Z|P=p)\,f_P(p)\,dp\\
&= \int_0^1 \left(\Pr\{L|Z,P=p\}\right)^L\left(1-\Pr\{L|Z,P=p\}\right)^{1-L} f_{Z|P}(Z|P=p;\theta^*)\,f_P^*(p)\,dp\\
&= \int_0^1 p^L(1-p)^{1-L} f_{Z|P}(Z|P=p;\theta^*)\,f_P^*(p)\,dp
\end{aligned}$$

With this expression for the observed density, our goal is to "invert" the above functional equation to identify both θ^* and $f_P^*(p)$. This problem is made difficult because the functional equation is nonlinear in θ and f_P.

For any $\theta\in\Sigma$, define the operator $H_\theta: D_P\to D_Z$ mapping densities over the space of P into densities over the space of Z by

$$\left[H_\theta(f_P)\right](z)=\int_0^1 f_P(p)\,f_{Z|P}(z|p;\theta)\,dp$$

and let $Z(P|\theta)$ denote the random variable distributed with p.d.f./p.m.f $f_{Z|P}(z|P;\theta)$. Given θ, H_θ is linear in f_P. Therefore, we can impose standard invertibility conditions on $f_{Z|P}$.

Assumption 4. *H_{θ^*} is injective at the true $\theta = \theta^*$ so that $H_{\theta^*}(f_1) = H_{\theta^*}(f_2) \implies f_1 = f_2$*

Injectivitity of H_{θ^*} assumes that if θ^* is known, then the distribution f_P is identified from the density f_Z. This is a mostly standard assumption in linear nonparametric identification (Newey and Powell, 2003 Newey and Powell [2003]; Hu and Schennach, 2008 Hu and Schennach [2008]).

Given this assumption, define the generalized inverse correspondence, H_θ^{-1}, to map f_Z to the set of functions f_P satisfying $H_\theta(f_P) = f_Z$.

$$H_\theta^{-1}(f_Z) = argmin \| f_Z(z) - \int f_{Z|P}(Z|P=p;\theta) f_P(p)\, dp \|$$

where the argmin is taken with respect to densities $f_P \in D_P$. Our assumption of injectivity implies that $H_\theta^{-1}(f_Z)$ is unique if f_Z lies in the range of H_θ and, in particular, is unique at the true value of $\theta = \theta^*$. H_θ^{-1} maps p.d.f.s in the Z space to a set of p.d.f.s in the P space. We also define the corresponding functions, $\hat{H}_\theta$ and $\hat{H}_\theta^{-1}$ which operate on random variables, so that $\hat{H}_\theta(P)$ maps the random variable with p.d.f. f_P to the random variable with p.d.f. $H_\theta(f_P)$.

Now, assumptions 1 and 2 imply that we can write the joint distribution of P and L in two ways by conditioning on $L = 1$,

$$\begin{aligned} f_{P|L}(p|L=1)\Pr\{L=1\} &= f_{P,L}(p,1) \\ &= \Pr\{L=1|P=p\} f_P(p) \\ &= p f_P(p) \end{aligned}$$

Since P has realizations on $[0,1]$, it has a moment generating function, so that we can write the above expression in moment form,

$$E\left[P^N|L=1\right]\Pr\{L\} = E\left[P^{N+1}\right] \quad \forall N \geq 0 \tag{12}$$

which provides a simple relationship between the moments of P given $L = 1$ and the unconditional distribution of P. Equation 12 provides an infinite set of moment conditions which aid in identification of θ and f_P.

At $\theta = \theta^*$, we have

$$E\left[\left(\hat{H}_{\theta^*}^{-1}(Z)\right)^N |L=1\right]\Pr\{L=1\} = E\left[\left(\hat{H}_{\theta^*}^{-1}(Z)\right)^{N+1}\right] \quad \forall N \geq 0 \tag{13}$$

The model is identified if and only if θ^* is the only such θ to generate this equality for all $N \geq 0$. Note that once we have θ^*, we have $f_P^* = H_{\theta^*}^{-1}(f_Z)$.

Because θ is finite-dimensional, the model is generally over-identified. Intuitively, equation 13 for $N = 0$ provides identification of the mean of the elicitation error

$$\Pr\{L\} = E\left[\hat{H}_\theta^{-1}(Z)\right]$$

and the equation for $N = 1$ provides identification of the dispersion in the elicitation error,

$$E\left[\left(\hat{H}_{\theta^*}^{-1}(Z)\right)|L=1\right]\Pr\{L=1\} = E\left[\left(\hat{H}_{\theta^*}^{-1}(Z)\right)^2\right]$$

so that, intuitively, the RHS of the equation varies with the dispersion in the error, holding the mean of the error constant. We recognize that this intuition is fairly abstract because it relies on properties of the operator $H_{\theta^*}^{-1}$, which is a difficult object to know a priori. We proceed on two fronts. First, we provide a formal proof that a close analogue to our specification in Section 7 for which the inverse operator has well-known properties and is identified using only equations $N = 0$ and $N = 1$ of equation 13 (so that the moments $N > 1$ provide a theoretical over-identification test). Second, since our specification in Section 7 does not have such a well-known inverse operator, we provide Monte Carlo tests of our specification. This allows us not only to confirm identification, but also assess the robustness of our results to various possible mis-specifications of the true elicitation error distribution, $f(Z|P)$.

B.2.1 Identification without censoring

Here, we consider an analogue to our model in which non-focal elicitations are not censored on $[0, 1]$. For this specification, the nonlinear inverse problem has well-known properties and θ^* and f_P^* are identified only the observed density, f_Z and equation 13 for $N = 0$ and $N = 1$, leaving equations $N > 1$ as over-identifying conditions.

In particular, suppose Z is distributed $(1-\lambda) N\left(P+\alpha, \sigma^2\right) + \lambda OP\left(P+\alpha, \sigma^2, \kappa\right)$, where $OP\left(P+\alpha, \sigma^2, \kappa\right)$ is an ordered probit with variance σ^2, latent mean $P+\alpha$, and cutoff regions $[0,\kappa]$, $(\kappa, 1-\kappa)$, and $[1-\kappa, 1]$ corresponding to values $Z = 0, 0.5, 1$. Note that our specification in Section 7 is similar but assumes non-focal values follow a censored normal, $CN\left(P+\alpha, \sigma^2\right)$, as opposed to a normal, $N\left(P+\alpha, \sigma^2\right)$, which captures the elicitations lie in $[0, 1]$.

We show identification as follows. First, since Z is continuously distributed for non-focal values, responses of $Z = 0, 0, 5, 1$ occur (with probability 1) as draws from the ordered probit, not the normal distribution. Moreover, because values of $Z = 0, 0.5, 1$ drawn from $N\left(P+\alpha, \sigma^2\right)$ occur with probability zero, we can consider identification of α and σ from the observed density of non-focal values of Z, which is drawn from $N\left(P+\alpha, \sigma^2\right)$. Thus, we now consider this simpler elicitation error distribution and return to the identification of λ and κ after discussing identification of α and σ.

Let $H_{\alpha,\sigma}(f_P)$ map the p.d.f. of a random variable P, f_P, to the p.d.f. of the random variable $Z = P + e\left(\alpha, \sigma^2\right)$ where $e = N\left(\alpha, \sigma^2\right)$ is independent of P. Thus, Z is a mean preserving spread of $P+\alpha$, with σ indexing the degree of the "spread". Moreover, because the elicitation error, e, is normally distributed, the inverse operator, $H_{\alpha,\sigma}^{-1}$, maps a p.d.f. of the random variables Z to the p.d.f. of the random variable $\hat{P}(Z;\alpha,\sigma)$ with the shift in mean α and whose variance is strictly decreasing in σ.

