Essays on Information and Insurance Markets

by

Nathaniel Hendren

Submitted to the Department of Economics in partial fulfillment of the requirements for the degree of Doctor of Philosophy at the MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2012

© 2012 Nathaniel Hendren. All rights reserved.

The author hereby grants to MIT permission to reproduce and distribute publicly paper and electronic copies of this thesis document in whole or in part.

Author

Department of Economics

May 15, 2012

Certified by

Daron Acemoglu
Elizabeth and James Killian Professor of Economics
Thesis Supervisor

Certified by

Amy Finkelstein
Professor of Economics
Thesis Supervisor

Accepted by

Michael Greenstone
3M Professor of Economics
Chairman, Departmental Committee on Graduate Studies
Essays on Information and Insurance Markets

by

Nathaniel Hendren

Submitted to the Department of Economics
on May 15, 2012, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis studies the impact of private information on the existence of insurance markets. In the first chapter, I study the case of insurance rejections. Across a wide set of non-group insurance markets, applicants are rejected based on observable, often high-risk, characteristics. I explore private information as a potential cause by developing and testing a model in which agents have private information about their risk. I derive a new no-trade result that can theoretically explain how private information could cause rejections. I use the no-trade condition to generate measures of the barrier to trade private information imposes. I develop a new empirical methodology to estimate these measures that uses subjective probability elicitations as noisy measures of agents’ beliefs. I apply the approach to three non-group markets: long-term care (LTC), disability, and life insurance. Consistent with the predictions of the theory, in all three settings I find significant evidence of private information for those who would be rejected; I find that they have more private information than those who can purchase insurance; and I find that it is enough to cause a complete absence of trade. This presents the first empirical evidence that private information leads to a complete absence of trade.

In the second chapter, I show that private information explains the absence of a private unemployment insurance market. I provide the empirical evidence that a private UI market would be afflicted by private information and suggest the amount of private information is sufficient to explain a complete absence of trade. I present evidence a private market would still not arise even if the government stopped providing unemployment benefits.

Finally, in the third chapter I use the empirical and theoretical tools developed in the first chapter to explore the impact of an adjusted community rating policy that would force insurance companies to only price based on age. My results suggest such a policy would completely unravel the LTC insurance market. Not only would welfare not be improved for those who are currently rejected, but the regulation would prevent the healthy from being able to purchase long-term care insurance.

Thesis Supervisor: Daron Acemoglu
Title: Elizabeth and James Killian Professor of Economics

Thesis Supervisor: Amy Finkelstein
Title: Professor of Economics
Acknowledgments

I am indebted to many people for this thesis, and it is certainly not possible to thank them all. First and foremost, I thank my advisors, Daron Acemoglu and Amy Finkelstein. You pushed me to think harder than I ever thought I could, and for that I will be forever grateful. I also want to thank Jon Gruber and Robert Townsend, who were also very involved throughout the construction of this thesis and provided invaluable feedback and advice. I have also benefited from a host of other faculty members at MIT who comprise a list too long to mention. I also wish to thank Timothy Conley, who responded to an email from a naive 2nd year undergraduate at the University of Chicago looking for an RA job and was the first to expose me to economic research.

I also wish to thank my friends and fellow graduate students. This list is again too long to mention, but Emily Breza, Gabriel Carroll, Ashley Swanson, and Joe Shapiro made the past five years so incredibly enjoyable. You’re not only friends, but sometimes even co-authors (Ashley, that joke is dedicated to you!).

I was fortunate to spend a large fraction of the past five years on research. For this, I’d like to acknowledge financial support of the National Science Foundation Graduate Research Fellowship and the National Bureau of Economic Research Health and Aging Fellowship, funded under the National Institute of Aging Grant Number T32-AG000186.

