ECONOMETRICA

JOURNAL OF THE ECONOMETRIC SOCIETY

An International Society for the Advancement of Economic Theory in its Relation to Statistics and Mathematics

http://www.econometricsociety.org/

Econometrica, Vol. 81, No. 5 (September, 2013), 1713–1762

PRIVATE INFORMATION AND INSURANCE REJECTIONS

NATHANIEL HENDREN
Harvard University, Cambridge, MA 02138, U.S.A. and NBER

Econometrica, Vol. 81, No. 5 (September, 2013), 1713–1762

PRIVATE INFORMATION AND INSURANCE REJECTIONS

BY NATHANIEL HENDREN[1]

Across a wide set of nongroup insurance markets, applicants are rejected based on observable, often high-risk, characteristics. This paper argues that private information, held by the potential applicant pool, explains rejections. I formulate this argument by developing and testing a model in which agents may have private information about their risk. I first derive a new no-trade result that theoretically explains how private information could cause rejections. I then develop a new empirical methodology to test whether this no-trade condition can explain rejections. The methodology uses subjective probability elicitations as noisy measures of agents' beliefs. I apply this approach to three nongroup markets: long-term care, disability, and life insurance. Consistent with the predictions of the theory, in all three settings I find significant amounts of private information held by those who would be rejected; I find generally more private information for those who would be rejected relative to those who can purchase insurance, and I show it is enough private information to explain a complete absence of trade for those who would be rejected. The results suggest that private information prevents the existence of large segments of these three major insurance markets.

KEYWORDS: Private information, adverse selection, insurance.

1. INTRODUCTION

NOT EVERYONE CAN PURCHASE INSURANCE. Across a wide set of nongroup insurance markets, companies choose not to sell insurance to potential customers who have certain observable, often high-risk, characteristics. In the nongroup health insurance market, one in seven applications to the four largest insurance companies in the United States was rejected between 2007 and 2009, a figure that excludes those people who would be rejected but were deterred from even applying.[2] In U.S. long-term care insurance, 12–23% of 65 year olds

[1]An earlier version of this paper is contained in the first chapter of my MIT graduate thesis. I am very grateful to Daron Acemoglu, Amy Finkelstein, Jon Gruber, and Rob Townsend for their guidance and support in writing this paper. I also thank Victor Chernozhukov, Sarah Miller, Whitney Newey, Ivan Werning, two anonymous referees, an extensive list of MIT graduate students, and seminar participants at University of California–Berkeley, Chicago Booth, University of Chicago, Columbia, Harvard, Microsoft Research New England, Northwestern, University of Pennsylvania, Princeton, and Stanford for helpful comments and suggestions. I would also like to thank several anonymous insurance underwriters for helpful assistance. Financial support from an NSF Graduate Research Fellowship and the NBER Health and Aging Fellowship, under the National Institute of Aging Grant T32-AG000186 is gratefully acknowledged.

[2]Figures were 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. The congressional report was released on October 12, 2010. The one in seven figure does not subtract duplicate applications of people who applied to more than one of these four firms.

 DOI: 10.3982/ECTA10931

have health conditions that would preclude them from being able to purchase insurance (Murtaugh, Kemper, and Spillman (1995)).[3]

It is surprising that a company would choose not to offer its products to a certain subpopulation. Although the rejected applicants generally have higher expected expenditures, they still face unrealized risk.[4] Regulation does not generally prevent risk-adjusted pricing in these markets.[5] So why not simply offer them a higher price?

In this paper, I argue that private information, held by the potential applicant pool, explains rejections. In particular, I provide empirical evidence in three insurance market settings that people who have observable conditions that prevent them from being able to purchase insurance also have additional knowledge about their risk beyond what is captured by their observable characteristics. To develop some intuition for this finding, consider the risk of going to a nursing home, one of the three settings that will be studied in this paper. Someone who has had a stroke, which renders them ineligible to purchase long-term care (LTC) insurance, may know not only her personal medical information (which is largely observable to an insurer), but also many specific factors and preferences that are derivatives of her health condition and affect her likelihood of entering a nursing home. These could be whether her kids will take care of her in her condition, her willingness to engage in physical therapy or other treatments that would prevent nursing home entry, or her desire to live independently with the condition as opposed to seek the aid of a nursing home. Such factors and preferences affect the cost of insuring nursing home expenses, but are often difficult for an insurance company to obtain and verify. This paper argues that, because of the private information held by applicants who have rejection conditions, if an insurer were to offer contracts to these individuals, they would be so heavily adversely selected that it would not deliver positive profits, at any price.

To make this argument formally, I begin with a theory of how private information could lead to rejections. The setting is the familiar binary loss environment introduced by Rothschild and Stiglitz (1976), generalized to incorporate an arbitrary distribution of privately informed types. In this environment, I ask under what conditions can anyone obtain insurance against the loss. I derive new a "no-trade" condition that characterizes when insurance companies would be unwilling to sell insurance on terms that anyone would accept. This condition has an unraveling intuition similar to that introduced in Akerlof (1970). The market unravels when the willingness to pay for a small

[3]Appendix F presents the rejection conditions from Genworth Financial (one of the largest U.S. LTC insurers), gathered from their underwriting guidelines that are provided to insurance agents for use in screening applicants.

[4]For example, in long-term care, I show that those who would be rejected have an average 5-year nursing home entry rate of less than 25%.

[5]The Civil Rights Act is a singular exception as it prevents purely race-based pricing.

amount of insurance is less than the pooled cost of providing this insurance to people of equal or higher risk. 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 make them unprofitable. Thus, the theory explains rejections as market segments (segmented by observable characteristics) in which the no-trade condition holds.

I then use the no-trade condition to identify properties of type distributions that are more likely to lead to no trade. This provides a vocabulary for quantifying private information. In particular, I characterize the barrier to trade imposed by a distribution of types in terms of the implicit tax rate, or markup, individuals would have to be willing to pay on insurance premiums in order for the market to exist. The comparative statics of the theory suggests that the implicit tax rates should be higher for the rejectees relative to nonrejectees and high enough for the rejectees to explain an absence of trade for plausible values of the willingness to pay for insurance.

I then develop a new empirical methodology to test the predictions of theory. I use information contained in subjective probability elicitations[6] to infer properties of the distribution of private information. I do not assume individuals can necessarily report their true beliefs. Rather, I use information in the joint distribution of elicitations and the realized events that correspond to these elicitations to deal with potential errors in elicitations.

I proceed with two complementary empirical approaches. First, I estimate the explanatory power of the subjective probabilities on the subsequent realized event, conditional on public information. I show that measures of their predictive power provide nonparametric lower bounds on theoretical metrics of the magnitude of private information. In particular, whether the elicitations are predictive at all provides a simple test for the presence of private information. I also provide a test in the spirit of the comparative statics of the theory that asks whether people who would be rejected are better able to predict their realized loss.

Second, I estimate the distribution of beliefs by parameterizing the distribution of elicitations given true beliefs (i.e., the distribution of measurement error). I then quantify the implicit tax individuals would need to be willing to pay in order for an insurance company to be able to profitably sell insurance against the corresponding loss. I then ask whether it 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.

I apply this approach to three nongroup markets: long-term care (LTC), disability, and life insurance. I combine two sources of data. First, I use data from

[6]A subjective probability elicitation about a given event is a question: "What is the chance (0–100%) that [event] will occur?"

the Health and Retirement Study, which elicits subjective probabilities that correspond to losses insured in each of these three settings, and that contains a rich set of observable demographic and health information. Second, I construct and merge a classification of those applicants who would be rejected (henceforth rejectees[7]) in each market from a detailed review of underwriting guidelines from major insurance companies.

Across all three market settings and a wide set of specifications, I find significant amounts of private information held by the rejectees: the subjective probabilities are predictive of the realized loss conditional on observable characteristics. Moreover, I find that they are more predictive for the rejectees than for the nonrejectees; indeed, once I control for observable characteristics used by insurance companies to price insurance, I cannot reject the null hypothesis of no private information where the market exists in any of the three markets I consider. Quantifying the amount of private information in each market, I estimate rejectees would need to be willing to pay an implicit tax of 82% in LTC, 42% in life, and 66% in disability insurance in order for a market to exist. In contrast, I estimate smaller implicit taxes for the nonrejectees that are not statistically different from zero in any of the three market settings.

The general empirical finding from the three settings is that there is one way to be healthy, but many (unobservable) ways to be sick. This may help explain patterns of rejections in other insurance markets. In nongroup health insurance, this can explain why those who have preexisting health conditions are rejected. In annuity markets, this can explain the *absence of* rejections. Very few individuals are informed about having exceptionally low mortality risk (there's only one way to be healthy). Thus, the population of healthy individuals can obtain annuities without a significant number of even lower mortality risks adversely selecting their contract.

This paper is related to several distinct literatures. In the theoretical dimension, it is, to my knowledge, the first paper to show that private information can eliminate all gains to trade in an 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 tradable contracts, which is unappealing in the context of insurance since insurers generally offer a menu of premiums and deductibles. Consequently, this 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 tradable contracts, this literature generally restricts the distribution of types (e.g., "two types" or a bounded support) and generally argues that trade

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

will always occur (Riley (1979), Chade and Schlee (2012)). But by considering an arbitrary distribution of types, I show this not to be the case. Indeed, not only is no trade theoretically possible; I argue it is the outcome in significant segments of three major insurance markets.

Empirically, this 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, Jullien, Salanié, and Salanié (2006), Finkelstein and Poterba (2002, 2004); see Einav, Finkelstein, and Levin (2010) 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. While this approach can potentially identify private information among those served by the market, my approach can study private information for the entire population, including rejectees. Thus, my results provide a new explanation for why previous studies have not found evidence of significant adverse selection in life insurance (Cawley and Philipson (1999)) or LTC insurance (Finkelstein and McGarry (2006)).[8] The most salient impact of private information may not be the adverse selection of existing contracts, but rather the existence of the insurance market.

Finally, this 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, this paper presents, to the best of my knowledge, the first direct empirical evidence that private information leads 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 the data. Section 6 presents the empirical specification and results for the nonparametric lower bounds. Section 7 presents the empirical specification and results for the estimation of the implicit tax imposed by private information. Section 8 places the results in the context of existing literature and discusses directions for future work. Section 9 concludes. To keep the main text to a reasonable length, the theoretical proofs and empirical estimation details are deferred to the Supplemental Material (Hendren (2013)).

[8]Although Finkelstein and McGarry (2006) found no evidence of a positive correlation between insurance purchase and claims in LTC insurance, they did find evidence of private information about nursing home entry using the same subjective probabilities I use in this paper. They subsequently argued that negatively correlated preference heterogeneity must be preventing adverse selection. However, I show that the predictive content of the elicitations is held solely by those unable to purchase insurance because of rejections. Hence my results suggests the rejection practices of LTC insurers prevents adverse selection.

2. THEORY

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

2.1. *Environment*

There exists a unit mass of agents endowed with nonstochastic wealth $w > 0$. All agents face a potential loss of size $l > 0$ that occurs with privately known probability p, which is distributed with cumulative distribution function (c.d.f.) $F(p|X)$ in the population, where X is the observable information insurers could use to price insurance (e.g., age, gender, observable health conditions, etc.). For the theoretical section, it will suffice to condition on a particular value for the observable characteristics, $X = x$, and let $F(p) = F(p|X = x)$ denote the distribution of types conditional on this value. I impose no restrictions on $F(p)$; it may be a continuous, discrete, or mixed distribution, and have full or partial support, denoted by $\Psi \subset [0, 1]$.[9] Throughout the paper, an uppercase P denotes the random variable that represents a random draw from the population (with c.d.f. $F(p)$); a lowercase p denotes a specific agent's probability (i.e., their realization of P).

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

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

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

2.2. *Implementable Allocations*

Under what conditions can anyone obtain insurance against the occurrence of loss? To ask this question in a general manner, I consider the set of implementable allocations.

DEFINITION 1: An allocation $A = \{c_{\mathrm{L}}(p), c_{\mathrm{NL}}(p)\}_{p\in\Psi}$ is *implementable* if the following statements hold.

[9] By choosing particular distributions $F(p)$, the environment nests type spaces used in many previous models of insurance. For example, $\Psi = \{p_{\mathrm{L}}, p_H\}$ 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 (2012) provided 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 $\Psi \subset [a, b] \subset (0, 1)$.

1. Allocation A is resource feasible:

$$\int \left[w - pl - pc_{\mathrm{L}}(p) - (1-p)c_{\mathrm{NL}}(p)\right] dF(p) \geq 0.$$

2. Allocation A is incentive compatible:

$$pu\bigl(c_{\mathrm{L}}(p)\bigr) + (1-p)u\bigl(c_{\mathrm{NL}}(p)\bigr) \geq pu\bigl(c_{\mathrm{L}}(\tilde{p})\bigr) + (1-p)u\bigl(c_{\mathrm{NL}}(\tilde{p})\bigr) \quad \forall p, \tilde{p} \in \Psi.$$

3. Allocation A is individually rational:

$$pu\bigl(c_{\mathrm{L}}(p)\bigr) + (1-p)u\bigl(c_{\mathrm{NL}}(p)\bigr) \geq pu(w-l) + (1-p)u(w) \quad \forall p \in \Psi.$$

It is easy to verify that these constraints must be satisfied in most, if not all, institutional environments such as competition or monopoly. Therefore, to ask when agents can obtain any insurance, it suffices to ask when the endowment, $\{(w-l, w)\}_{p\in\Psi}$, is the only implementable allocation.[10]

2.3. *The No-Trade Condition*

The key friction in this environment is that if a type p prefers an insurance contract relative to her endowment, then the pool of risks $P \geq p$ will also prefer this insurance contract relative to their endowment. Theorem 1 says that unless some type is willing to pay this pooled cost of worse risks so as to obtain some insurance, there can be no trade. Any insurance contract, or menu of insurance contracts, would be so adversely selected that it would not yield a positive profit.

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

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

where $\Psi \setminus \{1\}$ *denotes the support of* P, *excluding the point* $p = 1$. *Conversely, if* (1) *does not hold, then there exists an implementable allocation that strictly satisfies resource feasibility and individual rationality for a positive mass of types*.

[10]Focusing on implementable allocations, as opposed to explicitly modeling the market structure, also circumvents problems that arise from the potential nonexistence of competitive Nash equilibriums, as highlighted in Rothschild and Stiglitz (1976).

See Appendix A.1 for the proof.[11]

The left-hand side of (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 transfer of consumption to the event of a loss from the event of no loss. The actuarially fair cost of this transfer to a type p agent is $\frac{p}{1-p}$. However, if the worse risks $P \geq p$ also select this contract, the cost of this transfer would be $\frac{E[P|P\geq p]}{1-E[P|P\geq p]}$, which is the right-hand side of (1). The theorem shows that if no agent is willing to pay this pooled cost of worse risks, the endowment is the only implementable allocation.