Recall that at the true values, $\theta^* = (\alpha^*, \sigma^*)$, we have the equations

$$E\left[\left(\hat{P}(Z;\alpha^*,\sigma^*)\right)^N |L=1\right] \Pr\{L=1\} = E\left[\left(\hat{P}(Z;\alpha^*,\sigma^*)\right)^{N+1}\right] \quad \forall N \geq 0$$

So, for $N = 0$, we have the equation

$$\begin{aligned} \Pr\{L=1\} &= E\left[\hat{P}(Z;\alpha^*,\sigma^*)\right] \\ &= E[Z] - \alpha^* \end{aligned}$$

so that $\alpha^* = E[Z] - \Pr\{L=1\}$. Intuitively, the mean bias is identified as the difference between the average elicitation, $E[Z]$, and the realized probability of a loss, $\Pr\{L=1\}$.

For $N = 1$, we have the equation

$$E\left[\left(\hat{P}\left(Z;\alpha^*,\sigma^*\right)\right)|L=1\right]\Pr\left\{L=1\right\} = E\left[\left(\hat{P}\left(Z;\alpha^*,\sigma^*\right)\right)^2\right]$$

Now, notice that the LHS does not vary with σ. Moreover, the RHS is monotonically decreasing in σ, since $\sigma_1 < \sigma_2$ implies that $\hat{P}\left(Z;\alpha,\sigma_1\right)$ is a mean preserving spread of $\hat{P}\left(Z;\alpha,\sigma_2\right)$. Thus, the equation for $N = 1$ identifies σ^*.

Now that we have identified α and σ, we return to our original distribution with focal point responses. First, notice that λ is identified using the fraction of responses Z which are equal to 0, 0.5, or 1. Then, κ is identified by the relative frequency of $Z = 0$, $Z = 0.5$, and $Z = 1$ using the already identified values of α and σ.

This example with uncensored non-focal point values is identical to our specification in Section 7, except that we use a censored normal, as opposed to normal distribution, to take into account the fact that elicitations are restricted to the interval, $[0, 1]$. The impact of such censoring on the quality of our estimation is difficult to assess theoretically; we thus turn to Monte Carlo evidence to verify the performance of our estimation strategy.

B.2.2 Monte Carlo Results

This section presents Monte Carlo analysis of our estimator for the distribution of private information. First, we verify that our approach works well under correct model specification for the elicitation error parameters. Second, we assess the impact of mis-specification of the distribution of elicitation error. Throughout this section, we assume that $F(p)$ is a censored normal distribution with mean 0.3 and standard deviation 0.1. Our first simulation assumes Z is follows our specification. With probability 0.6, Z is drawn from a censored normal distribution with mean $P + 0.03$ and standard deviation 0.2. With probability 0.4, Z is a focal point value of 0, 0.5, or 1, drawn from an ordered probit distribution with mean $P + 0.03$, standard deviation 0.2, and cutoff regions $[0, 0.3], (0.3, 0.7), [0.7, 1]$. In Figure B1, we present the true c.d.f. of private information, along with the median, 5%, and 95% estimates from $N = 100$ Monte Carlo simulations (of a sample size of 2,000) where we estimate the distribution using a mixture of 2 beta distributions, as in our empirical analysis above. As the figures show, our estimation approach yields unbiased estimates.

Now we consider the impact of mis-specification of the elicitation error. To do so, we assume that the latent Z is drawn from a mixture of two normals, allowing us to incorporate skewness and kurtosis in the distribution. First, we assume with probability 0.4. the latent Z is drawn from a normal with mean $P + 0.03$ and standard deviation 0.2 (as before). But with probability 0.6, the latent Z is drawn from a normal with mean $P - 0.05$ and standard deviation 0.4. Focal values (again we assume 40% focal values) are then generated from this latent Z with the same cutoff regions, $[0, 0.3], (0.3, 0.7), [0.7, 1]$, and non-focal values are generated as the censored portion of this distribution on $[0, 1]$. The Monte Carlo results are presented in Figure B2. As we can see, our estimation perhaps introduces a slight median bias towards a less dispersed distribution of private information, but performs quite well given this substantial mis-specification.

Finally, we assess the robustness to excess kurtosis in the distribution of Z. We assume Z is again drawn from a mixture of normals, but this time assume these normals have the same mean of 0.03. With probability 0.5, the standard deviation is 0.2 and with probability 0.5 the standard deviation is 0.05.

We assume 40% focal responses with the same cutoffs regions of $[0, 0.3], (0.3, 0.7), [0.7, 1]$. Figure B3 presents the Monte Carlo results. As we can see, our estimation performs quite well (better than the skewed estimation) despite the mis-specification. In short, our estimation procedure appears robust to alternative specifications for $f(Z|P)$ which relax normality by including skewness and kurtosis.

C Selected Pages from Genworth Financial Underwriting Guidelines

The following 4 pages contain a selection from Genworth Financial's LTC underwriting guideline which is provided to insurance agents for use in screening applicants. Although marked "Not for use with consumers or to be distributed to the public", these guidelines are commonly left in the public domain on the websites of insurance brokers. The printed version here was found in public circulation at http://www.nyltcb.com/brokers/pdfs/Genworth_Underwriting_Guide.pdf on November 4, 2011. We present 4 pages of the 152 pages of the guidelines. The conditions documented below are not exhaustive for the list of conditions which lead to rejection - they constitute the set of conditions which solely lead to rejection (independent of other health conditions); combinations of other conditions may also lead to rejections and the details for these are provided in the remaining pages not shown here.

Genworth®
Financial

LONG TERM CARE INSURANCE UNDERWRITING GUIDE

PROVIDED BY THE GENWORTH UNDERWRITING DEPARTMENT

Long Term Care Insurance Underwritten
by Genworth Life Insurance Company,
and in New York
by Genworth Life Insurance Company of New York
Administrative Offices: Richmond, VA.

For agent use only. Not for use with consumers or to be distributed to the public.

85440 06/10/09

INTRODUCTION

Underwriting is the process by which an applicant's current health, medical history and lifestyle are evaluated to determine a risk profile. The underwriter's decision to accept or decline an applicant is determined by matching the profile to guidelines, which outline the limits of acceptable risk to the company.

We underwrite applicants in the age range 18-79. We do not modify the coverage applied for, nor do we apply extra premiums. We make every attempt to issue the desired coverage at the corresponding published premium.

The information in this manual reflects over 30 years of experience...the longest in the Long Term Care insurance industry. While not all-inclusive, enough information is presented to help you in most situations you will encounter. A hotline number is included should you have questions or run into an unusual circumstance.

An appeal process is also outlined in the event you disagree with our underwriting evaluation. We are always willing to have a second look, especially when additional information not included in the original application file is made available.

We value our relationship with you and look forward to providing high quality service and underwriting for you and your clients.