I’d like to thank Sarah Miller for her love, friendship, and quick-witted humor that always keeps life in perspective. And finally, I’d like to thank my family: my parents, Sarah and Steve, and my brother, Christopher, who have always supported me to pursue my dreams.
## Contents

1 Private Information and Insurance Rejections

1.1 Introduction ......................................................... 13
1.2 Theory ............................................................ 16
   1.2.1 Environment ................................................... 16
   1.2.2 The No-Trade condition ................................... 17
1.3 Comparative Statics and Testable Predictions .................. 20
   1.3.1 Two Measures of Private Information ..................... 20
   1.3.2 High-Risk Distributions .................................. 22
   1.3.3 Moving Towards Data: Testable Hypotheses ............. 22
1.4 Empirical Methodology ........................................... 23
   1.4.1 Nonparametric Lower Bound Approach .................... 24
   1.4.2 Semiparametric Approach: Estimation of the Distribution of Private Information ........................................ 27
1.5 Setting and Data .................................................. 29
   1.5.1 The Three Non-Group Market Settings .................... 29
   1.5.2 Data ......................................................... 30
     1.5.2.1 Loss Variables and Subjective Probability Elicitations .................... 31
     1.5.2.2 Rejection Classification ................................ 33
     1.5.2.3 Public Information ..................................... 35
     1.5.2.4 Sample Selection ...................................... 38
     1.5.2.5 Relation to Ideal Data ................................ 41
1.6 Lower Bound Estimation .......................................... 41
   1.6.1 Empirical Estimation ...................................... 41
   1.6.2 Statistical Inference ...................................... 42
   1.6.3 Graphical Results for \( P_2 - E[P_2|X] \) .................. 43
   1.6.4 Lower Bound Results ...................................... 43
   1.6.5 Subsample Analysis ...................................... 47
   1.6.6 Robustness .................................................. 49
     1.6.6.1 Insurance Ownership Sample Selection ............... 49
1.6.6.2 Organ Controls for Life Specification ........................................... 50
1.6.7 Summary ................................................................................. 50
1.7 Estimation of Distribution of Private Information ......................... 51
1.7.1 Empirical Specification .......................................................... 51
1.7.1.1 Elicitation Error Structure ...................................................... 51
1.7.1.2 Flexible Approximation for the Distribution of Private Information 52
1.7.1.3 Pooled Price Ratio (and its Minimum) ..................................... 54
1.7.2 Estimation Results ................................................................. 54
1.7.2.1 Graphs of the Distribution of Private Information .................. 54
1.7.2.2 Minimum Pooled Price Ratio ................................................ 54
1.7.2.3 Results for Elicitation Error Distribution ............................... 61
1.7.2.4 Annuities ......................................................................... 61
1.8 Conclusion ................................................................................. 63
1.A Theory Appendix ......................................................................... 64
1.A.1 Proof of No-Trade Theorem ...................................................... 64
1.A.1.1 Utility Space .......................................................................... 64
1.A.1.2 Converse ............................................................................. 65
1.A.1.3 Lemmas ............................................................................. 67
1.A.1.4 Finite Types .......................................................................... 68
1.A.1.5 Arbitrary Distribution ................................................................ 70
1.A.2 Informal Insurance ................................................................... 73
1.B Empirical Methodology Appendix ............................................... 75
1.B.1 Properties of the Lower Bound Estimator .................................... 75
1.B.1.1 Proof of Proposition ................................................................ 77
1.B.1.2 "Tight" lower bound ............................................................... 78
1.B.1.3 Measurement Error Example .................................................. 79
1.B.2 Semiparametric Identification ................................................... 79
1.B.2.1 Identification without censoring .............................................. 82
1.B.2.2 Monte Carlo Results ............................................................... 83
1.C Rejections Appendix: Selected Pages from Genworth Financial Underwriting Guidelines ................. 84
2 Private Information and Unemployment Insurance ............................. 89
2.1 Introduction ................................................................................. 89
2.2 Theory ......................................................................................... 92
2.2.1 Setup ...................................................................................... 92
2.2.2 Implementable Allocations ....................................................... 93
2.2.3 Concavity Assumption ............................................................. 94
## List of Tables

1.1 Rejection Classification .......................................................... 34
1.2 Covariates ................................................................................. 37
1.3 Sample Selection ....................................................................... 39
1.4 Magnitude of Private Information (Lower Bound) ....................... 45
1.5 Robustness Checks: Sample Selection ........................................ 50
1.6 Robustness Checks: Organ Controls (Life Setting) ...................... 51
1.7 Minimum Pooled Price Ratio ...................................................... 56
1.8 Minimum Pooled Price Ratio (Robustness to $\tau$) .................... 57
1.9 Willingness to Pay Calibration .................................................... 59
1.10 Elicitation Error Parameters .................................................... 62
1.11 Minimum Pooled Price Ratio (Annuities) ................................... 63