Conversely, if (1) does not hold, there exists an implementable allocation that 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 selling insurance.[12] In this sense, the no-trade condition (1) characterizes when one would expect trade to occur.[13]

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, where the cost curve is a function of those who demand it. My model is different in the following sense: while Akerlof (1970) derived conditions under which a *given contract* would unravel and result in no trade, my model provides conditions under which *any contract or menu of contracts* would unravel.[14]

This distinction is important since previous literature has argued that trade must always occur in similar environments with no restrictions on the contract space so that firms can offer varying premium and deductible menus (Riley (1979), Chade and Schlee (2012)). The key difference in my environment is that I do not assume that types are bounded away from $p = 1$.[15] To see why this

[11]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 consumption bundles.

[12]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 on request).

[13]It is easily verified that the no-trade condition can hold for common distributions. For example, if $F(p)$ is uniform on [0, 1], then $E[P|P \geq p] = \frac{1+p}{2}$ so that the no-trade condition reduces to $\frac{u'(w-l)}{u'(w)} \leq 2$. Unless individuals are willing to pay a 100% tax for insurance, there can be no trade when $F(p)$ is uniform over [0, 1].

[14]This is also a difference between my approach and the literature on extreme adverse selection in finance contexts that exogenously restricts the set of tradable assets. Mailath and Noldeke (2008) provided a condition, with similar intuition to the unraveling condition in Akerlof (1970), under which a given asset cannot trade in any nonzero quantity. However, it is easy to verify in their environment that derivatives of the asset could always be traded, even when their no-trade condition holds. In contrast, by focusing on the set of implementable allocations, my approach rules out the nonzero trading of any asset derived from the loss.

[15]Both Riley (1979) and Chade and Schlee (2012) assumed $\sup \Psi < 1$.

matters, recall that the key friction that can generate no trade is the unwillingness of any type to pay the pooled cost of worse risks. This naturally requires the perpetual existence of worse risks. Otherwise, the highest risk type, say $\bar{p} = \sup \Psi$, would be able to obtain an actuarially fair full insurance allocation, $c_{\mathrm{L}}(\bar{p}) = c_{\mathrm{NL}}(\bar{p}) = w - \bar{p}l$, which would not violate the incentive constraints of any other type. Therefore, the no-trade condition requires some risks be arbitrarily close to $p = 1$.

COROLLARY 1: *Suppose condition* (1) *holds. Then* $F(p) < 1\ \forall p < 1$.

Corollary 1 highlights why previous theoretical papers have not found outcomes of no trade in the binary loss environment with no restrictions on the contract space; they assumed $\sup \Psi < 1$.

The presence of risks near $p = 1$ makes the provision of insurance more difficult because it increases the values of $E[P|P \geq p]$ at interior values of p. However, the need for P to have full support near 1 is not a very robust requirement for no trade. In reality, the cost of setting up a contract is nonzero, so insurance companies cannot offer an infinite set of contracts. Remark 1 shows that if each allocation other than the endowment must attract a nontrivial fraction of types, then risks arbitrarily close to 1 are not required for no trade.

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

$$\mu\big(\big\{p \big| \big(c_{\mathrm{L}}(p), c_{\mathrm{NL}}(p)\big) = \big(c_{\mathrm{L}}(q), c_{\mathrm{NL}}(q)\big)\big\}\big) \geq \alpha,$$

where μ is the measure defined by $F(p)$. Then the endowment is the only implementable allocation if and only if

$$\frac{p}{1-p}\frac{u'(w-l)}{u'(w)} \leq \frac{E[P|P \geq p]}{1 - E[P|P \geq p]} \quad \forall p \in \hat{\Psi}_{1-\alpha}, \tag{2}$$

where $\hat{\Psi}_{1-\alpha} = [0, F^{-1}(1-\alpha)] \cap (\Psi \setminus \{1\})$.[16] Therefore, the no-trade condition needs to hold only for values $p < F^{-1}(1-\alpha)$.

For any $\alpha > 0$, it is easy to verify that the no-trade condition not only does not require types near $p = 1$, but it actually imposes no constraints on the upper range of the support of P.[17] In this sense, the requirement of risks arbitrarily close to $p = 1$ is a theoretical requirement in a world with no other

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

[17] More precisely, for any $\alpha > 0$ and $\gamma \in (0, 1]$, there exist $u(\cdot)$ and $F(p)$ such that $F(\gamma) = 1$ and the no-trade condition in (2) holds.

frictions, but not an empirically relevant condition if one believes insurance companies cannot offer contracts that attract an infinitesimal fraction of the population. Going forward, I retain the benchmark assumption of no such frictions or transactions costs, but return to this discussion in the empirical work in Section 7.

In sum, the no-trade condition (1) provides a theory of rejections: individuals who have observable characteristics, X, such that the no-trade condition (1) holds are rejected; individuals who have observable characteristics, X, such that (1) does not hold are able to purchase insurance. This is the theory of rejections that the remainder of this paper will seek to test.

3. COMPARATIVE STATICS AND TESTABLE PREDICTIONS

So as to generate testable implications of this theory of rejections, this section derives properties of distributions, $F(p)$, that are more likely to lead to no trade. I provide two such metrics that are used in the subsequent empirical analysis.

3.1. *Two Measures of Private Information*

To begin, multiply the no-trade condition (1) by $\frac{1-p}{p}$, yielding

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

The left-hand side is the ratio of marginal utilities in the loss versus no loss state, evaluated at the endowment. The right-hand side is independent of the utility function, u, and is the markup that would be imposed on type p if she had to cover the cost of worse risks, $P \geq p$. I define this term the pooled price ratio.

DEFINITION 2: For any $p \in \Psi \setminus \{1\}$, the *pooled price ratio at* p is given by

$$T(p) = \frac{E[P|P \geq p]}{1-E[P|P \geq p]} \frac{1-p}{p}. \tag{3}$$

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*

$$\frac{u'(w-l)}{u'(w)} \leq \inf_{p \in \Psi \setminus \{1\}} T(p). \tag{4}$$

Whether or not there can 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, $\inf_{p\in\Psi\setminus\{1\}} T(p)$. I call $\inf_{p\in\Psi\setminus\{1\}} T(p)$ the *minimum pooled price ratio*.

The minimum pooled price ratio 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. Thus, $\inf_{p\in\Psi\setminus\{1\}} T(p) - 1$ is the implicit tax rate imposed by private information. Given any distribution of risks, $F(p)$, it quantifies the implicit tax individuals would need to be willing to pay so that a market could exist.

Equation (4) leads to a simple comparative static.

COROLLARY 3—Comparative Static in the Minimum Pooled Price Ratio: *Consider two market segments*, 1 *and* 2, *with pooled price ratios $T_1(p)$ and $T_2(p)$, and common von Neumann–Morgenstern* (*vNM*) *preferences u. Suppose*

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

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, it will also be helpful to have another metric to guide portions of the empirical analysis.

DEFINITION 3: For any $p \in \Psi$, define the *magnitude of private information at* p by

$$m(p) = E[P|P \geq p] - p. \tag{5}$$

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*, 1 *and* 2, *with magnitudes of private information $m_1(p)$ and $m_2(p)$, and common support Ψ and common vNM preferences u. Suppose*

$$m_1(p) \leq m_2(p) \quad \forall p \in \Psi.$$

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 $p \in \Psi$; in this sense, Corollary 4 is a less precise comparative static than Corollary 3.

3.2. *Testable Hypotheses*

The goal of the rest of the paper is to test whether the no-trade condition (1) can explain rejections by estimating properties of the distribution of private information, $F(p|X)$, for rejectees and nonrejectees. Assuming, for the moment, that $F(p|X)$ is observable to the econometrician, the ideal tests are as follows. First, do rejectees have private information (i.e., is $F(p|X)$ a nontrivial distribution for the rejectees)? Second, do rejectees have more private information than the nonrejectees, as suggested by the comparative statics in Corollaries 3 and 4? Finally, is the quantity of private information, as measured by the minimum pooled price ratio, large (small) enough to explain (the absence of) rejections for plausible values of agents' willingness to pay, $\frac{u'(w-l)}{u'(w)}$, as suggested by Corollary 2?

Note that these tests do not involve any observation of adverse selection (i.e., a correlation between insurance purchases and realized losses). Instead, these ideal tests simulate the extent to which private information would afflict a hypothetical insurance market that pays $1 in the event that a loss occurs and prices policies using the observable characteristics, X.

To implement these tests, one must estimate properties of the distribution of private information, $F(p|X)$, to which I now turn.

4. EMPIRICAL METHODOLOGY

I develop an empirical methodology to study private information and operationalize the tests in Section 3.2. I rely primarily on four pieces of data. First, let L denote an event (e.g., dying in the next 10 years) that is commonly insured in some insurance market (e.g., life insurance).[18] Second, let Z denote an individual's subjective probability elicitation about event L (i.e., Z is a response to the question, "What is the chance (0–100%) that L will occur?"). Third, let X continue to denote the set of public information insurance companies would use to price insurance against the event L. Finally, let Θ^{Reject} and Θ^{NoReject} partition the space of values of X into those for whom an insurance company does and does not offer insurance contracts that provide payment if L occurs (e.g., if L is the event of dying in the next 10 years, Θ^{Reject} would be the value of observables, X, that render someone ineligible to purchase life insurance).

[18]To condense notation, L will denote both a probabilistic event and the binary random variable equal to 1 if the event occurs and 0 if the event does not occur (i.e., $\Pr\{L\} = \Pr\{L = 1\} = E[L]$).

The premise underlying this approach is that the elicitations, Z, are nonverifiable to an insurance company. Therefore, they can be excluded from the set of public information insurance companies would use to price insurance, X, and used to infer properties of the distribution of private information.

I maintain the implicit assumption in Section 2 that individuals behave as if they have true beliefs, P, about the occurrence of the loss, L.[19] But there are many reasons to expect individuals not to report exactly these beliefs on surveys.[20] Therefore, I do not assume $Z = P$. Instead, I use information contained in the joint distribution of Z and L (that are observed) to infer properties about the distribution of P (that is not directly observed).

I conduct two complementary empirical approaches. Under relatively weak assumptions rooted in economic rationality, I provide a test for the presence of private information and a nonparametric lower bound on the average magnitude of private information, $E[m(P)]$. Loosely, this approach asks how predictive are the elicitations of the loss L, conditional on observable information, X. Second, I use slightly stronger structural assumptions to estimate the distribution of beliefs, $F(p|X)$, and the minimum pooled price ratio. I then test whether it is larger for the rejectees and large (small) enough to explain a complete absence of trade for plausible values of $\frac{u'(w-l)}{u'(w)}$, as suggested by Corollary 2.

In this section, I introduce these empirical approaches in the abstract. I defer a discussion of the empirical specification and statistical inference in my particular settings to Sections 6 and 7, after discussing the data and settings in Section 5.

4.1. *Nonparametric Lower Bounds*

Instead of assuming people necessarily report their true beliefs, I begin with the weaker assumption that people cannot report *more* information than they know.

[19]The approach therefore follows the view of personal probability expressed in the seminal work of Savage (1954). The existence of beliefs P is guaranteed as long as people behave consistently (in the sense of Savage's axioms) in response to gambles over L.

[20]For example, they may not have the training to know how to answer probabilistic questions, they may intentionally lie to the surveyor, or they may simply be lazy in thinking about their response. Indeed, existing research suggests the way in which the elicitation is conducted affects the reported belief elicitation (Gigerenzer and Hoffrage (1995), Miller, Kirlik, Tsai, and Kosorukoff (2008)), which suggests elicitations do not measure true beliefs exactly. Previous literature has also argued that the elicitations in my settings should not be viewed as true beliefs due to excess concentrations at 0, 50%, and 100% (Gan, Hurd, and McFadden (2005), Hurd (2009)).

ASUMPTION 1: *The elicitation Z contains no additional information than P about the loss L, so that*

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

This assumption states that if the econometrician were trying to forecast whether or not an agents' loss would occur, and knew both the observable characteristics X and the agents' true beliefs, P, the econometrician could not improve the forecast of L by also knowing the elicitation Z. All of the predictive power that Z has about L must come from agents' beliefs, P.[21] Proposition 1 follows.

PROPOSITION 1: *Suppose* $\Pr\{L|X, Z\} \neq \Pr\{L|X\}$ *for a positive mass of realizations of Z. Then* $\Pr\{L|X, P\} \neq \Pr\{L|X\}$ *for a positive mass of realizations of P.*

PROOF: Assumption 1 implies $E[\Pr\{L|X, P\}|X, Z] = \Pr\{L|X, Z\}$. *Q.E.D.*

Proposition 1 says that if Z has predictive information about L conditional on X, then agents' true beliefs P have predictive information about L conditional on X, that is, agents have private information. This motivates my test for the presence of private information.

TEST 1—Presence of Private Information: Are the elicitations, Z, predictive of the loss, L, conditional on observable information, X?

Although this test establishes the presence of private information, it does not provide a method of asking whether one group has more private information than another. Intuitively, the predictiveness of Z should be informative of how much private information people have. Such a relationship can be established with an additional assumption about how realizations of L relate to beliefs P.

ASUMPTION 2: *Beliefs P are unbiased*: $\Pr\{L|X, P\} = P$.

Assumption 2 states that if the econometrician could hypothetically identify an individual who has beliefs P, then the probability that the loss occurs equals P. As an empirical assumption, it is strong, but commonly made in existing literature (e.g., Einav, Finkelstein, and Schrimpf (2010)); indeed, it provides perhaps the simplest link between the realized loss L and beliefs P.[22]

[21]This assumption would be clearly implied in a model in which agents formed rational expectations from an information set that included X and Z. In this case, $\Pr\{L|X, P, Z\} = P$. But it also allows agents' beliefs to be biased, so that $\Pr\{L|X, P, Z\} = h(P)$, where h is any function that is not dependent on Z. In particular, $h(P)$ could be an S-shaped function as suggested by Kahneman and Tversky (1979).

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

Under Assumptions 1 and 2, the predictiveness of the elicitations forms a distributional lower bound on the distribution of P. To see this, define P_Z to be the predicted value of L given the variables X and Z,

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

Under Assumptions 1 and 2, it is easy to verify (see Appendix B) that

$$P_Z = E[P|X, Z],$$

so that the true beliefs P are a mean-preserving spread of the distribution of predicted values P_Z. In this sense, the true beliefs are more predictive of the realized loss than are the elicitations. In particular, if P_Z is dispersed for the rejectees, then P must be even more dispersed for the rejectees.

This also motivates my first test of whether rejectees have more private information than nonrejectees. I plot the distribution of predicted values, P_Z, separately for rejectees ($X \in \Theta^{\text{Reject}}$) and nonrejectees ($X \in \Theta^{\text{NoReject}}$). I then assess whether it is more dispersed for the rejectees.

In addition to a visual inspection of P_Z, one can also construct a dispersion metric derived from the comparative statics of the theory. Recall from Corollary 4 that higher values of the magnitude of private information, $m(p)$, are more likely to lead to no trade. Consider the average magnitude of private information, $E[m(P)|X]$. This is a nonnegative measure of the dispersion of the population distribution of P. If an individual were drawn at random from the population, one would expect the risks higher than his to have an average loss probability that is $E[m(P)|X]$ higher.

Although P is not observed, I construct the analogue using the P_Z distribution. First, I construct $m_Z(p)$ as the difference between p and the average predicted probability, P_Z, of those who have predicted probabilities higher than p:

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

The $Z|X$ subscript highlights that I am integrating over realizations of Z conditional on X.