UNINSURABLE CONDITIONS

Acquired Immune Deficiency Syndrome (AIDS)
ADL limitation, present
AIDS Related Complex (ARC)
Alzheimer's Disease
Amputation due to disease, e.g., diabetes or atherosclerosis
Amyotrophic Lateral Sclerosis (ALS) , Lou Gehrig's Disease
Ascites present
Ataxia, Cerebellar
Autonomic Insufficiency (Shy-Drager Syndrome)
Autonomic Neuropathy (excluding impotence)
Behçet's Disease
Binswanger's Disease
Bladder incontinence requiring assistance
Blindness due to disease or with ADL/IADL limitations
Bowel incontinence requiring assistance
Buerger's Disease (thromboangiitis obliterans)
Cerebral Vascular Accident (CVA)
Chorea
Chronic Memory Loss
Cognitive Testing, failed
Cystic Fibrosis
Dementia
Diabetes treated with insulin
Dialysis, Kidney (Renal)
Ehlers-Danlos Syndrome
Forgetfulness (frequent or persistent)
Gangrene due to diabetes or peripheral vascular disease
Hemiplegia
Hoyer Lift
Huntington's or other forms of Chorea
Immune Deficiency Syndrome
Korsakoff's Psychosis
Leukemia-except for Chronic Lymphocytic Leukemia (CLL) and Hairy Cell Leukemia (HCL)
Marfan's Syndrome
Medications

- Antabuse (disulfiram)
- Aricept (donepezil HCl)
- Campral (acamprosate calcium)
- Cognex (tacrine)
- Depade (naltrexone)
- Exelon (rivastigmine)
- Hydergine (ergoloid mesylate)
- Namenda (memantine)
- Razadyne (galantamine hydrobromide)
- Reminyl (galantamine hydrobromide)
- ReVia (naltrexone)
- Vivitrol (naltrexone)

Memory Loss, chronic
Mesothelioma
Multiple Sclerosis (MS)

Muscular Dystrophy (MD)
Myelofibrosis
Organ Transplants, except kidney transplants
Organic Brain Syndrome (OBS)
Oxygen use except if used for headaches or sleep apnea
Paralysis/Paraplegia
Parkinson's Disease
Pneumocystis Pneumonia
Polyarteritis Nodosa
Postero-Lateral Sclerosis
Quad Cane use
Quadriplegia
Senility
Spinal Cord Injury with ADL/IADL limitations
Stroke (CVA)
Surgery scheduled or anticipated (except cataract surgery under local anesthesia)
Takayasu's Arteritis
Thalassemia Major
Total Parenteral Nutrition (TPN) for regular or supplementary feeding or administration of medication
Waldenstrom's Macroglobulinemia
Walker use
Wegener's Granulomatosis
Wernicke-Korsakoff Syndrome
Wheelchair use
Wilson's Disease

Table 1: Rejection Classification

	Long-Term Care		Disability		Life	
Classification	Condition	% Sample	Condition	% Sample	Condition	% Sample
Rejection	Any ADL/IADL Restriction	6.5%	Back Condition	22.7%	Cancer[4] (Current)	13.1%
	Past Stroke	7.8%	Obesity (BMI > 40)	1.7%	Stroke (Ever)	7.3%
	Past Nursing/Home Care	12.4%	Psychological Condition	6.3%		
	Over age 80	18.4%				
Uncertain	Lung Disease	10.0%	Arthritis	36.9%	Diabetes	13.8%
	Heart Condition	28.4%	Diabetes	7.7%	High Blood Pressure	50.7%
	Cancer (Current)	14.7%	Lung Disease	5.1%	Lung Disease	10.9%
	Hip Fracture	1.3%	High Blood Pressure	35.3%	Cancer (Ever, not current)	12.1%
	Memory Condition[1]	0.8%	Heart Condition	6.1%	Heart Condition	26.5%
	Other Major Health Problems[2]	26.7%	Cancer (Ever Have)	4.6%	Other Major Health Problems[2]	23.5%
			Blue-collar/high-risk Job[3]	23.3%		
			Other Major Health Problems[2]	16.2%		

[1]Memory conditions generally lead to rejection, but were not explicitly asked in waves 2-3; we classify memory conditions as uncertain for consistency, since they would presumably be considered an "other" condition in waves 2-3.

[2]Wording of the question varies slightly over time, but generally asks: "Do you have any other major/serious health problems which you haven't told me about?"

[3]We define blue collar/high-risk jobs as non-self employed jobs in the cleaning, foodservice, protection, farming, mechanics, construction, and equipment operators

[4]We exclude minor basel cell cancers

Note that percentages will not add to the total fraction of the population classifed as rejection and uncertain because of people with multiple conditions

Table 2: Covariate Specifications

Long-Term Care		Disability		Life	
Price Controls	Extended Controls	Price Controls	Extended Controls	Price Controls	Extended Controls
Age, Age^2, Gender Gender*age Gender*age^2 Word Recall Performance[1] Indicators for ADL/IADL Restriction Psych Condition Diabetes Lung Disease Arthritis Heart Disease Cancer Stroke High blood pressure	Full interactions of Age Gender Word Recall Performance[1] Indicators for ADL/IADL Restriction Psychological Condition Diabetes Lung Disease Arthritis Heart Disease Cancer Stroke High blood pressure Interactions between 5 yr age bins and the presence of: Number of Health Conditions (High bp, diabetes, heart condition, lung disease, arthritis, stroke, obesity, psych condition) Number of ADL / IADL Restrictions Number of living relatives (<=3) Past home care usage Census region (1-5) Income Decile	Age, Age^2, Gender Gender*age Gender*age^2 Indicators for Self Employed Obese Psych condition Back condition Diabetes Lung Disease Arthritis Heart Condition Cancer Stroke High Blood Pressure BMI Wage Decile	Full interactions of Age Gender Full interactions of wage decile part time indicator job tenure quartile self-employment indicator Interactions between 5 yr age bins and the presence of: Arthritis Diabettes Lung disease Cancer Heart condition Psychological condition Back condition BMI Quartile Full interactions of BMI quartile 5 year age bins Full interactions of Job requires stooping Job requires lifting Job requires phys activity Full Interactions of 5 year age bins Census region (1-5)	Age, Age^2, Gender Gender*age Gender*age^2 Smoker Status Indicator for years to question[2] Indicator for death of parent before age 60 BMI Indicators for Psychological Condition Diabetes Lung Disease Arthritis Heart Disease Cancer Stroke High blood pressure Income decile	Full interactions of Age Gender Full Interactions of age AGE in subj prob question Interactions of 5 yr age bins with: Smoker Status Income Decile Heart condition Stroke Cancer Lung disease Diabetes High blood pressure Census Region BMI Indicator for death of parent before age 60

[1]Indicator for lowest quartile performance on word recall test
[2]Full indicator variables for number of years to AGE reported in subjective probability question

Table 3: Sample Summary Statistics

	Long-Term Care			Disability			Life		
	No Reject	Reject	Uncertain	No Reject	Reject	Uncertain	No Reject	Reject	Uncertain
Subj. Prob (mean)[1]	0.111	0.168	0.132	0.292	0.385	0.352	0.366	0.556	0.491
(std dev)	(0.194)	(0.249)	(0.207)	(0.257)	(0.264)	(0.262)	(0.313)	(0.341)	(0.337)
Loss	0.039	0.175	0.054	0.156	0.441	0.339	0.273	0.572	0.433
	(0.195)	(0.38)	(0.227)	(0.363)	(0.497)	(0.473)	(0.446)	(0.495)	(0.496)
Demographics									
Age	71.7	79.4	72.2	54.7	55.0	55.6	70.4	75.3	72.9
	(4.366)	(6.934)	(4.303)	(4)	(4.016)	(3.679)	(7.627)	(7.785)	(7.548)
Female	0.622	0.631	0.564	0.606	0.602	0.551	0.595	0.564	0.588
	(0.485)	(0.483)	(0.496)	(0.489)	(0.49)	(0.497)	(0.491)	(0.496)	(0.492)
Health Status Indicators									
Arthritis	0.479	0.616	0.552	0.000	0.553	0.510	0.351	0.435	0.443
	(0.5)	(0.486)	(0.497)	(0)	(0.497)	(0.5)	(0.477)	(0.496)	(0.497)
Diabetes	0.140	0.172	0.147	0.000	0.090	0.121	0.000	0.163	0.185
	(0.347)	(0.377)	(0.354)	(0)	(0.287)	(0.326)	(0)	(0.369)	(0.388)
High Blood Pressure	0.505	0.598	0.535	0.280	0.392	0.378	0.000	0.574	0.685
	(0.5)	(0.49)	(0.499)	(0.449)	(0.488)	(0.485)	(0)	(0.495)	(0.465)
Sample Size									
Observations (Ind x wave)	9,051	10,108	10,690	2,540	2,216	3,757	2,689	2,362	6,800
Unique Individuals	4,418	3,215	5,190	1,480	1,280	1,929	1,720	1,371	4,270
Unique Households	3,283	2,620	3,860	1,112	975	1,540	1,419	1,145	3,545
Fraction Insured[2]	13.9%	10.7%	14.8%				65.1%	63.3%	64.2%