Now, I construct the average magnitude of private information implied by Z in segment X, $E[m_Z(P_Z)|X]$. This is the average difference in segment X between an individual's predicted loss and the predicted losses of those who have higher predicted probabilities. Proposition 2 follows from Assumptions 1 and 2.

PROPOSITION 2—Lower Bound: *We have* $E[m_Z(P_Z)|X] \leq E[m(P)|X]$.

See Appendix B for the proof.

Proposition 2 states that the average magnitude of private information implied by Z is a lower bound on the true average magnitude of private information. Therefore, using only Assumptions 1 and 2, one can provide a lower bound to the answer to the question, “If an individual is drawn at random, on average how much worse are the higher risks?”

Given this theoretical measure of dispersion, $E[m_Z(P_Z)|X]$, I conduct a test in the spirit of the comparative statics given by Corollary 4. I test whether rejectees have higher values of $E[m_Z(P_Z)|X]$:

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

Stated loosely, (6) asks whether the subjective probabilities of the rejectees better explain the realized losses than do those of the nonrejectees, where “better explain” is measured using the dispersion metric $E[m_Z(P_Z)|X]$.[23] I now summarize the tests for more private information for the rejectees relative to the nonrejectees.

TEST 2—More Private Information for Rejectees: Are the elicitations, Z, more predictive of L for the rejectees: (a) is P_Z more dispersed for rejectees and (b) is $\Delta_Z > 0$?

Discussion

In sum, I conduct two sets of tests motivated by Assumptions 1 and 2. First, I ask whether the elicitations are predictive of the realized loss conditional on X (Test 1); this provides a test for the presence of private information as long as people cannot unknowingly predict their future loss (Assumption 1). Second, I ask whether the elicitations are more predictive for rejectees relative to nonrejectees (Test 2). To do so, I analyze whether the predicted values, P_Z, are more dispersed for rejectees relative to nonrejectees. In addition to assessing this visually, I collapse these predicted values into the average magnitude of private information implied by Z, $E[m_Z(P_Z)]$, and ask whether it is larger for those who would be rejected relative to those who can purchase insurance ((6)).

The approach is nonparametric in the sense that I have made no restrictions on how the elicitations Z relate to the true beliefs P. For example, P_Z and $m_Z(p)$ are invariant to one-to-one transformations in Z: $P_Z = P_{h(Z)}$ and $m_Z(p) = m_{h(Z)}(p)$ for any one-to-one function h. Thus, I 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 P_Z or $E[m_Z(P_Z)|X]$.[24]

[23] Note that the expectations in (6) condition on X and then aggregate across values of X in a given sample (either Θ^{Reject} or Θ^{NoReject}). Hence, the average magnitudes of private information implied by Z provide an aggregated measure of the explanatory power of Z for L conditional on X.

[24] In principle, Z need not even be a number. Some individuals could respond to the elicitation question in a crazy manner by saying they like red cars and others that they like Buffy the Vampire

But while the lower bound approach relies on only minimal assumptions on how subjective probabilities relate to true beliefs, the resulting empirical test in (6) suffers several significant limitations as a test of the theory that private information causes insurance rejections. First, comparisons of lower bounds of $E[m(P)|X]$ across segments do not necessarily imply comparisons of its true magnitude. Second, orderings of $E[m(P)|X]$ do not imply orderings of $m(p)$ for all p, which was the statement of the comparative static in $m(p)$ in Corollary 4. Finally, in addition to having limitations as a test of the comparative statics, this approach cannot quantify the minimum pooled price ratio. These shortcomings motivate a complementary empirical approach, which imposes structure on the relationship between Z and P and estimates of the distribution of private information, $F(p|X)$.

4.2. *Estimation of the Distribution of Private Information*

The second approach estimates the distribution of private information and the minimum pooled price ratio. For expositional ease, fix an observable $X = x$ and let $f_P(p)$ denote the probability density function (p.d.f.) of the distribution of beliefs, P, given $X = x$, which is assumed to be continuous. For this approach, I expand the joint p.d.f./p.m.f. (probability mass function) of the observed variables L and Z, denoted $f_{L,Z}(L, Z)$, by integrating over the unobserved beliefs, P,

$$\begin{aligned} f_{L,Z}(L, Z) &= \int_0^1 f_{L,Z}(L, Z|P = p) f_P(p)\, dp \\ &= \int_0^1 \bigl(\Pr\{L|Z, P = p\}\bigr)^L \bigl(1 - \Pr\{L|Z, P = p\}\bigr)^{1-L} \\ &\quad \times f_{Z|P}(Z|P = p) f_P(p)\, dp \\ &= \int_0^1 p^L (1 - p)^{1-L} f_{Z|P}(Z|P = p) f_P(p)\, dp, \end{aligned}$$

where $f_{Z|P}(Z|P = p)$ is the distribution of elicitations given beliefs. 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. The third equality follows from Assumptions 1 and 2.

The goal of this approach is to specify a functional form for $f_{Z|P}$, say $f_{Z|P}(Z|P; \theta)$, and a flexible approximation for f_P, say $f_P(p; \nu)$, and estimate θ and ν using maximum likelihood from the observed data on L and Z. To do

Slayer. The empirical approach would proceed to analyze whether a stated liking of red cars versus Buffy the Vampire Slayer is predictive of L conditional on X. Of course, such elicited information may have low power to identify private information about L.

so, one must impose sufficient restrictions on $f_{Z|P}$ so that θ and ν are identified. Because the discussion of functional form for $f_{Z|P}$ and its identification is more straightforward after discussing the data, I defer a detailed discussion of my choice of specification and the details of identification to Section 7.1. At a high level, identification of the elicitation error parameters, θ, comes from the relationship between L and Z, and identification of the distribution of P is a deconvolution of the distribution of Z, where θ contains the parameters that govern the deconvolution. Therefore, a key concern for identification is that the measurement error parameters are well identified from the relationship between Z and L; I discuss how this is the case in my particular specification in Section 7.1.[25]

With an estimate of f_P, the pooled price ratio follows from the identity $T(p) = \frac{E[P|P\geq p]}{1-E[P|P\geq p]}\frac{1-p}{p}$. I then construct an estimate of its minimum, $\inf_{p\in[0,1)} T(p)$. Although $T(p)$ can be calculated at each p using estimates of $E[P|P \geq p]$, as p increases, $E[P|P \geq p]$ relies on a smaller and smaller effective sample size. Thus, the minimum of $T(p)$ is not well identified over a domain that includes the uppermost points of the support of P. To overcome this extreme quantile estimation problem, I construct the minimum of $T(p)$ over the restricted domain, $\hat{\Psi}_\tau = [0, F_P^{-1}(\tau)] \cap (\Psi \setminus \{1\})$. For a fixed quantile, estimates of the minimum pooled price ratio over $\hat{\Psi}_\tau$ are continuously differentiable functions of the maximum likelihood estimator (MLE) parameter estimates of $f_P(p)$ for $p \leq F_P^{-1}(\tau)$.[26] Derived MLE estimates of $\inf_{p\in\hat{\Psi}_\tau} T(p)$ are consistent and asymptotically normal, provided $F_P(p)$ is continuous.[27] One can assess the robustness to the choice of τ, but the estimates will become unstable as $\tau \to 1$.

While the motivation for restricting attention to $\hat{\Psi}_\tau$ as opposed to Ψ is primarily because of statistical limitations, Remark 1 in Section 2.3 provides an economic rationale for why $\inf_{p\in\hat{\Psi}_\tau} T(p)$ may not only be a suitable substitute for $\inf_{p\in\Psi\setminus\{1\}} T(p)$, but also may actually be more economically relevant. If contracts must attract a nontrivial fraction $1-\tau$ of the market to be viable, then $\inf_{p\in\hat{\Psi}_\tau} T(p)$ characterizes the barrier to trade imposed by private information.

Given estimates of $\inf_{p\in\hat{\Psi}_\tau} T(p)$ for rejectees and nonrejectees, I test whether it is larger (smaller) for the rejectees (Corollary 3) and whether it is large (small) enough to explain a complete absence of (presence of) trade for plausible values of people's willingness to pay, $\frac{u'(w-l)}{u'(w)}$, as suggested by Corollary 2.

[25] Indeed, not all distributions $f_{Z|P}$ are identified from data on L and Z since, in general, $f_{Z|P}$ is an arbitrary two-dimensional function, whereas L is binary.

[26] Nondifferentiability could hypothetically occur at points where the infimum is attained at distinct values of p.

[27] To see this, note that if $F_P(p)$ is continuous, then $T(p) = \frac{1-pF_P(p)-\int_0^p F_P(\hat{p})\,d\hat{p}}{1-F_P(p)}$, so that $T(p)$ is continuous in the estimated parameters of F_P.

TEST 3—Quantification of Private Information: Is the minimum pooled price ratio larger for the rejectees relative to the nonrejectees and is it large enough (small enough) to explain an absence of (presence of) trade for plausible values of agents' willingness to pay?

5. SETTING AND DATA

I ask whether private information can explain rejections in three nongroup insurance market settings: long-term care, disability, and life insurance.

5.1. *Short Background on the Three Nongroup 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: expenditures in the United States total over $135 billion in 2004. Moreover, expenditures are heavily skewed: less than half of the population will ever move to a nursing home (Congressional Budget Office (CBO) (2004)). Despite this, the LTC insurance market is small: roughly 4% of all nursing home expenses are paid by private insurance, compared to 31% paid out-of-pocket (CBO (2004)).[28]

Private disability insurance protects against lost income that results from a work-limiting disability. It is primarily sold through group settings, such as one's employer; more than 30% of nongovernment workers have group-based disability policies. In contrast, the nongroup market is quite small. Only 3% of nongovernment workers own a nongroup disability policy, most of whom are self-employed or professionals who do not have access to employer-based group policies (American Council of Life Insurers (ACLI) (2010)).[29]

Life insurance provides payments to ones' heirs or estate upon death, insuring lost income or other expenses. In contrast to the nongroup disability and LTC markets, the private nongroup life insurance market is quite big. More than half of the adult U.S. population owns life insurance, 54% of which is sold in the nongroup market.[30]

Previous Evidence of Private Information

Previous research has found minimal or no evidence of adverse selection in these three markets. In life insurance, Cawley and Philipson (1999) found no evidence of adverse selection. He (2009) revisited this topic with a different

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

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

[30]Life insurance policies either expire after a fixed length of time (term life) or cover one's entire life (whole life). Of the nongroup policies in the United States, 43% of these are term policies, while the remaining 57% are whole life policies (ACLI (2010)).

sample focusing on new purchasers and did find evidence of adverse selection under some empirical specifications. In long-term care, Finkelstein and McGarry (2006) found direct evidence of private information by showing that subjective probability elicitations are correlated with subsequent nursing home use. However, they found no evidence that this private information leads to adverse selection: conditional on the observables used to price insurance, those who buy LTC insurance are no more likely to go to a nursing home than are those who do not purchase LTC insurance.[31] To my knowledge, there is no previous study of private information in the nongroup disability market.

5.2. *Data*

To implement the empirical approach in Section 4, the ideal data set contains four pieces of information for each setting:

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

The 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 older individuals (mostly over age 55) and their spouses. It contains a rich set of health and demographic information. Moreover, it asks respondents three subjective probability elicitations about future events that correspond to a commonly insured loss in each of the three 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).[32]

Figure 1(a)–(c) displays histograms of these responses (divided by 100 to scale to $[0, 1]$).[33] These histograms highlight one reason why it would be problematic to view these elicitations as true beliefs. As noted in previous literature using these subjective probabilities (Gan, Hurd, and McFadden (2005), Finkelstein

[31]They suggest that heterogeneous preferences, in which good risks also have a higher valuation of insurance, can explain why private information does not lead to adverse selection.

[32]I construct the corresponding elicitation to be $Z^{\text{die}} = 100\% - Z^{\text{live}}$, where Z^{live} is the survey elicitation for the probability of living to AGE.

[33]The histograms use the sample selection described in Section 5.2.3. The bar heights are normalized so that their areas sum to 1 in each sample.

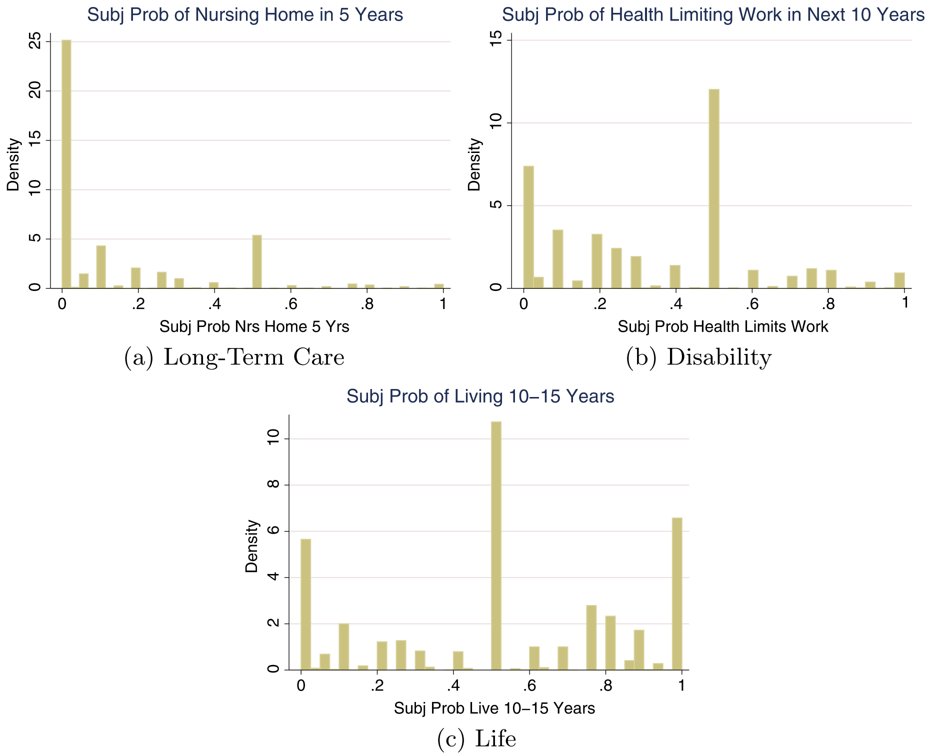


FIGURE 1.—Subjective probability histograms.

and McGarry (2006)), many respondents report 0, 50, or 100. Taken literally, responses of 0 or 100 imply an infinite degree of certainty. The lower bound approach remains agnostic on the way in which focal point responses relate to true beliefs. The parametric approach will take explicit account of this focal point response bias in the specification of $f_{Z|P}(Z|P;\theta)$, discussed further in Section 7.1.1.

Corresponding to each subjective probability elicitation, I 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.[34] In disability, L denotes the event that the respondent reports that their health limits their work activity in the subsequent 10–11 years.[35] In life, L denotes the event that

[34] Although the HRS surveys every 2 years, I use information from the third subsequent interview (6 years post), which provides date of nursing home entry information to construct the exact 5-year indicator of nursing home entry.

[35] The loss is defined as occurring when the individual reports yes to the question "Does your health limit your work activity?" over the subsequent five surveys, which is 10 years for all waves

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.[36]

5.2.1. *Public Information*

To identify private information, it is essential to control for the public information, X, that would be used by insurance companies to price contracts. For nonrejectees, this is a straightforward requirement that involves analyzing existing contracts. But for rejectees, I must make an assumption about how insurance companies would price these contracts if they were to offer them. My preferred approach is to assume insurance companies price rejectees separately from those to whom they currently offer contracts, but use a similar set of public information. Thus, the 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 that allows me to approximate the set of information that insurance companies use to price insurance. Indeed, previous literature has used the HRS to replicate the observables used by insurance companies to price insurance in LTC and life (for LTC, see Finkelstein and McGarry (2006) and for life, see He (2009)), and I primarily follow this literature in constructing this set of covariates. Appendix C.1 provides a detailed listing of the control specifications used in each market setting.

The quality of the approximation to what insurers actually use to price insurance is quite good, but does vary by market. For long-term care, I replicate the information set of the insurance company quite well. For example, perhaps the most obscure 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, I 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 I will not observe in the HRS. Conversations with underwriters in these markets suggest these tests are primarily to confirm application information, which I can approximate quite well with the HRS. But I cannot rule out the potential that

except AHEAD wave 2, which corresponds to a time interval of 11 years because of a slightly different survey spacing. Although the HRS has other measures of disability (e.g., Social Security Disability Insurance (SSDI) claims), I use this measure because the wording corresponds exactly to the subjective probability elicitation, which will be important for the structural assumptions made to estimate the minimum pooled price ratio.

[36]The HRS collects date of death information that allows me to establish the exact age of death.

there is additional information that can be gathered by insurance companies in the disability and life settings.[37]

While the preferred specification attempts to replicate the variables used by insurance companies in pricing, I also assess the robustness of the estimates to larger and smaller sets of controls.[38] As a baseline, I consider a specification with only age and gender. As an extension, I also consider an extended controls specification that adds a rich set of interactions between health conditions and demographic variables that could be, but are not currently, used in pricing insurance. I conduct the lower bound approach for all three sets of controls. For brevity, I focus exclusively on the preferred specification of pricing controls for the parametric approach.

5.2.2. *Rejection Classification*

Not everyone can purchase insurance in these three nongroup markets. To identify conditions that lead to rejection, I obtain underwriting guidelines used by underwriters and provided to insurance agents for use in screening applicants. An insurance company's underwriting guidelines list the conditions for which underwriters are instructed to not offer insurance at any price and for which insurance agents are expected to discourage applications. These guidelines are generally viewed as a public relations liability and are not publicly available.[39] Thus, the extent of my access varies by market. In long-term care, I obtained a set of guidelines used by an insurance broker for 18 of the 27 largest long-term care insurance companies, which comprise a majority of the U.S. market.[40] In disability and life, I obtained several underwriting guidelines and supplemented this information with interviews with underwriters at several major U.S. insurance companies. Appendix F provides several pages from the LTC underwriting guideline from Genworth Financial, one of the largest LTC insurers in the United States.[41]

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

[38] 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' that are uncorrelated with L or Z has no effect on the population value of $E[m(P)|X \in \Theta]$.

[39] An example of these guidelines is presented in Appendix F and a collection of these guidelines is available on my website. Also, many underwriting guidelines are available via internet searches of "underwriting guideline not-for-public-use pdf." These are generally left on the websites of insurance brokers who leave them electronically available to their sales agents and, potentially unknowingly, available to the general public.

[40] I thank Amy Finkelstein for making these broker-collected data available.

[41] A collection of underwriting guidelines from these three markets is available from the author on request and is posted on my website.

I then use the detailed health and demographic information available in the HRS to identify individuals who have these rejection conditions. While the HRS contains a relatively comprehensive picture of respondents' health, sometimes the rejection conditions are too precise to be matched to the HRS. For example, individuals who have advanced stages of lung disease would be unable to purchase life insurance, but some companies will sell policies to individuals who have a milder case of lung disease; however, the HRS only provides information for the presence of a lung disease.

Instead of attempting to match all cases, I construct a third classification in each setting, "uncertain," to which I classify those individuals who may be rejected, but for whom data limitations prevent a solid assessment. This allows me to be relatively confident in the classification of rejectees and nonrejectees. For completeness, I present the lower bound analysis for all three classifications.

Table I presents the list of conditions for the rejection and the uncertain classifications, along with the frequency of each condition in the sample (using the sample selection outlined below in Section 5.2.3). LTC insurers generally reject applicants who have conditions that would make them more likely to use a nursing home in the relatively near future. Activity of daily living (ADL) restrictions (e.g., needs assistance walking, dressing, using toilet, etc.), a previous stroke, any previous home nursing care, and anyone over the age of 80 would be rejected regardless of health status. Disability insurers reject applicants with back conditions, obesity (body mass index > 40), and doctor-diagnosed psychological conditions such as depression or bipolar disorder. Finally, life insurers reject applicants who have had a past stroke or currently have cancer.

Table I also lists the conditions that may lead to rejection dependent on the specifics of the disease. People who have these conditions are allocated into the uncertain classification.[42] In addition to health conditions, disability insurers also have stringent income and job characteristics underwriting. Individuals who earn less than $30,000 (or wages below $15/hr) and individuals who work in blue-collar occupations are often rejected regardless of health condition due to their employment characteristics. I therefore allocate all such individuals to the uncertain category in the disability insurance setting.

Given these classifications, I construct the reject, no reject, and uncertain samples by first taking anyone who has a known rejection condition in Table I and classifying them into the reject sample in each setting. I then classify anyone with an uncertain rejection condition into the uncertain classification, so that the remaining category is the set of people who can purchase insurance (the no reject classification).

[42]I also attempt to capture the presence of rarer conditions that are 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, I allocate to the uncertain classification individuals who report having an additional major health problems that was not explicitly asked about in the survey.

TABLE I

REJECTION CLASSIFICATION[a]

	Long-Term Care		Disability		Life	
Classification	Condition	% Sample	Condition	% Sample	Condition	% Sample
Rejection	Any ADL/IADL restriction	7.5%	Back condition	22.7%	Cancer[e] (current)	13.1%
	Past stroke	8.3%	Obesity (BMI > 40)	1.7%	Stroke (ever)	7.3%
	Past nursing/home care	13.6%	Psychological condition	6.3%		
	Over age 80	20.0%				
Uncertain	Lung disease	10.7%	Arthritis	36.9%	Diabetes	13.8%
	Heart condition	29.6%	Diabetes	7.7%	High blood pressure	50.7%
	Cancer (current)	15.4%	Lung disease	5.1%	Lung disease	10.9%
	Hip fracture	1.3%	High blood pressure	31.3%	Cancer (ever, not current)	12.1%
	Memory condition[b]	0.9%	Heart condition	6.9%	Heart condition	26.5%
	Other major health problems[c]	26.8%	Cancer (ever have)	4.6%	Other major health problems[c]	23.5%
			Blue-collar/high-risk job[d]	23.3%		
			Wage < $15 or income < $30K	65.5%		
			Other major health problems[c]	16.2%		

[a] Percentages will not add to the total fraction of the population classified as rejection and uncertain because of people with multiple conditions.

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

[c] 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?"

[d] I define blue collar/high-risk jobs as nonself employed jobs in the cleaning, foodservice, protection, farming, mechanics, construction, and equipment operators.

[e] Basel cell (skin) cancers are excluded from the cancer classification.

5.2.3. *Sample Selection*

For each sample, I begin with years 1993–2008 of the HRS. The selection process varies across each of the three market settings due to varying data constraints. Appendix C.2 discusses the specific data construction details for each setting. The primary sample restrictions arise from requiring that the subjective elicitation be asked (e.g., only individuals over age 65 are asked about future nursing home use) and needing to observe individuals in the panel long enough to construct the loss indicator L in each setting.[43] For LTC, the sample consists of individuals aged 65 and older; for disability, the sample consists of individuals aged 60 and under;[44] and for life, the sample consists of individuals over age 65. Table II presents the summary statistics for each sample. I include multiple observations for a given individual (which are spaced roughly 2 years apart) to increase power.[45]

There are several broad patterns across the three samples. First, there is a sizable sample of rejectees in each setting. Because the HRS primarily surveys older individuals, the sample is older, and therefore sicker, than the average insurance purchaser in each market. Obtaining this large sample size of rejectees is a primary benefit of the HRS; but it is important to keep in mind that the fraction of rejectees in the HRS is not a measure of the fraction of the applicants in each market who are 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 that has less stringent underwriting requirements than the nongroup market.

However, the fact that some own insurance raises the concern that moral hazard could generate heterogeneity in loss probabilities from differential insurance ownership. Therefore, I also perform robustness checks in LTC and life on samples that exclude those who currently own insurance.[46] Since Med-

[43]Note that death during this subsequent time horizon does not exclude an individual from the sample; I classify the event of dying before the end of the time horizon as $L = 0$ for the LTC and disability settings as long as an individual did not report the loss (i.e., nursing home entry or health limiting work) prior to death.

[44]The disability question is asked of individuals up to age 65, but I exclude individuals aged 61–65 because of the near presence of retirement. Ideally, I would focus on a sample of even younger individuals, but unfortunately the HRS contains relatively few respondents below age 55.

[45]All standard errors will be clustered at the household level. Because the multiple observations within a person will always have different X values (e.g., different ages), including multiple observations per person does not induce bias in the construction of $F(p|X)$.

[46]Since rejection conditions are generally absorbing states, this rules out the path through which insurance contract choice could generate heterogeneity for the rejectees. For the nonrejectees, this removes the heterogeneity induced by current contract choice, but it does not remove heterogeneity introduced from expected future purchase of insurance contracts. However, for my purposes this remaining moral hazard impact only biases against finding more private information among the rejectees.

TABLE II

SAMPLE SUMMARY STATISTICS

	Long-Term Care			Disability			Life		
	No Reject	Reject	Uncertain	No Reject	Reject	Uncertain	No Reject	Reject	Uncertain
Subj Prob (mean)[a]	0.112	0.171	0.132	0.276	0.385	0.335	0.366	0.556	0.491
(std. dev.)	(0.195)	(0.252)	(0.207)	(0.245)	(0.264)	(0.263)	(0.313)	(0.341)	(0.337)
Loss	0.052	0.225	0.073	0.115	0.441	0.286	0.273	0.572	0.433
	(0.222)	(0.417)	(0.26)	(0.32)	(0.497)	(0.452)	(0.446)	(0.495)	(0.496)
Demographics									
Age	71.7	79.7	72.3	54.6	55.0	55.3	70.4	75.3	72.9
	(4.368)	(6.961)	(4.319)	(4.066)	(4.016)	(3.795)	(7.627)	(7.785)	(7.548)
Female	0.618	0.619	0.557	0.453	0.602	0.590	0.595	0.564	0.588
	(0.486)	(0.486)	(0.497)	(0.498)	(0.49)	(0.492)	(0.491)	(0.496)	(0.492)
Health Status Indicators									
Arthritis	0.479	0.613	0.551	0.000	0.553	0.346	0.351	0.435	0.443
	(0.5)	(0.487)	(0.497)	(0)	(0.497)	(0.476)	(0.477)	(0.496)	(0.497)
Diabetes	0.141	0.181	0.150	0.000	0.090	0.082	0.000	0.163	0.185
	(0.348)	(0.385)	(0.357)	(0)	(0.287)	(0.274)	(0)	(0.369)	(0.388)
Heart condition	0.000	0.401	0.432	0.000	0.083	0.061	0.000	0.375	0.332
	(0)	(0.49)	(0.495)	(0)	(0.275)	(0.24)	(0)	(0.484)	(0.471)
Sample Size									
Observations (Ind × wave)	9027	11,259	10,976	763	2216	5534	2689	2362	6800
Unique individuals	4379	3587	5291	391	1280	3018	1720	1371	4270
Unique households	3206	2887	3870	290	975	2362	1419	1145	3545
Fraction insured[b]	14.0%	10.5%	14.6%				65.1%	63.3%	64.2%
Fraction insured (incl Medicaid)	19.5%	20.6%	19.7%						

[a] I transform the life insurance variable to 1 − Pr{living to AGE} to correspond to the loss definition.

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

icaid also pays for nursing home use, I also exclude Medicaid enrollees from this restricted LTC sample. Unfortunately, the HRS does not ask about disability insurance ownership, so I cannot conduct this robustness check for the disability setting.

Finally, although the rejectees have, on average, a higher chance of experiencing the loss than the nonrejectees, it is not certain that they would experience the loss. For example, only 22.5% of rejectees in LTC actually end up going to a nursing home in the subsequent 5 years. This suggests there is substantial unrealized risk among the rejectees.

5.2.4. *Relation to Ideal Data*

Before turning to the results, it is important to be clear about the extent to which the data resemble the ideal data set in each market setting. In general, I approximate the ideal data set quite well, aside from the necessity to classify a relatively large fraction of the sample to the uncertain rejection classification. In disability and in life, I classify a smaller fraction of the sample as rejected or not rejected as compared with LTC. Also, for disability and life, I rely on a smaller set of underwriting guidelines (along with underwriter interviews) to obtain rejection conditions, as opposed to LTC, where I obtain a fairly large fraction of the underwriting guidelines used in the market. In disability and life, I also do not observe medical tests that may be used by insurance companies to price insurance (although conversations with underwriters suggest this is primarily to verify application information, which I approximate quite well using the HRS). In contrast, in LTC, I classify a relatively large fraction of the sample, I closely approximate the set of public information, and I can assess the robustness of the 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 reiterating that all three of the samples approximate the ideal data set quite well, the LTC sample is arguably the best of the three.

6. LOWER BOUND ESTIMATION

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

6.1. *Specification*

All of the empirical estimation is conducted separately for each of the settings and for rejection classifications within each setting. Here I provide an overview of the preferred specification, which controls for the variables used by insurance companies to price insurance. I defer a detailed discussion of all three control specifications to Appendix D.1.[47]

I estimate the distribution of $P_Z = \Pr\{L|X, Z\}$ using a probit specification

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

where X are the control variables (i.e., the pricing controls listed in Table A-I) and $\Gamma(\text{age}, Z)$ captures the relationship between L and Z, allowing it to depend on age.[48] With this specification, the null hypothesis of no private information, $\Pr\{L|X, Z\} = \Pr\{L|X\}$, is tested by restricting $\Gamma = 0$.[49] I choose a flexible functional form for $\Gamma(\text{age}, Z)$ that uses full interactions of basis functions in age and Z. For the basis in Z, I use second-order Chebyshev polynomials plus separate indicators for focal point responses at $Z = 0$, 50, and 100. For the basis in age, I use a linear specification.

With infinite data, one could estimate $E[m_Z(P_Z)|X]$ at each value of X. However, the high dimensionality of X requires being able to aggregate across values of X. To do this, I assume that conditional on ones' age and rejection classification, the distribution of $P_Z - \Pr\{L|X\}$ does not vary with X. This allows the rich set of observables to flexibly affect the mean loss probability, but allows for aggregation of the dispersion of the distribution across values of X.[50]

I then estimate the conditional expectation $m_Z(p) = E[P_Z|P_Z \geq p, X] - p$ using the estimated distribution of $P_Z - \Pr\{L|X\}$ within each age grouping and rejection classification. After estimating $m_Z(p)$, I use the estimated distribution of P_Z to construct its average $E[m_Z(P_Z)|X \in \Theta]$, where Θ is a given sample (e.g., LTC rejectees). I construct the difference between the reject and no reject estimates,

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

and test whether I can reject a null hypothesis that $\Delta_Z \leq 0$.