[1]We transform the life insurance variable to 1-Pr{living to AGE} to correspond to the loss definition

[2]Calculated based on full sample prior to excluding individuals who purchased insurance

Table 4: Magnitude of Private Information (Lower Bound)

	LTC			Disability			Life		
	Age & Gender	Price Controls	Extended Controls	Age & Gender	Price Controls	Extended Controls	Age & Gender	Price Controls	Extended Controls
Difference: Δ_Z	**0.0234*****	**0.0245*****	**0.0213*****	**0.0445*****	**0.0255****	**0.0234****	**0.0449*****	**0.0338*****	**0.0397*****
Bootstrap s.e.[1]	(0.0041)	(0.004)	(0.004)	(0.0115)	(0.0109)	(0.0098)	(0.0113)	(0.0109)	(0.0097)
p-value[2]	0.0000	0.0000	0.0000	0.0000	0.0120	0.0200	0.0000	0.0020	0.0000
No Reject	0.0041	0.0041	0.0040	0.0282***	0.0257***	0.027***	0.031**	0.025	0.021
Bootstrap s.e.	(0.0015)	(0.0017)	(0.0019)	(0.0061)	(0.006)	(0.0058)	(0.0074)	(0.0067)	(0.0061)
Wald test p-value[3]	0.3880	0.4330	0.2156	0.0009	0.0056	0.0055	0.0102	0.1187	0.2395
Reject	0.0274***	0.0286***	0.0253***	0.0727***	0.0512***	0.0504***	0.0759***	0.0587***	0.0604***
Bootstrap s.e.	(0.0038)	(0.0036)	(0.0034)	(0.0098)	(0.0089)	(0.0082)	(0.0085)	(0.0088)	(0.0081)
Wald test p-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000
Uncertain	0.0058**	0.0056**	0.0053**	0.0567***	0.0421***	0.0407***	0.0463***	0.0294***	0.028***
Bootstrap s.e.	(0.002)	(0.0018)	(0.0019)	(0.0072)	(0.0066)	(0.0064)	(0.0058)	(0.0053)	(0.0051)
Wald test p-value	0.0121	0.0472	0.0147	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001

[1]Bootstrapped standard errors computed using block re-sampling at the household level (results shown for N=500 repetitions)

[2]p-value is the sum of the p-value for the rejection group having no private information and the p-value for the hypothesis that the difference is less than or equal to zero, where the latter is computed using bootstrap (N=50 repetitions)

[3]p-value for the Wald test which restricts coefficients on subjective probabilities equal to zero. Standard errors clustered at the household level

*** p<0.01, ** p<0.05, * p<0.10

Table 5: Robustness Checks: Sample Selection

	LTC, Price Controls		Life, Price Controls	
	Primary Sample	Excluding Insured	Primary Sample	Excluding Insured
Difference: Δ_Z	**0.0245*** **	**0.0257*** **	**0.0338*** **	**0.011**
Bootstrap s.e.[1]	(0.004)	(0.0043)	(0.0109)	(0.0166)
p-value[2]	0.0000	0.0000	0.0020	0.2960
No Reject	0.0041	0.0033	0.0249	0.0377
Bootstrap s.e.	(0.0017)	(0.0018)	(0.0067)	(0.0112)
Wald test p-value[3]	0.4330	0.8280	0.1187	0.2334
Reject	0.0286***	0.029***	0.0587***	0.0491*
Bootstrap s.e.	(0.0036)	(0.0039)	(0.0088)	(0.0116)
Wald test p-value	0.0000	0.0000	0.0000	0.0523
Uncertain	0.0056**	0.0056	0.0294***	0.0269
Bootstrap s.e.	(0.0018)	(0.0019)	(0.0053)	(0.008)
Wald test p-value	0.0472	0.1352	0.0001	0.1560

[1]Bootstrapped standard errors computed using block re-sampling at the household level (results shown for N=500 repetitions)

[2]p-value is the sum of the p-value for the rejection group having no private information and the p-value for the hypothesis that the difference is less than or equal to zero, where the latter is computed using bootstrap (N=500 repetitions)

[3]p-value for the Wald test which restricts coefficients on subjective probabilities equal to zero. Standard errors clustered at the household level

*** p<0.01, ** p<0.05, * p<0.10

Table 6:Robustness Checks: Cancer Organ Controls (Life Setting)

	Preferred Specification	Organ + Extended Controls (1993/1994 Only)
Difference: Δ_Z	**0.0338*** **	**0.0308** **
Bootstrap s.e.[1]	(0.0109)	(0.0121)
p-value[2]	0.0020	0.0140
No Reject	0.0249	0.0218
Bootstrap s.e.	(0.0067)	(0.007)
Wald test p-value[3]	0.1187	0.3592
Reject	0.0587***	0.0526***
Bootstrap s.e.	(0.0088)	(0.01)
Wald test p-value	0.0000	0.0024
Uncertain	0.0294***	0.0342***
Bootstrap s.e.	(0.0053)	(0.0061)
Wald test p-value	0.0001	0.0003

[1]Bootstrapped standard errors computed using block re-sampling at the household level (results shown for N=500 repetitions)

[2]p-value is the sum of the p-value for the rejection group having no private information and the p-value for the hypothesis that the difference is less than or equal to zero, where the latter is computed using bootstrap (N=500 repetitions)

[3]p-value for the Wald test which restricts coefficients on subjective probabilities equal to zero. Standard errors clustered at the household level

*** p<0.01, ** p<0.05, * p<0.10

Table 7: Minimum Pooled Price Ratio

	LTC				Disability				Life			
Quantile of Index, q	Mean	20%	50%	80%	Mean	20%	50%	80%	Mean	20%	50%	80%
Reject	**1.715**	**1.681**	**1.711**	**1.730**	**1.954**	**1.900**	**1.937**	**2.282**	**1.727**	**1.642**	**1.751**	**2.269**
5%	1.538	1.525	1.523	1.602	1.890	1.851	1.878	2.259	1.500	1.483	1.582	1.987
95%	1.784	1.837	1.777	1.768	2.076	2.025	2.043	2.315	2.241	2.286	2.330	2.415
Pr{L|reject}	0.175	0.094	0.157	0.244	0.441	0.293	0.430	0.578	0.572	0.351	0.589	0.791
No Reject	**1.206**	**1.337**	**1.247**	**1.147**	**1.611**	**1.703**	**1.626**	**1.572**	**1.361**	**1.640**	**1.406**	**1.345**
5%	1.000	1.000	1.000	1.000	1.243	1.256	1.248	1.250	1.000	1.000	1.000	1.000
95%	1.311	1.793	1.453	1.251	2.346	3.017	2.442	2.274	1.397	1.768	1.444	1.487
Pr{L|No Reject}	0.039	0.017	0.030	0.057	0.156	0.109	0.146	0.197	0.273	0.073	0.194	0.458
Difference (Reject - Nc	**0.509**	**0.344**	**0.464**	**0.582**	**0.343**	**0.198**	**0.312**	**0.710**	**0.366**	**0.002**	**0.345**	**0.923**
5%	0.329	0.144	0.150	0.437	-1.181	-1.938	-1.439	-0.016	0.116	-0.246	0.146	0.681
95%	0.582	0.508	0.551	0.622	0.695	0.639	0.684	1.032	0.950	0.444	1.013	1.254

Note: 5/95% CI computed using bootstrap block re-sampling at the household level (N=250 Reps); 5% level extended to include 1 if p-value of F-test for presence of private information is less than .05; Bootstrap CI is bias corrected using the non-accelerated procedure in Efron (1982).