[47]The central estimation challenge for all specifications and settings is the high dimensionality of the observables, X. This makes it difficult to flexibly estimate the full distribution of P_Z separately for every possible value of X. Throughout, I adopt specifications aimed at flexibly nesting the null hypothesis of no private information, $P_Z = \Pr\{L|X\}$. In other words, I allow the first moment of P_Z to vary flexibly with X. However, the sample size and dimensionality of X limit the extent to which one can allow the higher moments of P_Z to vary flexibly across values of X.

[48]Note that age is an element of X, so that Γ captures the interaction term of age with Z.

[49]At various points in the estimation, I require an estimate of $\Pr\{L|X\}$, which I obtain with the same specification as above, but restricting $\Gamma = 0$.

[50]Note also that I only impose this assumption *within* a setting/rejection classification—I do not require the dispersion of the rejectees to equal that of the nonrejectees. Also note that 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 $\Gamma = 0$). A priori, this assumption would be especially worrisome for the LTC no reject sample, for which the mean loss is near 0.05 and near 0.01 for younger ages. However, the estimates will suggest Z has no predictive content for L in this sample; hence $P_Z = \Pr\{L|X\}$, so that the value of $E[m_Z(P_Z)]$ will be approximately zero for any assumption made about the shape of the distribution of P_Z given X.

6.2. *Statistical Inference*

Statistical inference for $E[m_Z(P_Z)|X \in \Theta]$ 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, estimates of $E[m_Z(P_Z)|X \in \Theta]$ will be positive (Z will always have some predictive power in finite samples). Provided the true value of $E[m_Z(P_Z)|X \in \Theta]$ is positive, the bootstrap provides consistent, asymptotically normal, standard errors for $E[m_Z(P_Z)|X \in \Theta]$ (Newey (1997)). But if the true value of $E[m_Z(P_Z)|X \in \Theta]$ is zero (as would occur if there were no private information among those with $X \in \Theta$), then the bootstrap distribution is not asymptotically normal and does not provide adequate finite-sample inference.[51] Therefore, I supplement the bootstrap with a Wald test that restricts $\Gamma(\text{age}, Z) = 0$.[52] The Wald test is the key statistical test for the presence of private information, as it tests whether Z is predictive of L conditional on X. I report results from both the Wald test and the bootstrap.

I conduct inference on Δ_Z in a similar manner. To test the null hypothesis that $\Delta_Z \leq 0$, I construct conservative p-values by taking the maximum p-value from two tests: (i) a Wald test of no private information held by the rejectees, $E[m_Z(P_Z)|X \in \Theta^{\text{Reject}}] = 0$, and (ii) the p-value from the bootstrapped event of less private information held by the rejectees, $\Delta \leq 0$.[53]

6.3. *Results*

I begin with graphical evidence of the predictive power of the subjective probability elicitations in each sample. Figure 2(a)–(c) plots the estimated distribution of $P_Z - E[P_Z|X]$ aggregated by rejection classification for the rejectees and nonrejectees, using the preferred pricing control specification.[54]

Across all three market settings, the distribution of $P_Z - \Pr\{L|X\}$ appears more dispersed for the rejectees relative to nonrejectees.[55] In this sense, the subjective probability elicitations contain more information about L for the rejectees than for the nonrejectees.[56]

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

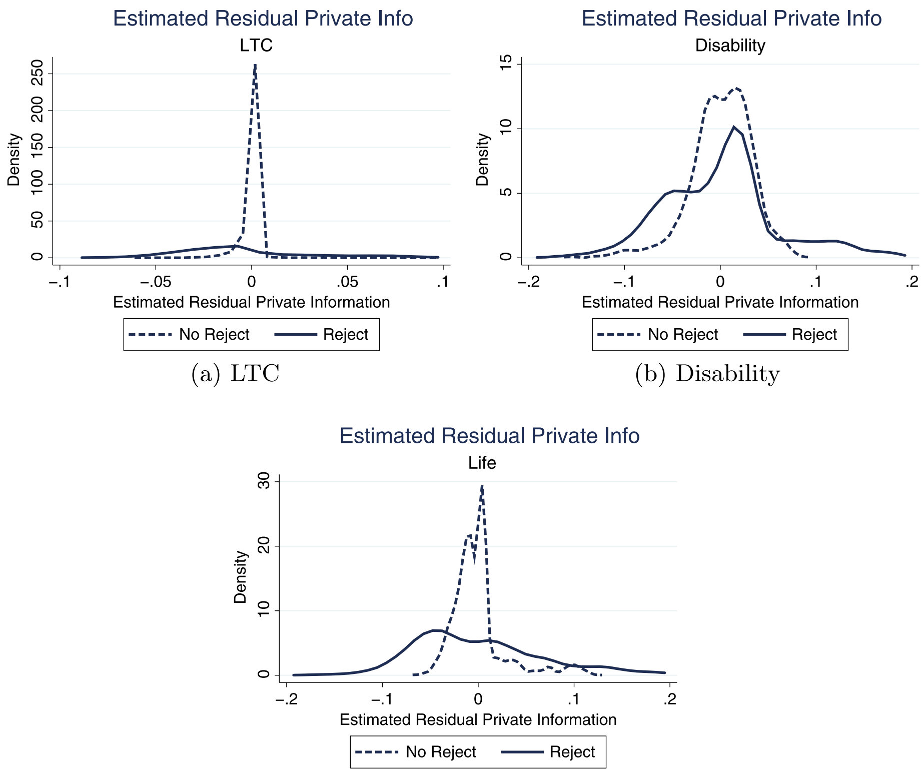
[52]The event $\Gamma(\text{age}, Z) = 0$ in sample Θ is equivalent to both the event $\Pr\{L|X, Z\} = \Pr\{L|X\}$ for all $X \in \Theta$ and the event $E[m_Z(P_Z)|X \in \Theta] = 0$.

[53]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 $\Pr\{E[m_Z(P_Z)|X \in \Theta^{\text{Reject}}] = 0|\Delta \leq 0\}$. Since this weight is unknown, I use these conservative p-values that are robust to any weight in [0, 1].

[54]Subtracting $E[P_Z|X]$ or, equivalently, $\Pr\{L|X\}$, allows for simple aggregation across X within each sample.

[55]The area under each curve sums to 1. Note that the vertical scale in the LTC graph is larger because of the greater mass near zero for the no reject sample.

[56]Appendix D.2 shows that the greater predictive power of the elicitations for rejectees is, from a statistical perspective, driven by a combination of (i) a larger slope of $\Pr\{L|X, Z\}$ with respect to Z for rejectees and (ii) greater dispersion in Z, given X, for the rejectees.



(a) LTC

(b) Disability

(c) Life

FIGURE 2.—Distribution of $P_Z - \Pr\{L|X\}$.

Table III presents the measurements of this dispersion using the average magnitude of private information implied by Z. The first set of rows, labelled "Reject," presents the estimates for the rejectees in each setting and control specification. Across all settings and control specifications, I find significant evidence of private information among the rejectees ($p < 0.001$); the subjective probabilities are predictive of the realized loss, conditional on the set of criteria insurance companies use to price insurance, and also are predictive conditional on the baseline controls (age and gender) and the extended controls.

In addition, the estimates provide an economically significant lower bound on the average magnitude of private information. For example, the estimate of 0.0358 for the LTC price controls specification indicates that if a rejectee was drawn at random, one would expect the average probability of higher risks (with the same observables, X) to be at least 3.58% higher, which is 16% higher than the mean loss probability of 22.5% for LTC rejectees.

TABLE III
LOWER BOUND RESULTS

	LTC			Disability			Life		
Classification	Age & Gender	Price Controls	Extended Controls	Age & Gender	Price Controls	Extended Controls	Age & Gender	Price Controls	Extended Controls
Reject	0.0336***	0.0358***	0.0313***	0.0727***	0.0512***	0.0504***	0.0759***	0.0587***	0.0604***
s.e.[a]	(0.0038)	(0.0037)	(0.0036)	(0.0092)	(0.0086)	(0.0083)	(0.0088)	(0.0083)	(0.0078)
p-value[b]	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
No Reject	0.0048	0.0049	0.0041	0.036	0.024	0.023	0.031**	0.025	0.021
s.e.[a]	(0.0018)	(0.0018)	(0.0018)	(0.0116)	(0.009)	(0.0072)	(0.0076)	(0.007)	(0.0066)
p-value[b]	0.2557	0.3356	0.3805	0.6843	0.8525	0.9324	0.0102	0.1187	0.2395
Difference: Δ_Z	0.0288***	0.0309***	0.0272***	0.0365*	0.027	0.0274*	0.0449***	0.0338***	0.0397***
s.e.[a]	(0.0041)	(0.0041)	(0.0039)	(0.0146)	(0.0127)	(0.0109)	(0.0112)	(0.0107)	(0.0103)
p-value[c]	0.000	0.000	0.000	0.091	0.121	0.092	0.000	0.000	0.001
Uncertain	0.009***	0.0086***	0.0079***	0.0506***	0.0409***	0.0363***	0.0463***	0.0294***	0.028***
s.e.[a]	(0.0024)	(0.0025)	(0.0024)	(0.0058)	(0.0047)	(0.0051)	(0.0058)	(0.0054)	(0.0051)
p-value[b]	0.0001	0.0014	0.0001	0.0000	0.0000	0.0000	0.0000	0.0001	0.0001

[a] Bootstrapped standard errors computed using block re-sampling at the household level (results shown for $N = 1000$ repetitions).

[b] p-value for the Wald test which restricts coefficients on subjective probabilities equal to zero.

[c] p-value is the maximum of the p-value for the rejection group having no private information (Wald test) and the p-value for the hypothesis that the difference is less than or equal to zero, where the latter is computed using bootstrap.

*** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

The third set of rows in Table III provides the estimates of Δ_Z. Again, across all specifications and market settings, I estimate larger lower bounds on the average magnitude of private information for the rejectees relative to those served by the market. These differences are statistically significant at the 1% level in LTC and life, and positive (but not significant at standard levels) in disability.[57]

Not only do I find smaller amounts of private information for the nonrejectees, but I cannot actually reject the null hypothesis of no private information among this group once the set of variables insurers use to price insurance is included, as indicated by the second set of rows in Table III.[58] This provides a new explanation for why previous research has not found significant amounts of adverse selection of insurance contracts in LTC (Finkelstein and McGarry (2006)) and life insurance (Cawley and Philipson (1999)). The practice of rejections by insurance companies limits the extent to which private information manifests itself in adverse selection of contracts.

6.4. *Age 80 in LTC insurance*

LTC insurers reject applicants above age 80, regardless of health status. This provides an opportunity for a finer test of the theory by exploring whether those who do not have rejection health conditions start to obtain private information at age 80. To do so, I construct a series of estimates of $E[m_Z(P_Z)]$ by age for the set of people who do not have a rejection health condition and thus would only be rejected if their age exceeded 80.[59]

[57]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 III. In general, the theory does not have a prediction for the uncertain group. However, if $E[m_Z(P_Z)|X]$ takes on similar values for all rejectees (e.g., $E[m_Z(P_Z)|X] \approx m^{\mathrm{R}}$) and nonrejectees (e.g., $E[m_Z(P_Z)|X] \approx m^{\mathrm{NR}}$), then linearity of the expectation implies

$$E\big[m_Z(P_Z)|X \in \Theta^{\text{Uncertain}}\big] = \lambda m^{\mathrm{R}} + (1-\lambda)m^{\mathrm{NR}}, \tag{7}$$

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

[58]Of course, the difference between the age and gender specification and the price controls specification is not statistically significant. Also, the inability to reject a null of no private information is potentially driven by the small sample size in the disability setting, but the LTC sample of nonrejectees is quite large (>9K) and the sample of nonrejectees in life is larger than the sample of rejectees.

[59]To ensure that no information from those who have rejection health conditions is used in the construction of $E[m_Z(P_Z)]$ for those who do not have health conditions and are above age 80, I split the reject sample into two groups: those who do not have a rejection health condition (and thus would only be rejected because their age is above 80) and those who do have a rejection

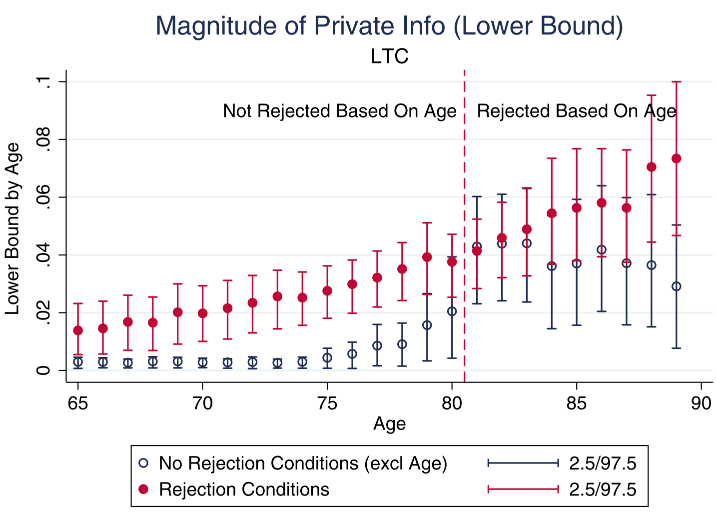


FIGURE 3.—Magnitude of private information by age and rejection classification.

Figure 3 plots the results for those who do not have health conditions (hollow circles), along with a comparison set of results for those who do have rejection health conditions (filled circles).[60] The figure suggests that the subjective probability elicitations of those who do not have rejection health conditions become predictive of L right around age 80—exactly the age at which insurers choose to start rejecting applicants based on age, regardless of health status. Indeed, from the perspective of $E[m_Z(P_Z)]$, a healthy 81 year old looks a lot like a 70 year old who had a stroke. This is again consistent with the theory that private information limits the existence of insurance markets.

6.5. *Robustness*

Moral Hazard

As discussed in Section 5.2.3, one alternative hypothesis is that the private information I estimate is the result of moral hazard from insurance contract

condition. I estimate P_Z separately on these two samples using the pricing specification outlined in Section 6.1.

[60]The graph presents bootstrapped 95% confidence intervals adjusted for bias using the nonaccelerated procedure suggested in Efron and Gong (1983). These are appropriate confidence intervals as long as the true magnitude of private information is positive. In the aggregate sample of rejectees, I reject the null hypothesis of no private information (see Table III). However, for any particular age, I am unable to reject a null hypothesis of no private information using the Wald test. As a result, the shown standard errors do not incorporate the null hypothesis of no private information separately for each age.

TABLE IV

ROBUSTNESS TO MORAL HAZARD: NO INSURANCE SAMPLE

	LTC, Price Controls		Life, Price Controls	
	Primary Sample	Excluding Insured	Primary Sample	Excluding Insured
Reject	0.0358***	0.0351***	0.0587***	0.0491*
s.e.[a]	(0.0037)	(0.0041)	(0.0083)	(0.0115)
p-value[b]	0.0000	0.0000	0.0000	0.0523
No Reject	0.0049	0.0038	0.0249	0.0377
s.e.[a]	(0.0018)	(0.0019)	(0.007)	(0.0107)
p-value[b]	0.3356	0.8325	0.1187	0.2334
Difference: Δ_Z	0.0309***	0.0313***	0.0338***	0.011
s.e.[a]	(0.0041)	(0.0046)	(0.0107)	(0.0157)
p-value[c]	0.000	0.000	0.000	0.301
Uncertain	0.0086***	0.0064	0.0294***	0.0269
s.e.[a]	(0.0025)	(0.0024)	(0.0054)	(0.0078)
p-value[c]	0.0014	0.1130	0.0001	0.1560

[a]Bootstrapped standard errors computed using block re-sampling at the household level (results shown for $N =$ 1000 repetitions).

[b] p-value for the Wald test which restricts coefficients on subjective probabilities equal to zero.

[c] p-value is the maximum of the p-value for the rejection group having no private information (Wald test) and the p-value for the hypothesis that the difference is less than or equal to zero, where the latter is computed using bootstrap.

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

choice, not an underlying heterogeneity in loss probabilities. To assess whether this is driving any of the results, I reestimate the average magnitude of private information implied by Z on samples in LTC and life that exclude those who currently own insurance. For LTC, I exclude those who own private LTC insurance along with those who are currently enrolled in Medicaid, since it pays for nursing home stays. As shown in Table II, this excludes 20.6% of the sample of rejectees and 19.5% of nonrejectees. For life, I exclude those who have any life insurance policy. Unfortunately, this excludes 63% of the rejectees and 65% of the nonrejectees; thus the remaining sample is quite small.

Table IV presents the results. For LTC, I continue to find amounts of private information for the rejectees ($p < 0.001$) that is significantly greater than for the nonrejectees ($\Delta_Z = 0.0313$, $p < 0.001$), and I cannot reject the null hypothesis of no private information for the nonrejectees ($p = 0.8325$). For life, I estimate marginally significant amounts of private information for the rejectees ($p = 0.0523$) of a magnitude similar to what is estimated on the full sample (0.0491 versus 0.0587). I estimate more private information for the rejectees relative to the nonrejectees; however, the difference is no longer statistically significant ($\Delta_Z = 0.011$, $p = 0.301$), which is arguably a result of the reduced sample size. I also continue to be unable to reject the null hypothesis of no private information for the nonrejectees ($p = 0.2334$). In short, the results suggest that moral hazard is not driving my findings of private information for the rejectees and more private information for the rejectees relative to the nonrejectees.

Additional Robustness Checks

Appendix D.3 contains a couple of additional robustness checks. I present the age-based plots, similar to Figure 3, for the disability and life settings, and show that I generally find greater amounts of private information across all age groups for the rejectees in each setting. I also present an additional specification in life insurance that includes additional cancer controls, discussed in Appendix C.1, that are available for a smaller sample of the HRS data; I show that the estimates are similar when these additional controls are introduced.

6.6. *Summary*

In all three market settings, I estimate a significant amount of private information held by the rejectees that is robust to a wide set of controls for public information. I find more private information held by the rejectees relative to the nonrejectees, and I cannot reject a null hypothesis of no private information held by those who actually are served by the market. Moreover, a deaggregated analysis of the practice of LTC insurers of rejecting all applicants above age 80 (regardless of health) reveals that healthy individuals begin to have private information right around age 80—precisely the age chosen by insurers to stop selling insurance. In sum, the results are consistent with the theory that private information leads to insurance rejections.

7. ESTIMATION OF THE DISTRIBUTION OF PRIVATE INFORMATION

While the lower bound results, in particular, the stark pattern of the presence of private information, provide support for the theory that private information would afflict a hypothetical insurance market for the rejectees, it does not establish whether the amount of private information is sufficient to explain why insurers do not sell policies to the rejectees. This requires an estimate of the minimum pooled price ratio and, hence, an estimate of the distribution of private information, $F(p|X)$. To do so, I follow the second approach, which is outlined in Section 4.2: I impose additional structure on the relationship between elicitations Z and true beliefs P that allows for a flexible estimation of $F(p|X)$.

7.1. *Empirical Specification*

7.1.1. *Elicitation Error Model*

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 who have the same beliefs may have different elicitations. Moreover, as shown in Figure 1(a)–(c) and recognized in previous literature (e.g., Gan, Hurd, and McFadden (2005)), people may have a tendency to report focal point values at 0, 50, and 100%. My model of elicitations capture all three of these forms of elicitation error.

To illustrate the model, first define the random variable

$$\tilde{Z} = P + \varepsilon,$$

where $\varepsilon \sim N(\alpha, \sigma^2)$. The variable $\tilde{Z}$ is a noisy measure of beliefs with bias α and noise variance σ^2, where the error follows a normal distribution. I assume there are two types of responses: focal point responses and nonfocal point responses. With probability $1 - \lambda$, an agent gives a nonfocal point response

$$Z^{\mathrm{nf}} = \begin{cases} \tilde{Z} & \text{if } \tilde{Z} \in [0, 1], \\ 0 & \text{if } \tilde{Z} < 0, \\ 1 & \text{if } \tilde{Z} > 1, \end{cases}$$

which is $\tilde{Z}$ censored to the interval $[0, 1]$. These responses are continuously distributed over $[0, 1]$ with some mass at 0 and 1.

The second type of responses is focal point responses. With probability λ, an agent reports

$$Z^{\mathrm{f}} = \begin{cases} 0 & \text{if } \tilde{Z} \le \kappa, \\ 0.5 & \text{if } \tilde{Z} \in (\kappa, 1 - \kappa), \\ 1 & \text{if } \tilde{Z} \ge 1 - \kappa, \end{cases}$$

where $\kappa \in [0, 0.5)$ captures the focal point window. With this structure, focal point responses have the same underlying structure as nonfocal point responses, but are reported on a scale of low, medium, and high as opposed to a continuous scale on $[0, 1]$.[61] As a result, nonfocal point responses will contain more information about P than do focal point responses. Therefore, most of the identification for the distribution of P will come from responses that report nonfocal point values.

Given this model, I have four elicitation parameters to be estimated, $\{\alpha, \sigma, \kappa, \lambda\}$, which will be estimated separately in each market setting and classification. This allows for the potential that rejectees have a different elicitation error process than nonrejectees.

[61]Note that I do assume that the act of providing a focal point response is not informative of P (λ is not allowed to be a function of P). Ideally, one would allow focal point respondents to have differing beliefs from nonfocal point respondents, yet the focal point bias inherently limits the extent of information that can be extracted from their responses.

7.1.2. *Flexible Approximation for the Distribution of Private Information*

With infinite data, one could flexibly estimate $f(p|X)$ separately for every possible value of X and p. Faced with finite data and a high dimensional X, this is not possible. Since the minimum pooled price ratio is essentially a function of the shape of the distribution of $f(p|X)$ across values of p, I choose a specification that allows for considerable flexibility across p. In particular, I assume $f(p|X)$ is well approximated by a mixture of beta distributions,

$$f(p|X) = \sum_i w_i \,\text{Beta}\big(p|a_i + \Pr\{L|X\}, \psi_i\big), \tag{8}$$

where $\text{Beta}(p|\mu, \psi)$ is the p.d.f. of the beta distribution with mean μ and shape parameter ψ.[62] With this specification, $\{w_i\}$ governs the weights on each beta distribution, $\{a_i\}$ governs the noncentrality of each beta distribution, and ψ_i governs the dispersion of each beta distribution. The flexibility of the beta distributions ensures that I impose no restrictions on the size of the minimum pooled price ratio.[63] For the main specification, I include three beta distributions.[64] Additional details of the specification are provided in Appendix E.1.

7.1.3. *Pooled Price Ratio (and Its Minimum)*

With an estimate of $f(p|X)$, the pooled price ratio is easily constructed as $T(p) = \frac{E[P|X]}{1-E[P|X]}\frac{1-p}{p}$ for each p, where $E[P|P \geq p, X]$ is computed using the estimated $f(p|X)$. Throughout, I focus on estimates evaluated for a mean loss characteristic, $\Pr\{L|X\}$. In principle, one could analyze the pooled price ratio

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

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

where $B(\alpha, \beta)$ is the beta function. 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 $\psi = \alpha + \beta$.

[63]In principle, the event of no private information is captured with $\psi_1 \to \infty$, $a_1 = 0$, and $w_1 = 1$. For computational reasons, I need to impose a cap on ψ_i in the estimation. In the initial estimation, this cap binds for the central most beta distribution in both the LTC no reject and disability no reject samples. Intuitively, the model wants to estimate a large fraction of very homogeneous individuals around the mean. Therefore, for these two samples, I also include a point-mass distribution with weight w_0 in addition to the three beta distributions. This allows me to capture a large concentration of mass in a way that does not require integrating over a distribution $f(p|X)$ with very high curvature. Appendix E.1 provides further details.

[64]While (8) allows for a very flexible shape of $f(p|X)$ across p, it is fairly restrictive in how this shape varies across values of X. Indeed, I do not allow the distribution parameters to vary with X. This is a practical necessity due to the size of my samples and the desire to allow for a very flexible shape for $f(p|X)$. Moreover, it is important to stress that I will still separately estimate $f(p|X)$ for the rejectees and the nonrejectees using the separate samples.

across all values of X, but given the specification, focusing on differing values of X or $\Pr\{L|X\}$ does not yield an independent test of the theory. In Appendix E.2, I show that the results are generally robust to focusing on values of $\Pr\{L|X\}$ at the 20, 50, and 80th percentiles of its distribution.

As described in Section 4.2, I estimate the analogue of the minimum pooled price ratio, $\inf_{p\in\hat{\Psi}_\tau} T(p)$, for the restricted domain $\hat{\Psi}_\tau = [0, F^{-1}(\tau)]$. My 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, I 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, $F^{-1}(0.9)$.

7.1.4. *Identification*

Before turning to the results, it is important to understand the sources of identification for the model. As discussed above, much of the model is identified from the nonfocal point responses. If the elicitation error parameters were known, then identification of the distribution of P is a deconvolution of the distribution of Z^{nf}; thus, the empirical distribution of nonfocal elicitations provides a strong source of identification for the distribution of P conditional on having identified the elicitation error parameters.[65]

To identify the elicitation error parameters, the model relies on the relationship between Z^{nf} and L. To see this, note that Assumptions 1 and 2 imply

$$E\big[Z^{\text{nf}} - P\big] = E\big[Z^{\text{nf}}\big] - E[L],$$

so that the mean elicitation bias is the difference between the mean elicitation and the mean loss probability. This provides a strong source of identification for α.[66] In practice, the model calculates α jointly with the distribution of P to adjust for the fact that the nonfocal elicitations are censored over [0, 1].

To identify σ, note that Assumptions 1 and 2 imply

$$\text{var}\big(Z^{\text{nf}}\big) - \text{cov}\big(Z^{\text{nf}}, L\big) = \text{var}\big(Z^{\text{nf}} - P\big) + \text{cov}\big(Z^{\text{nf}} - P, P\big), \tag{9}$$

where $\text{var}(Z^{\text{nf}} - P)$ is the variance of the nonfocal elicitation error and $\text{cov}(Z^{\text{nf}} - P, P)$ is the correction term that accounts for the fact that I al-

[65] If Z^{nf} were not censored on $[0, 1]$, then P would be nonparametrically identified from the observation of the distribution of $Z^{\text{nf}} = \tilde{Z}$ (this follows from the completeness of the exponential family of distributions). However, since I have modeled the elicitations as being censored at 0 and 1, some distributions of P, especially those leading to a lot of censored values, may not be nonparametrically identified solely from the distribution of Z^{nf} and may also rely on moments of the joint distribution of Z^{nf} and L for identification.

[66] Indeed, if Z^{nf} were not censored on $[0, 1]$, this quantity would equal α.

low nonfocal elicitations to be censored on $[0, 1]$.[67] The quantity $\text{var}(Z^{\text{nf}}) - \text{cov}(Z^{\text{nf}}, L)$ is the variation in Z that is not explained by L. Since the primary impact of changing σ is to change the elicitation error variance of $Z^{\text{nf}} - P$, the value of $\text{var}(Z^{\text{nf}}) - \text{cov}(Z^{\text{nf}}, L)$ provides a strong source of identification for σ.[68] Finally, the fraction of focal point respondents, λ, and the focal point window, κ, are identified from the distribution of focal points and the loss probability at each focal point.

7.1.5. *Statistical Inference*

Bootstrapping delivers appropriate confidence intervals for the estimates of $\inf_{p\in[0,F^{-1}(\tau)]} T(p)$ and the values of $f_P(p|X)$ and $F_P(p|X)$ as long as the estimated parameters are in the interior of their potential support. This assumption is violated in the potentially relevant case in which there is no private information. In this case, $\psi_1 \to \infty$, $w_1 = 1$, and $a_1 = 0$. As with the lower bound approach, the problem is that in finite samples, one may estimate a nontrivial distribution of P even if the true P is only a point mass. Because the parameters are at a boundary, one cannot use bootstrapped estimates to rule out the hypothesis of no private information.

To account for the potential that individuals have no private information, I again use the Wald test from the lower bound approach (see Table III) that tests whether $\Pr\{L|X, Z\} = \Pr\{L|X\}$ for all X in the sample (by restricting $\Gamma = 0$).[69] I construct 5/95% confidence intervals for $\inf_{p\in\hat{\Psi}_\tau} T(p)$ by combining bootstrapped confidence intervals (CI) and extending the 5% boundary to 1 in

[67]To see this, note that

$$\text{var}(Z^{\text{nf}}) = \text{var}(Z^{\text{nf}} - P) + \text{var}(P) + 2\,\text{cov}(Z^{\text{nf}} - P, P)$$

and

$$\text{cov}(Z^{\text{nf}}, L) = \text{cov}(Z^{\text{nf}} - P, P) + \text{cov}(P, L) = \text{cov}(Z^{\text{nf}} - P, P) + \text{var}(P),$$

where the latter equality follows from $\Pr\{L|P\} = P$. Subtracting these equations yields (9).

[68]More generally, Assumptions 1 and 2 impose an infinite set of moment conditions that can be used to identify the elicitation parameters:

$$E[P^N|L = 1]\Pr\{L\} = E[P^{N+1}].$$

It is easy to verify that $N = 0$ provides the source of identification for α mentioned above and $N = 1$ provides the source of identification for σ. This expression suggests one could, in principle, allow for a richer specification of the elicitation error; I leave the interesting but difficult question of the nonparametric identification conditions on the elicitation error for future work.

[69]This test also has the advantage that misspecification of $f_{Z|P}$ will not affect the test for private information. But, in principle, one could use the structural assumptions made on $f_{Z|P}$ to generate a more powerful test for the presence of private information. 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) provided a potential method for constructing an appropriate test; this is left for future work.

the event that I cannot reject a null hypothesis of no private information at the 5% level. Given the results in Table III, this amounts to extending the 5/95% CI to include 1 for the nonrejectees in each of the three settings.