Table 8: Minimum Pooled Price Ratio

	LTC			Disability			Life		
Quantile Region: Ψ_τ	0-70%	0-80%	0-90%	0-70%	0-80%	0-90%	0-70%	0-80%	0-90%
Reject	**1.715**	**1.715**	**1.715**	**2.350**	**1.954**	**1.727**	**1.865**	**1.727**	**1.572**
5%	1.538	1.538	1.627	2.216	1.890	1.682	1.626	1.500	1.415
95%	1.784	1.784	1.782	2.549	2.076	1.817	2.577	2.241	2.110
No Reject	**1.206**	**1.206**	**1.206**	**1.611**	**1.611**	**1.611**	**1.444**	**1.361**	**1.281**
5%	1.000	1.000	1.000	1.272	1.243	1.247	1.000	1.000	1.000
95%	1.311	1.311	1.311	2.346	2.346	2.089	1.457	1.397	1.286
Difference	**0.509**	**0.509**	**0.509**	**0.739**	**0.343**	**0.117**	**0.421**	**0.366**	**0.291**
5%	0.329	0.329	0.407	-0.252	-1.181	-0.582	0.170	0.116	0.115
95%	0.582	0.582	0.590	1.085	0.695	0.443	1.134	0.950	0.817

Note: 5/95% CI computed using bootstrap block re-sampling at the household level (N=250 Reps); 5% level extended to include 1 if p-value of F-test for presence of private information is less than .05; Bootstrap CI is bias corrected using the non-accelerated procedure in Efron (1982).

Table 9: Willingness to Pay Calibration

	Coeff. Rel. Risk Aversion		
Consumption drop	1	2	3
10%	11.1%	23.5%	37.2%
15%	17.6%	38.4%	62.8%
20%	25.0%	56.3%	95.3%

Table 10: Elicitation Error Parameters

	LTC		Disability		Life	
	No Reject	Reject	No Reject	Reject	No Reject	Reject
Standard Deviation	0.287	0.320	0.305	0.100	0.384	0.427
s.e.	(0.035)	(0.011)	(0.031)	(0.006)	(0.016)	(0.012)
Fraction Focal Respondents	0.372	0.496	0.343	0.508	0.386	0.392
s.e.	(0.055)	(0.017)	(0.023)	(0.011)	(0.015)	(0.014)
Focal Window	0.179	0.238	0.002	0.193	0.033	0.031
s.e.	(0.024)	(0.018)	(0.028)	(0.007)	(0.028)	(0.011)

Note: Bootstrapped standard errors computed using block re-sampling at the household level (results shown for N=250 repetitions)

Table 11: Minimum Pooled Price Ratio (Annuities)

	Annuities		
Quantile Region: Ψ_τ	0-70%	0-80%	0-90%
No Reject	**1.2227**	**1.1770**	**1.1523**
5%	1.0000	1.0000	1.0000
95%	1.4045	1.2651	1.2651
Reject	**1.405**	**1.334**	**1.268**
5%	1.248	1.221	1.227
95%	1.736	1.720	1.720

Note: 5/95% CI computed using bootstrap block re-sampling at the household level (N=250 Reps); 5% level extended to include 1 if p-value of F-test for presence of private information is less than .05; Bootstrap CI is bias corrected using the non-accelerated procedure in Efron (1982)