I will also present graphs of the estimated p.d.f., $f_P(p|X)$, cumulative distribution function (c.d.f.), $F_P(p|X)$, and pooled price ratio, $T(p)$, evaluated at the mean characteristic, $\Pr\{L|X\} = \Pr\{L\}$, in each sample. For these, I present the 95% confidence intervals and do not attempt to incorporate information from the Wald test. The reader should keep in mind that one cannot reject $F(p|X) = 1\{p \leq \Pr\{L|X\}\}$ at the 5% level for the nonrejectees in any of the three settings.[70] Also, for the estimated confidence intervals of $F_P(p|X)$, I impose monotonicity in a conservative fashion by defining $F_P^5(p|X) = \min_{\hat{p} \leq p} \hat{F}_P^5(p|X)$ and $F_P^{95}(p|X) = \max_{\hat{p} \geq p} \hat{F}_P^{95}(p|X)$, where $\hat{F}_P^5(p|X)$ and $\hat{F}_P^{95}(p|X)$ are the estimated pointwise 5/95% confidence thresholds from the bootstrap.

7.2. *Estimation Results*

Qualitatively, no trade is more likely for distributions with a thick upper tail of high risks, the presence of which inhibits the provision of insurance to lower risks by raising the value of $E[P|P \geq p]$. In each market setting, I find evidence consistent with this prediction. Figure 4 presents the estimated p.d.f. $f_P(p|X)$ and c.d.f. $F_P(p|X)$ for each market setting, plotted for a mean characteristic within each sample using the price controls, X.[71] The solid line presents estimates for the rejectees; the dotted line for nonrejectees. Across all three settings, there is qualitative evidence of a thick upper tail of risks as $p \to 1$ for the rejectees. In contrast, for the nonrejectees, there is less evidence of such an upper tail.

Figure 4 translates these estimates into their implied pooled price ratio, $T(p)$, for $p \leq F^{-1}(0.8)$, and Table V presents the estimated minimums over this same region, $\inf_{p \in [0, F^{-1}(0.8)]} T(p)$. Across all three market settings, I estimate a sizable minimum pooled price ratio for the rejectees: 1.82 in LTC (5/95% CI [1.657, 2.047]), 1.66 in disability (5/95% CI [1.524, 1.824]), and 1.42 in life (5/95% CI [1.076, 1.780]). In contrast, in all three market settings I esti-

[70] Estimates of the p.d.f., c.d.f., and minimum pooled price ratio exhibited considerable bias in the bootstrap estimation, especially among the life and disability settings since they have smaller samples. To be conservative, I present confidence intervals that are the union of bias-corrected confidence intervals (Efron and Gong (1983)) and the more traditional studentized-t confidence intervals. In practice, the studentized-t confidence intervals tended to be wider than the bias-corrected confidence intervals for the disability and life estimates. However, the use of either of these methods does not affect the statistical conclusions.

[71] This involves setting $\Pr\{L|X\} = \Pr\{L\}$ in (8) within each sample (e.g., $\Pr\{L\} = 0.052$ for the LTC no reject sample; the other means are reported in Table II). Appendix E.2 shows that the general conclusions are robust to focusing on other values of $\Pr\{L|X\}$ in each sample; I focus on the mean since it is the most in-sample estimate.

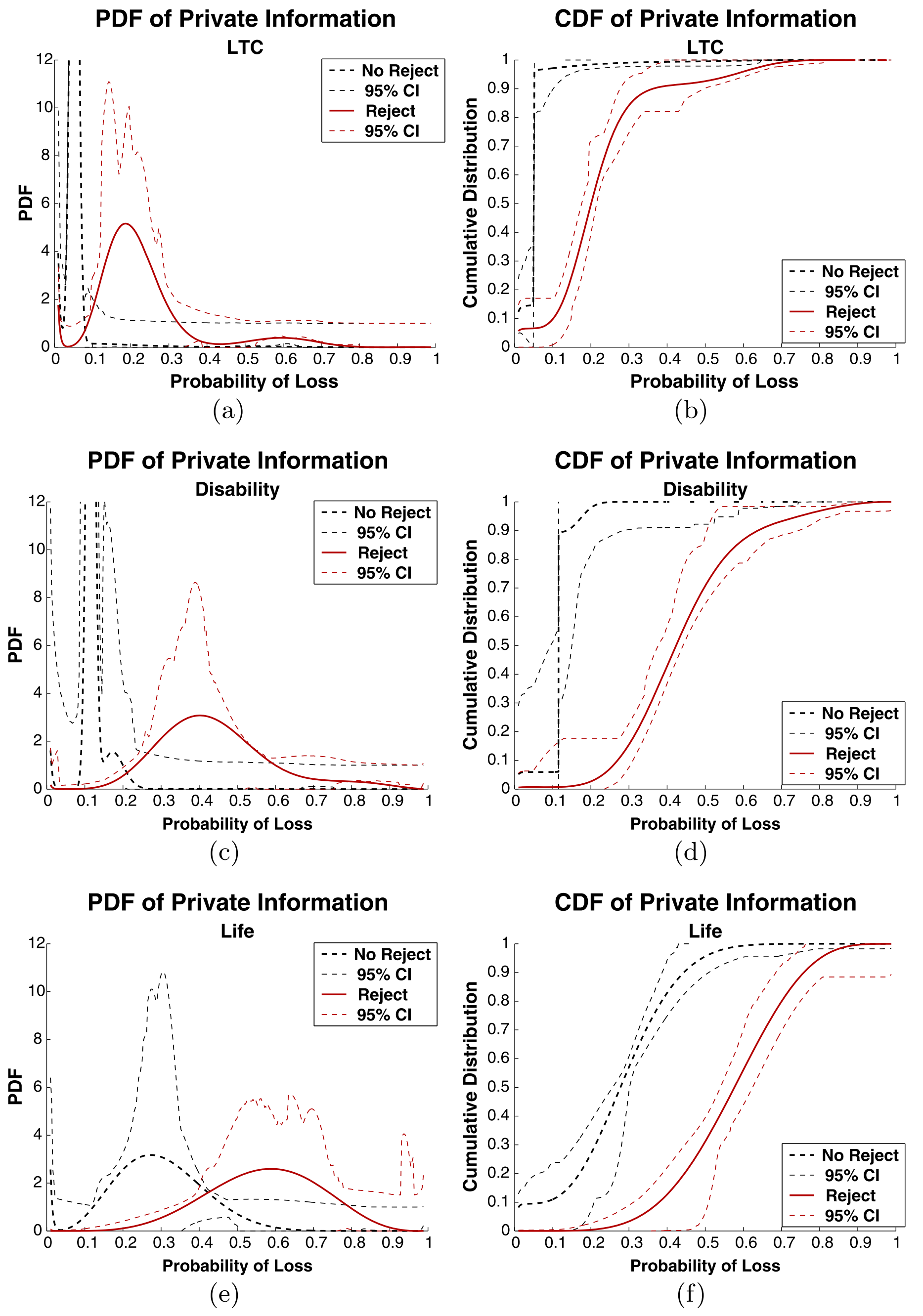


FIGURE 4.—Distribution of private information.

TABLE V

MINIMUM POOLED PRICE RATIO[a]

	LTC	Disability	Life
Reject	**1.827**	**1.661**	**1.428**
5%[b]	1.657	1.524	1.076
95%	2.047	1.824	1.780
No Reject	**1.163**	**1.069**	**1.350**
5%[b]	1.000	1.000	1.000
95%	1.361	1.840	1.702
Difference	**0.664**	**0.592**	**0.077**
5%[c]	0.428	0.177	−0.329
95%	0.901	1.008	0.535

[a]Minimum Pooled Price Ratio evaluated for X s.t. $\Pr\{L|X\} = \Pr\{L\}$ in each sample.

[b]5/95% CI computed using bootstrap block re-sampling at the household level ($N = 1000$ repetitions); 5% level extended to include 1.00 if p-value of F-test for presence of private information is less than. 0.05; Bootstrap CI is the union of the percentile-t bootstrap and bias corrected (nonaccelerated) percentile intervals from Efron and Gong (1983).

[c]5/95% CI computed using bootstrap block re-sampling at the household level ($N = 1000$ repetitions); 5% level extended to include 1.00 if p-value of F-test for presence of private information for the rejectees is less than 0.05; Bootstrap CI is the union of the percentile-t bootstrap and bias corrected (nonaccelerated) percentile intervals from Efron and Gong (1983).

mate smaller minimum pooled price ratios for the nonrejectees. Moreover, consistent with the prediction of Corollary 3, the estimated differences between rejectees and nonrejectees are large and significant in both LTC and disability (roughly 59%); for life, the difference is positive (8%) but not statistically different from zero.

The estimates suggest that an insurance market cannot exist for the rejectees unless they are willing to pay an 82% implicit tax in LTC, a 66% implicit tax in disability, and a 42% implicit tax in life. These implicit taxes are large enough relative to the magnitudes of willingness to pay found in existing literature and those implied by simple models of insurance. For LTC, there is no exact estimate that corresponds to the willingness to pay for a marginal amount of LTC insurance, but Brown and Finkelstein (2008) suggested most 65 year olds are not willing to pay more than a 60% markup for existing LTC insurance policies.[72] For disability, Bound, Cullen, Nichols, and Schmidt (2004) calibrated

[72]More specifically, the results of Brown and Finkelstein (2008) imply that an individual at the 60–70th percentile of the wealth distribution is willing to pay roughly a 27–62% markup for existing LTC insurance policies This is not reported directly, but can be inferred from Figure 1 and Table II. Figure 2 suggests the break-even point for insurance purchase is at the 60–70th percentile of the wealth distribution. Table II shows this corresponds to individuals being willing to pay a tax of 27–62%. Their model would suggest that those above the 80th percentile of the wealth distribution are willing to pay a substantially higher implicit tax; however, Lockwood (2012) showed that incorporating bequest motives significantly reduces the demand for LTC insurance in the upper income distribution.

the marginal willingness to pay for an additional unit of disability insurance to be roughly 46–109%. This estimate is arguably an overestimate of the willingness to pay for insurance because the model calibrates the insurance value using income variation, not consumption variation, which is known to be less variable than income. Nonetheless, the magnitudes are of a similar level to the implicit tax of 66% for the disability rejectees.[73] Finally, if a loss incurs a 10% drop in consumption and individuals have constant relative risk aversion (CRRA) preferences with a coefficient of 3, then $\frac{u'(w-l)}{u'(w)} = 1.372$, so that individuals would be willing to pay a 37.2% markup for insurance, a magnitude that roughly rationalizes the pattern of trade in all three market settings.[74] In short, the size of the estimated implicit taxes suggests that the barrier to trade imposed by private information is large enough to explain a complete absence of trade for the rejectees.

Robustness to Choice of τ

The results in Table V focus on the results for $\tau = 80\%$. Table VI assesses the robustness of the findings to the choice of τ by also presenting results for $\tau = 0.7$ and $\tau = 0.9$. In general, the results are quite similar. For LTC and disability, both the minimums for the rejectees and nonrejectees are obtained at an interior point of the distribution, so that the estimated minimum is unaffected by the choice of τ in the region $[0.7, 0.9]$. For life, the minimums are obtained at the end points, so that changes in τ do affect the estimated minimum. At $\tau = 0.7$, the minimum pooled price ratio rises to 1.488 for the rejectees and 1.423 for the nonrejectees; at $\tau = 0.9$, the minimum pooled price ratio drops to 1.369 for the rejectees and 1.280 for the nonrejectees. In general, the results are similar across values of τ.

Additional Robustness Checks

The results in Tables V and VI evaluate the minimum pooled price ratio for a characteristic, X, that corresponds to a mean loss probability within each sample, $\Pr\{L|X\} = \Pr\{L\}$. In Appendix E.2, I show that the estimates are quite similar if, instead of evaluating at the mean, one chooses X such that $\Pr\{L|X\}$ lies at the 20th, 50th, or 80th quantile of its within-sample distribution.[75] The

[73]See column 6 of Table 2 in Bound et al. (2004). The range results from differing samples. The lowest estimate is 46% for workers with no high school diploma and 109% for workers with a college degree. The sample age range of 45–61 is roughly similar to the age range used in my analysis.

[74]To the best of my knowledge, there does not exist a well estimated measure of the marginal willingness to pay for an additional unit of life insurance.

[75]Because of the choice of functional form for $f_P(p|X)$, these should not be considered to be separate statistical tests of the theory. The functional form is restrictive in the extent to which the shape of the distribution can vary across values of X within a rejection classification. But, nonetheless it is important to ensure that the results do not change simply by focusing on different levels of the index, $\Pr\{L|X\}$.

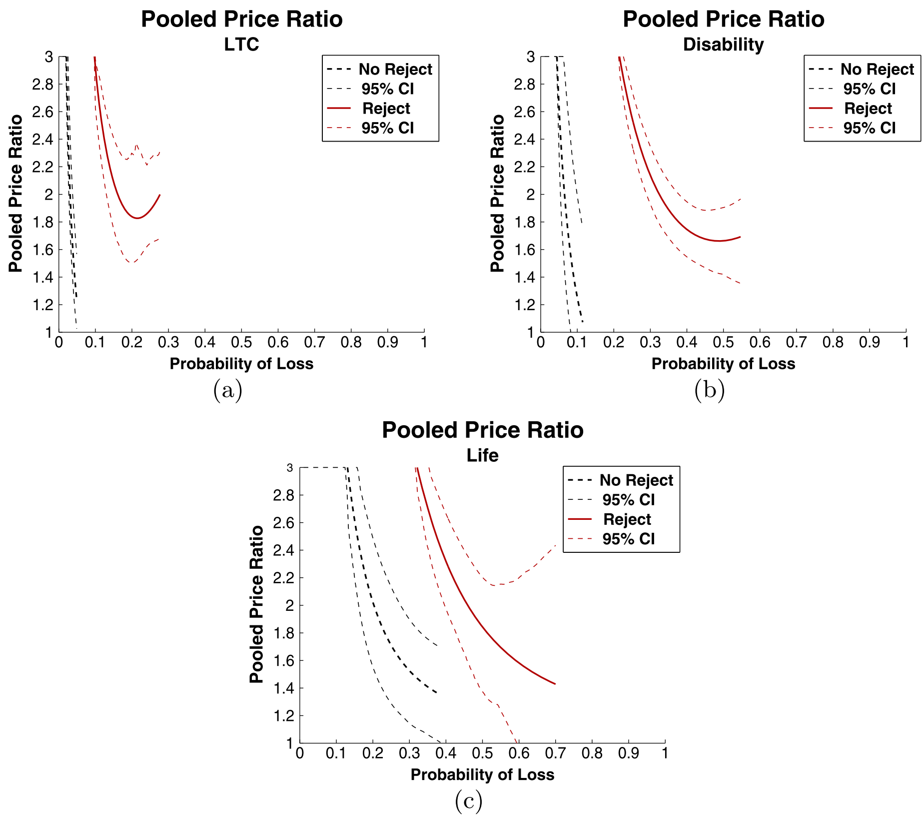


FIGURE 5.—Pooled price ratios.

minimum pooled price ratio for rejectees ranges from 1.77 to 2.09 in LTC, 1.659 to 1.741 in disability, and 1.416 to 1.609 in life. For the nonrejectees, I estimate significantly smaller magnitudes in LTC and disability, and the estimated differences between rejectees and nonrejectees for life remain statistically indistinct from zero.