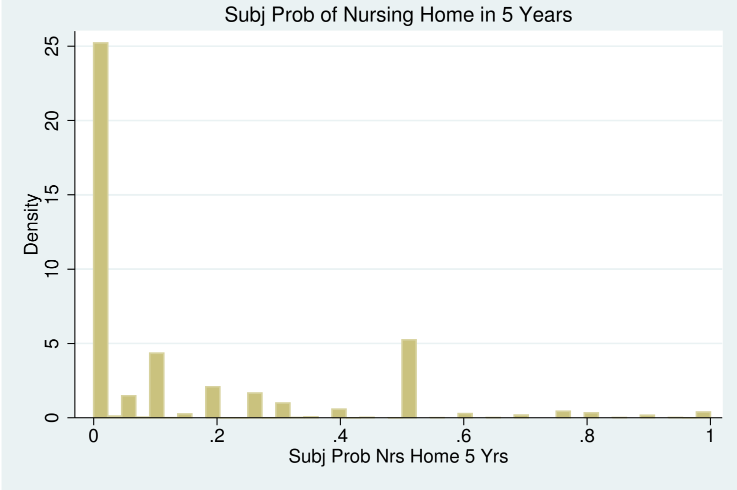


Figure 1a: Long–term Care

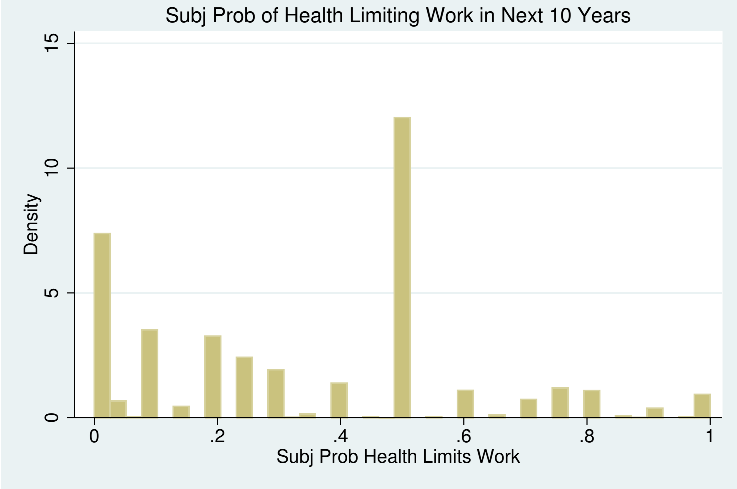


Figure 1b: Disability

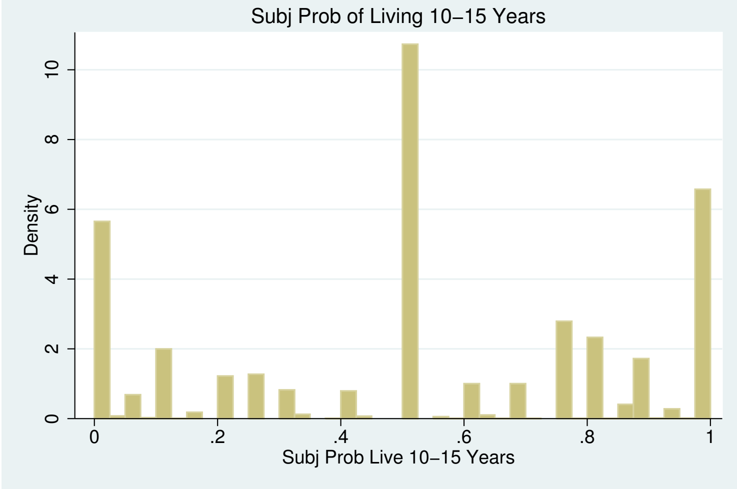


Figure 1c: Life

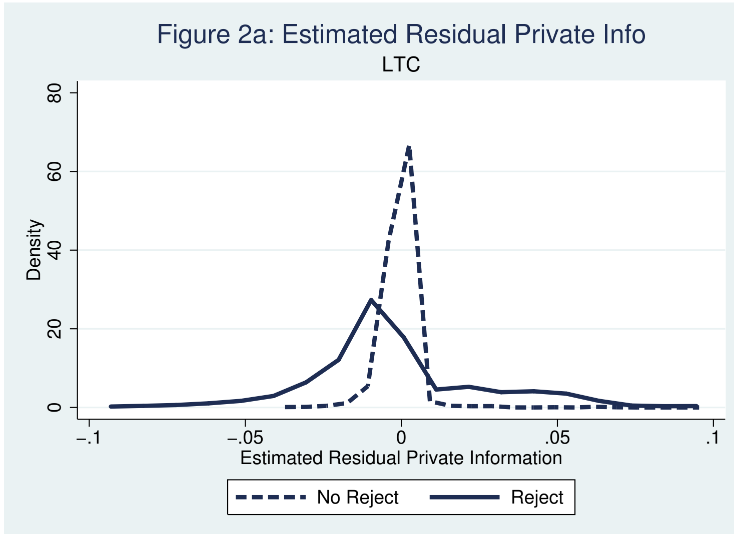


Figure 2a: Estimated Residual Private Info
LTC
Density
80
60
40
20
0
−.1
−.05
0
.05
.1
Estimated Residual Private Information
No Reject
Reject

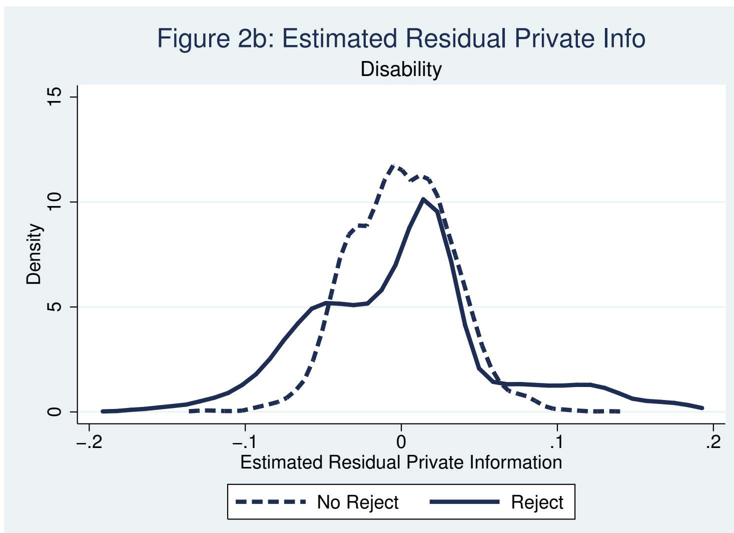


Figure 2b: Estimated Residual Private Info
Disability
Density
15
10
5
0
−.2
−.1
0
.1
.2
Estimated Residual Private Information
No Reject
Reject

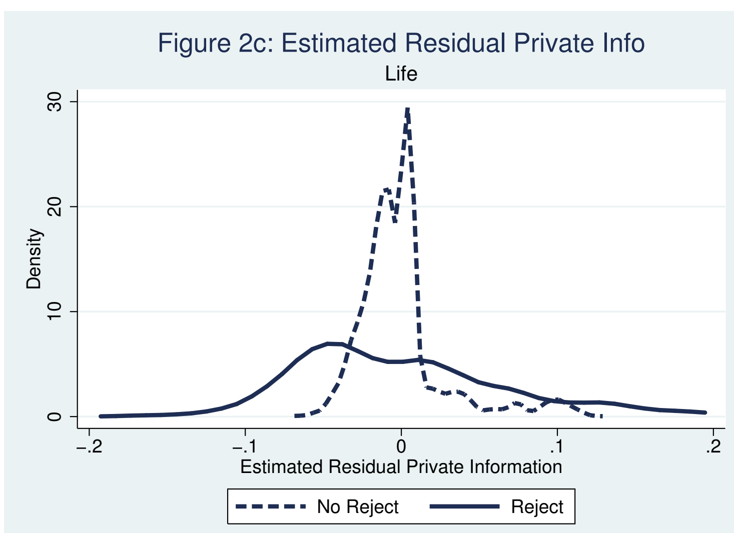


Figure 2c: Estimated Residual Private Info
Life
Density
30
20
10
0
−.2
−.1
0
.1
.2
Estimated Residual Private Information
No Reject
Reject

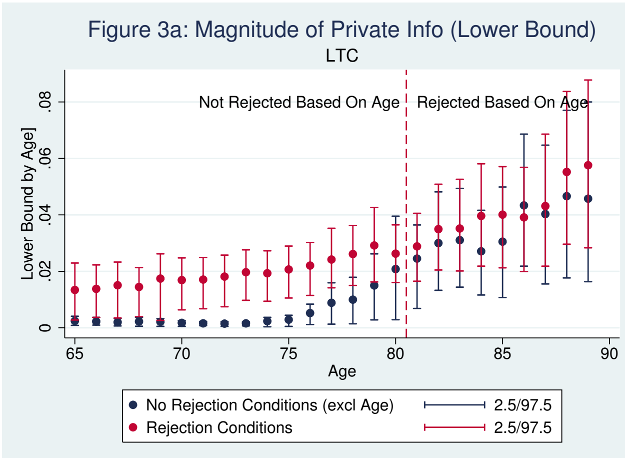


Figure 3a: Magnitude of Private Info (Lower Bound)
LTC
Lower Bound by Age]
Not Rejected Based On Age
Rejected Based On Age
Age
No Rejection Conditions (excl Age)
Rejection Conditions
2.5/97.5
2.5/97.5

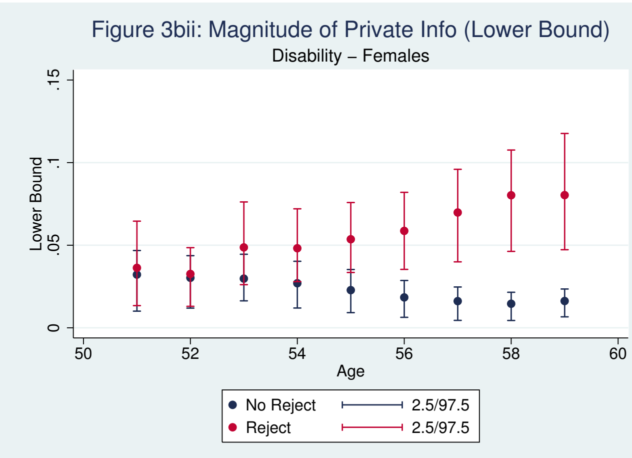


Figure 3bii: Magnitude of Private Info (Lower Bound)
Disability – Females
Lower Bound
Age
No Reject
Reject
2.5/97.5
2.5/97.5

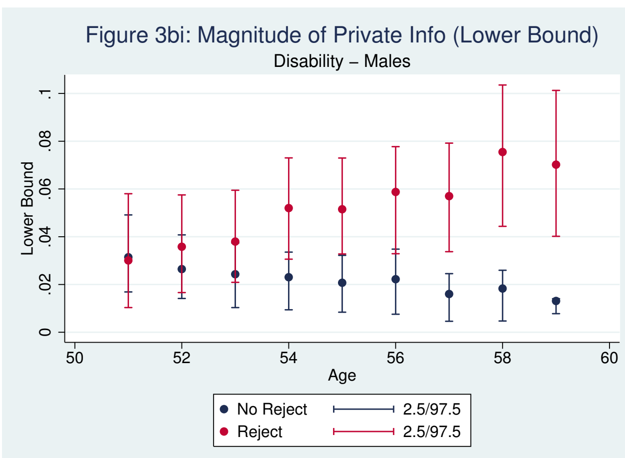


Figure 3bi: Magnitude of Private Info (Lower Bound)
Disability – Males
Lower Bound
Age
No Reject
Reject
2.5/97.5
2.5/97.5

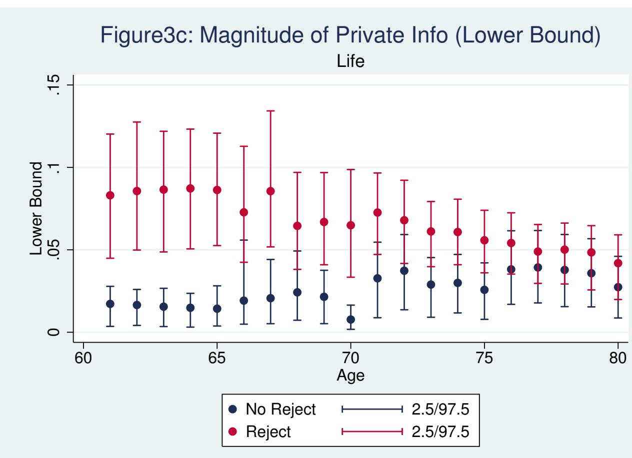


Figure3c: Magnitude of Private Info (Lower Bound)
Life
Lower Bound
Age
No Reject
Reject
2.5/97.5
2.5/97.5