8. DISCUSSION

The results shed new light on many patterns found in existing literature and pose new questions for future work.

8.1. *Generalizing the Results*

The general empirical finding from the three settings I consider can be summarized succinctly: *there is one way to be healthy, but many (unobservable) ways*

TABLE VI

MINIMUM POOLED PRICE RATIO: ROBUSTNESS TO t

	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	**1.827**	**1.827**	**1.827**	**1.661**	**1.661**	**1.661**	**1.488**	**1.428**	**1.369**
5%[a]	1.661	1.657	1.624	1.518	1.524	1.528	1.124	1.076	1.000
95%	2.250	2.047	2.030	1.824	1.824	1.795	1.815	1.780	1.754
No Reject	**1.163**	**1.163**	**1.163**	**1.069**	**1.069**	**1.069**	**1.423**	**1.350**	**1.280**
5%[a]	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
95%	1.361	1.361	1.366	1.918	1.840	1.728	1.750	1.702	1.665
Difference	**0.664**	**0.664**	**0.664**	**0.592**	**0.592**	**0.592**	**0.065**	**0.077**	**0.089**
5%[b]	0.430	0.428	0.407	0.158	0.177	0.215	−0.344	−0.329	−0.340
95%	1.026	0.901	0.922	1.026	1.008	0.970	0.505	0.535	0.558

[a] 5/95% CI computed using bootstrap block re-sampling at the household level ($N = 1000$ repetitions); 5% level extended to include 1.00 if p-value of F-test for presence of private information is less than 0.05; Bootstrap CI is the union of the percentile-t bootstrap and bias corrected (nonaccelerated) percentile intervals from Efron and Gong (1983).

[b] 5/95% CI computed using bootstrap block re-sampling at the household level ($N = 1000$ repetitions); 5% level extended to include 1.00 if p-value of F-test for presence of private information for the rejectees is less than 0.05; Bootstrap CI is the union of the percentile-t bootstrap and bias corrected (nonaccelerated) percentile intervals from Efron and Gong (1983).

to be sick. The sick are sick in their own unique ways; as a result, the potential for adverse selection prevents insurers from being able to offer insurance to the sick.

This general empirical finding can also explain the pattern of rejections in other insurance markets. For example, nongroup health insurers often reject the sick (i.e., individuals with so-called preexisting conditions); in contrast the observably healthy are generally offered insurance policies.

In addition to explaining general patterns of rejections, this can also explain why there are no rejections in annuity markets. Some people with health conditions know that they are exceptionally high mortality risks, but no one knows they are exceptionally low mortality risks (there is only one way to be healthy). Hence, annuity companies can sell to an average person who has no major health conditions without the risk of it being adversely selected by an even healthier subset of the population. Annuities may be adversely selected, as the sick choose not to buy them (as shown in Finkelstein and Poterba (2002, 2004)), but by reversing the direction of the incentive constraints, rejections no longer occur.[76]

[76] Moreover, the presence of private information among those who have health conditions may explain why annuity companies are generally reluctant to offer discounts to those who have health conditions.

8.2. *Welfare*

My results suggest that the practice of rejections by insurers is *constrained* efficient. Insurance cannot be provided without relaxing one of the three implementability constraints. Either insurers must lose money or be subsidized (relax the resource constraint), individuals must be convinced to be irrational (relax the incentive constraint), or agents' outside option must be adjusted via mandates or taxation (relax the participation constraint). However, policymakers must ask whether they like the constraints. Indeed, the first-best utilitarian allocation is full insurance for all, $c = W - E[p]L$, which could be obtained through subsidies or mandates that use government conscription to relax the participation constraints.

However, literal welfare conclusions based on the stylized model in this paper should be highly qualified. The model abstracts from many realistic features such as preference heterogeneity, moral hazard, and the dynamic aspect of insurance purchase. Indeed, the latter may be quite important for understanding welfare. Although my analysis asks why the insurance market shuts down, I do not address why those who face rejection did not purchase a policy before they obtained the rejection condition. Perhaps they do not value insurance (in which case mandates may lower welfare) or perhaps they face credit constraints (in which case mandates may be beneficial). Unpacking the decision of when to purchase insurance in the presence of potential future rejection is an interesting direction for future work.

8.3. *Group Insurance Markets*

Although this paper focuses on nongroup insurance markets, much insurance is sold in group markets, often through one's firm. For example, more than 30% of nongovernment U.S. workers have group-based disability insurance, whereas just 3% of workers have a nongroup disability policy (ACLI (2010)). Similarly, in health insurance, 49% of the U.S. population have an employer-based policy, whereas only 5% have a nongroup policy.[77]

While it is commonplace to assume that the tax advantage status for employer-sponsored health insurance causes more insurance to be sold in group versus nongroup health insurance markets, tax advantages cannot explain the same pattern in disability insurance. Disability benefits are always taxed regardless of whether the policy is sold in the group or nongroup market.[78] This suggests group markets may be more prevalent because of their ability to deal with informational asymmetries. Indeed, group markets can po-

[77]Figures according to Kaiser Health Facts, http://kff.org/statedata/.

[78]If premiums are paid with after-tax income, then benefits are not taxed. If premiums are paid with pre-tax income (as is often the case with an employer plan), then benefits are taxed.

tentially relax participation constraints by subsidizing insurance purchase for its members. Identifying and quantifying this mechanism is an important direction for future work, especially for understanding the impact of government policies that attempt to promote either the individual or the group-based insurance market.

8.4. *Private Information versus Adverse Selection*

There is a recent and growing literature that seeks to identify the impact of private information on the workings of insurance markets. Generally, this literature has searched for adverse selection, asking whether those who have more insurance have higher claims. Yet my theoretical and empirical results suggest that this approach is unable to identify private information precisely in cases where its impact is most severe: where the insurance market completely shuts down. This provides a new explanation for why previous literature has found mixed evidence of adverse selection and, in cases where adverse selection is found, estimated small welfare impacts (Cohen and Siegelman (2010), Einav, Finkelstein, and Levin (2010)).

Existing explanations for the frequent absence of adverse selection focus on preference heterogeneity (see Finkelstein and McGarry (2006) in LTC, Fang, Keane, and Silverman (2008) in Medigap, and Cutler, Finkelstein, and McGarry (2008) for a broader focus across five markets). At a high level, these papers suggest that in some contexts, the higher risk (e.g., the sick) may have a lower preference for insurance. Although this paper cannot directly shed more light on whether individuals who have different beliefs have different utility functions, u,[79] it is important to note that my results raise concerns about inferring that the sick have lower demand for insurance because they have lower ownership rates. Rather, one needs to consider the potential that the supply of insurance to the sick, especially those who have *observable* health conditions, is limited through rejections.[80] It may not be that the sick do not want insurance, but rather that the insurers do not want the sick.

[79]Future work could merge my empirical approach to identify beliefs with traditional revealed preference approaches to identify demand, thereby identifying the distribution of preferences for insurance conditional on beliefs and further exploring the role of preference heterogeneity in insurance markets.

[80]See footnote 8 for a discussion of Finkelstein and McGarry (2006) and preference heterogeneity in LTC insurance. For Medigap, Fang, Keane, and Silverman (2008) found evidence of advantageous selection based on observables: individuals who have observable health conditions are less likely to purchase Medigap insurance, despite having higher expected costs. However, their analysis does not address the potential that rejections by Medigap insurers drive the lower ownership among those who have observable health conditions. Although Medigap insurers are not allowed to reject applicants during a 6-month open enrollment period at the age of 65, beyond this grace period, rejections are allowed and are common industry practice in most states.

9. CONCLUSION

This paper argues that private information leads insurance companies to reject applicants who have certain observable, often high-risk, characteristics. My findings suggest that if insurance companies were to offer any contract or set of contracts to those who currently are rejected, they would be too adversely selected to yield a positive profit. More generally, the results suggest that 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

AKERLOF, G. (1970): "The Market for Lemons: Qualitative Uncertainty and the Market Mechanism," *Quarterly Journal of Economics*, 84 (3), 488–500. [1714,1716,1720]

AMERICAN COUNCIL OF LIFE INSURERS (2010): *Life Insurers Fact Book 2010*. Washington, D.C.: American Council of Life Insurers. [1731,1759]

ANDREWS, D. (2001): "Testing When a Parameter Is on the Boundary of the Maintained Hypothesis," *Econometrica*, 69, 683–734. [1752]

BLACKWELL, D. (1951): "Comparison of Experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley: University of California Press, 93–102. [1726]

——— (1953): "Equivalent Comparisons of Experiments," *The Annals of Mathematical Statistics*, 24, 265–272. [1726]

BOUND, J., J. B. CULLEN, A. NICHOLS, AND L. SCHMIDT (2004): "The Welfare Implications of Increasing Disability Insurance Benefit Generosity," *Journal of Public Economics*, 88 (12), 2487–2514. [1755,1756]

BROWN, J., AND A. FINKELSTEIN (2008): "The Interaction of Public and Private Insurance: Medicaid and the Long-Term Care Insurance Market," *American Economic Review*, 98, 1083–1102. [1755]

CAWLEY, J., AND T. PHILIPSON (1999): "An Empirical Examination of Information Barriers to Trade in Insurance," *American Economic Review*, 89 (4), 827–846. [1717,1731,1745]

CHADE, H., AND E. SCHLEE (2012): "Optimal Insurance With Adverse Selection," *Theoretical Economics*, 7 (3), 571–607. [1717,1718,1720]

CHIAPPORI, P., AND B. SALANIÉ (2000): "Testing for Asymmetric Information in Insurance Markets," *Journal of Political Economy*, 108 (1), 56–78. [1717]

CHIAPPORI, P., B. JULLIEN, B. SALANIÉ, AND F. SALANIÉ (2006): "Asymmetric Information in Insurance: General Testable Implications," *RAND Journal of Economics*, 37 (4), 783–798. [1717]

COHEN, A., AND P. SIEGELMAN (2010): "Testing for Adverse Selection in Insurance Markets," *The Journal of Risk and Insurance*, 77, 39–84. [1717,1760]

CONGRESSIONAL BUDGET OFFICE (2004): *Financing Long-Term Care for the Elderly*. Washington, D.C.: Congressional Budget Office. [1731]

CUTLER, D., A. FINKELSTEIN, AND K. MCGARRY (2008): "Preference Heterogeneity and Insurance Markets: Explaining a Puzzle of Insurance," *American Economic Review*, 98 (2), 157–162. [1760]

EFRON, B., AND G. GONG (1983): "A Leisurely Look at the Bootstrap, the Jackknife, and Cross-Validation," *The American Statistician*, 37 (1), 36–48. [1746,1753,1755,1758]

EINAV, L., A. FINKELSTEIN, AND J. LEVIN (2010): "Beyond Testing: Empirical Models of Insurance Markets," *Annual Review of Economics*, 2 (1), 311–336. [1717,1760]

EINAV, L., A. FINKELSTEIN, AND P. SCHRIMPF (2010): "Optimal Mandates and the Welfare Cost of Asymmetric Information: Evidence From the UK Annuity Market," *Econometrica*, 78 (3), 1031–1092. [1726]

FANG, H., M. KEANE, AND D. SILVERMAN (2008): "Sources of Advantageous Selection: Evidence From the Medigap Insurance Market," *Journal of Political Economy*, 116 (2), 303–350. [1760]

FINKELSTEIN, A., AND K. MCGARRY (2006): "Multiple Dimensions of Private Information: Evidence From the Long-Term Care Insurance Market," *American Economic Review*, 96 (4), 938–958. [1717,1732,1734,1745,1760]

FINKELSTEIN, A., AND J. POTERBA (2002): "Selection Effects in the United Kingdom Individual Annuities Market," *The Economic Journal*, 112 (476), 28–50. [1717,1758]

——— (2004): "Adverse Selection in Insurance Markets: Policyholder Evidence From the UK Annuity Market," *Journal of Political Economy*, 112 (1), 183–207. [1717,1758]

GAN, L., M. HURD, AND D. MCFADDEN (2005): "Individual Subjective Survival Curves," in *Analyses in the Economics of Aging*, ed. by D. Wise. Chicago: University of Chicago Press. [1725, 1732,1749]

GIGERENZER, G., AND U. HOFFRAGE (1995): "How to Improve Bayesian Reasoning Without Instruction: Frequency Formats," *Psychological Review*, 102, 684–704. [1725]

HE, D. (2009): "The Life Insurance Market: Asymmetric Information Revisited," *Journal of Public Economics*, 93 (9–10), 1090–1097. [1731,1734]

HENDREN, N. (2013): "Supplement to 'Private Information and Insurance Rejections'," *Econometrica Supplemental Material*, 81, http://www.econometricsociety.org/ecta/supmat/10931_proofs.pdf; http://www.econometricsociety.org/ecta/supmat/10931_data_and_programs.zip. [1717]

HURD, M. (2009): "Subjective Probabilities in Household Surveys," *Annual Review of Economics*, 1, 543–562. [1725]

KAHNEMAN, D., AND A. TVERSKY (1979): "Prospect Theory: An Analysis of Decision Under Risk," *Econometrica*, 47 (2), 263–292. [1726]

LOCKWOOD, L. (2012): "Incidental Bequests: Bequest Motives and the Choice to Self-Insure Late-Life Risks," Working Paper. [1755]

MAILATH, G. J., AND G. NOLDEKE (2008): "Does Competitive Pricing Cause Market Breakdown Under Extreme Adverse Selection?" *Journal of Economic Theory*, 140, 97–125. [1720]

MILLER, S., A. KIRLIK, A. KOSORUKOFF, AND J. TSAI (2008): "Supporting Joint Human–Computer Judgement Under Uncertainty," in *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society, 408–412. [1725]

MIYAZAKI, H. (1977): "The Rat Race and Internal Labor Markets," *The Bell Journal of Economics*, 8 (2), 394–418. [1720]

MURTAUGH, C., P. KEMPER, AND B. SPILLMAN (1995): "Risky Business: Long-Term Care Insurance Underwriting," *Inquiry*, 35 (3), 204–218. [1714]

NEWEY, W. K. (1997): "Convergence Rates and Asymptotic Normality for Series Estimators," *Journal of Econometrics*, 79 (1), 147–168. [1742]

RILEY, J. G. (1979): "Informational Equilibrium," *Econometrica*, 47 (2), 331–359. [1717,1718, 1720]

ROTHSCHILD, M., AND J. STIGLITZ (1976): "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics*, 90 (4), 629–649. [1714,1716,1718,1719]

SAVAGE, L. J. (1954): *The Foundations of Statistics*. New York: Wiley. [1725]

SPENCE, A. M. (1978): "Product Differentiation and Performance in Insurance Markets," *Journal of Public Economics*, 10 (3), 427–447. [1720]

Dept. of Economics, Harvard University, Cambridge, MA 02138, U.S.A. and NBER; nhendren@fas.harvard.edu.

Manuscript received July, 2012; final revision received January, 2013.