Figure 4a: Distribution of Private Information

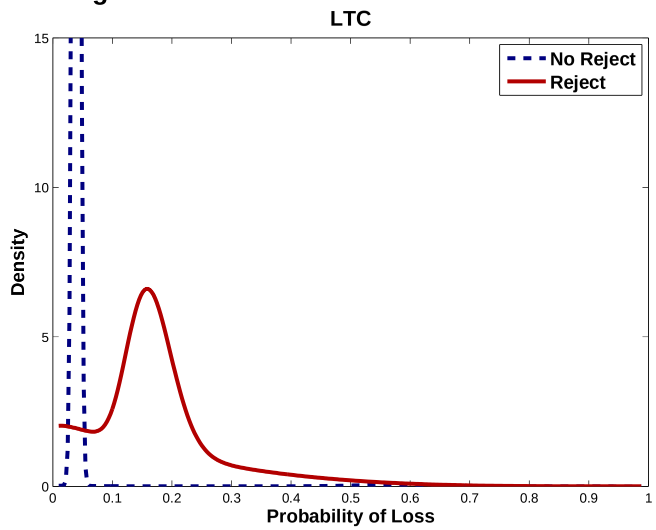


Figure 4b: Distribution of Private Information

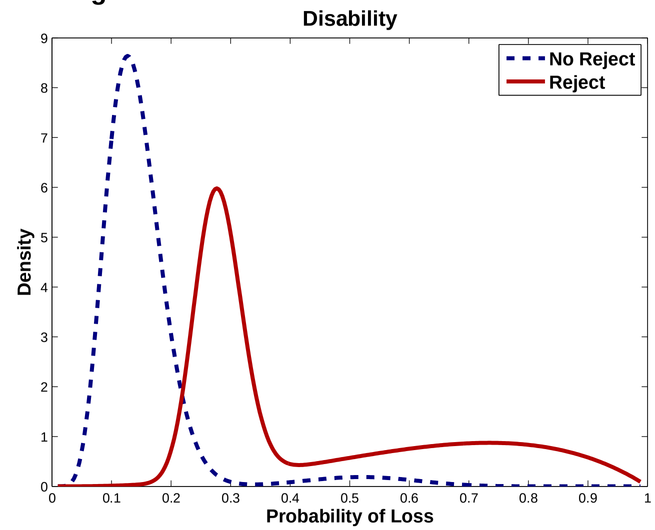


Figure 4c: Distribution of Private Information

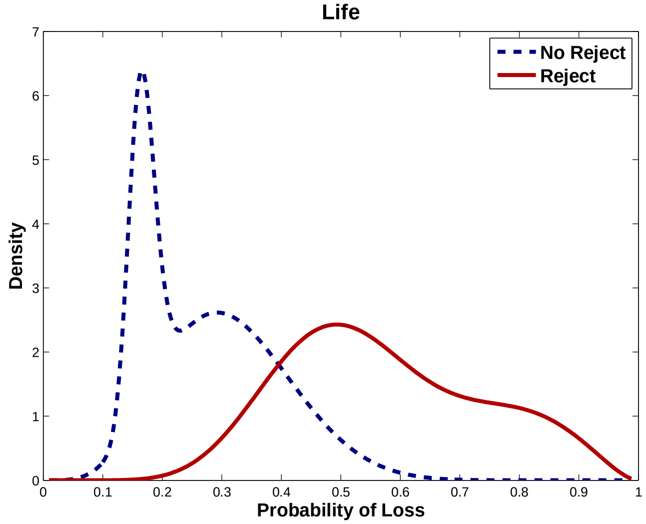


Figure 5a: CDF of Private Information

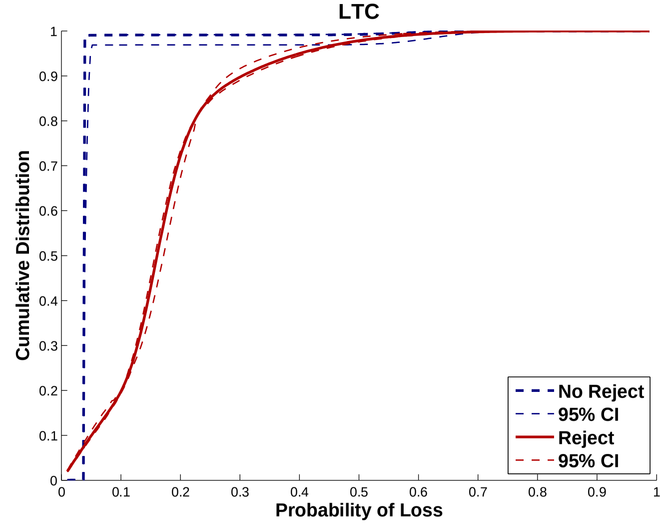


Figure 5b: CDF of Private Information

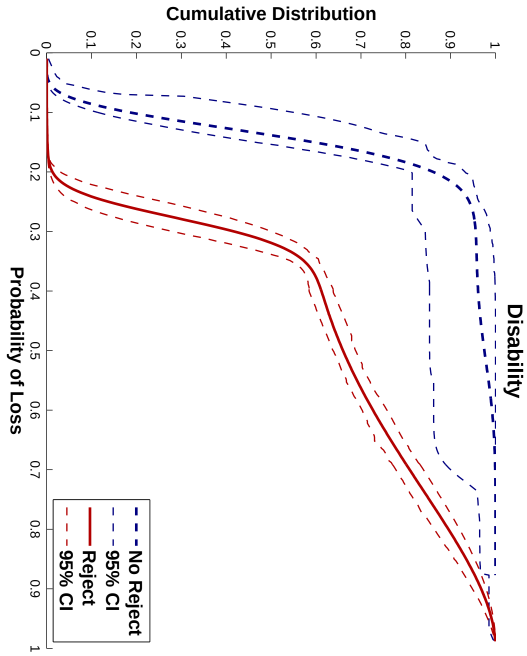


Figure 5c: CDF of Private Information

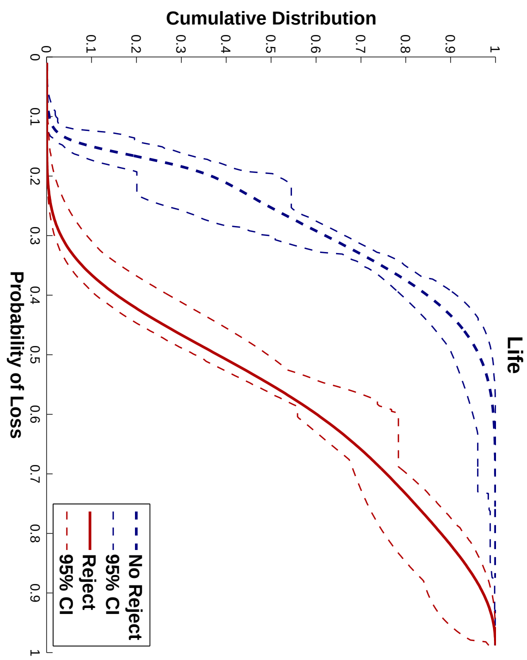
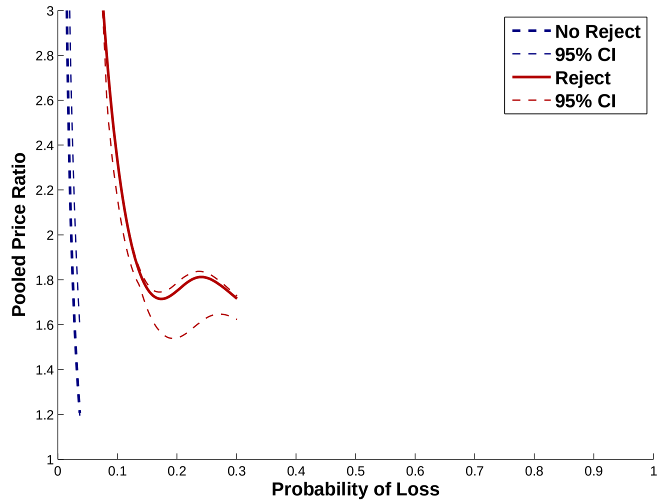


Figure 6a: Pooled Price Ratio

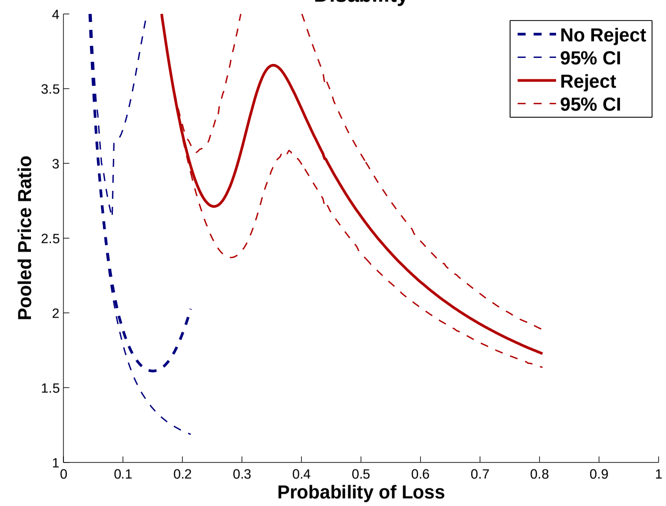
LTC



Note: Plot censored above 90th estimated pctile of F(p)

Figure 6b: Pooled Price Ratio

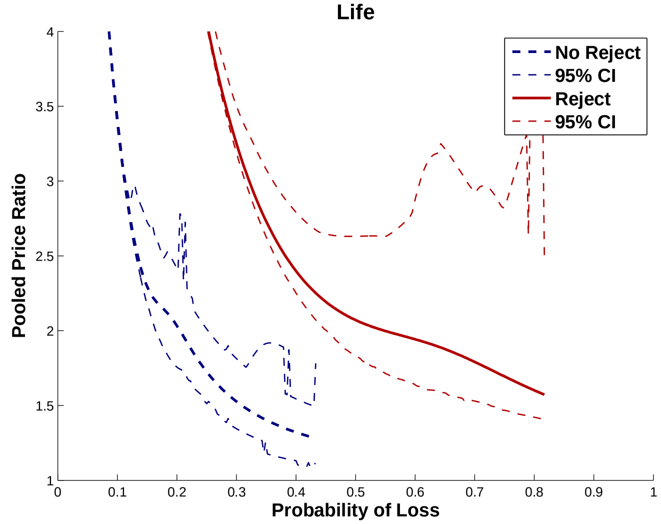
Disability



Note: Plot censored above 90th estimated pctile of F(p)

Figure 6c: Pooled Price Ratio

Life



Note: Plot censored above 90th estimated pctile of F(p)

Figure B1: Monte Carlo Simulation

Correct Specification

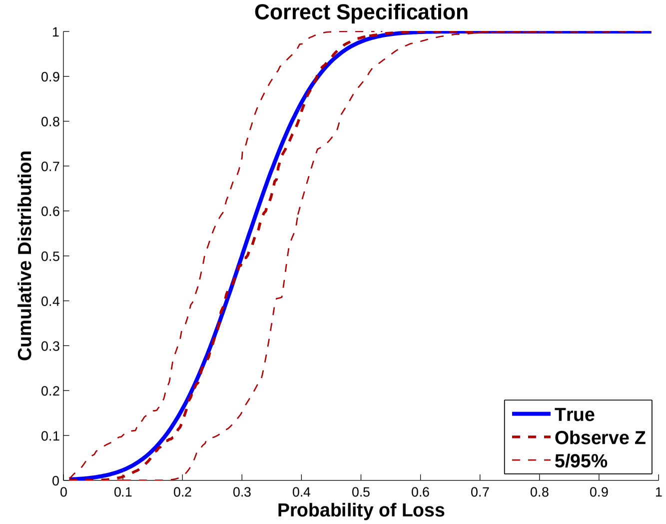


Figure B2: Monte Carlo Simulation

Skewed Specification

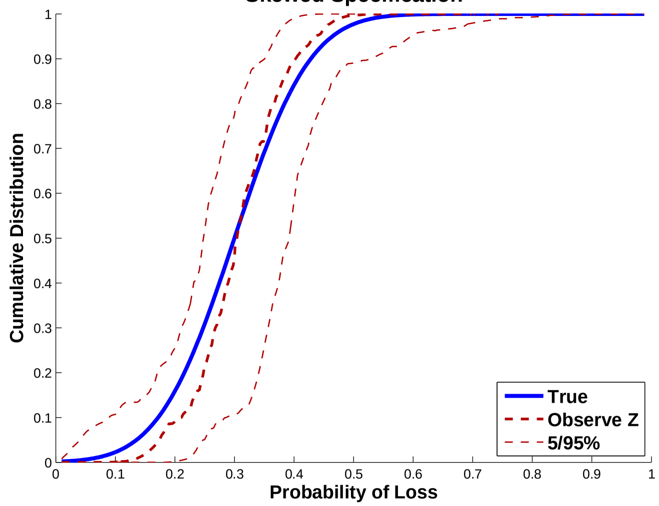


Figure B3: Monte Carlo Simulation

Excess Kurtosis Specification