2.1 Covariate Specification ............................................................... 102
2.2 Sample Summary Statistics ......................................................... 102
2.3 Presence of Private Information .................................................. 105
2.4 Minimum Pooled Price Ratio, $\inf T(p)$ ...................................... 108
2.5 Robustness to Choice of $\tau$ ....................................................... 108
2.6 Willingness to Pay ................................................................. 109
2.7 Elicitation Error Parameters ....................................................... 110

3.1 Sample Summary Statistics ......................................................... 127
3.2 Presence of Private Information .................................................. 129
3.3 Minimum Pooled Price Ratio, $\inf T(p)$ ...................................... 131
3.4 Robustness to Choice of $\tau$ ....................................................... 131
3.5 Elicitation Error Parameters ....................................................... 132
## List of Figures

1-1 Subjective Probability Histograms ........................................ 32  
1-2 Distribution of Residual Private Information .......................... 44  
1-3 Magnitude of Private Information (Lower Bound) ...................... 48  
1-4 PDF of Private Information ............................................... 55  
1-5 CDF of Private Information ............................................... 55  
1-6 Pooled Price Ratio ......................................................... 60  
1-7 Monte Carlo Simulations ................................................... 83  

2-1 Histogram of Subj. Prob. (Job Loss in 1 Year) .......................... 101  
2-2 CDF of Private Information ................................................ 108  

3-1 Subjective Probability Histogram (Nursing Home Entry in 5 Years) .... 128  
3-2 CDF of Private Information ................................................ 131
Chapter 1

Private Information and Insurance Rejections

1.1 Introduction

Not everyone can purchase insurance. Across a wide set of non-group insurance markets, companies choose to not sell insurance to potential customers with certain observable, often high-risk, characteristics. In the non-group health insurance market, 1 in 7 applications to the four largest insurance companies in the United States were rejected between 2007 and 2009, a figure that excludes those who would be rejected but were deterred from even applying.\(^1\) In US long-term care insurance, 12-23% of 65 year olds have health conditions that would preclude them from being able to purchase insurance (Murtaugh et al. (1995)).\(^2\)

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

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

---

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

\(^2\)Appendix 1.C presents the rejection conditions from Genworth Financial (one of the largest US LTC insurers), gathered from their underwriting guidelines provided to insurance agents for use in screening applicants.

\(^3\)For example, in long-term care we estimate those who would be rejected have an average five-year nursing home entry rate of less than 20%.
hold across market structures such as monopoly or competition.

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

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

We then develop a new empirical methodology for studying private information to test the predictions of theory. We use information contained in subjective probability elicitations to infer properties of the distribution of private information. At no point do we view these elicitations as true beliefs. Rather, we use information in the joint distribution of elicitations and the realized events corresponding to these elicitations to deal with potential errors in elicitations. We proceed with two complementary approaches. First, we make the weak assumption that agent's elicitations cannot contain more information about the subsequent loss than would the true beliefs. We estimate the explanatory power of the subjective probabilities on the subsequent realized event, conditional on public information. This allows us to generate nonparametric lower bounds on a measure of the magnitude of private information provided by the theory. With these bounds, we provide a simple test for the presence of private information, along with a test of whether those who would be rejected have larger estimates of this lower bound.

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

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

5If beliefs are generated through rational expectations given some information set, this assumption is equivalent to assuming the elicitations are a garbling of the agent's true beliefs in the sense of Blackwell (1951, 1953).
We apply our approach to three non-group markets: long-term care (LTC), disability, and life insurance. We combine two sources of data. First, we use data from the Health and Retirement Study, which elicits subjective probabilities corresponding to losses insured in each of these three settings and contains a rich set of demographic and health information commonly used by insurance companies in pricing insurance. We supplement this with a detailed review of underwriting guidelines from major insurance companies to identify those who would be rejected (henceforth “rejectees”) in each market.

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

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

Empirically, our paper is related to a recent and growing literature on testing for the existence and consequences of private information in insurance markets (Chiappori and Salanié (2000); Chiappori et al. (2006); Finkelstein and Poterba (2002, 2004); see Einav et al. (2010) and Cohen and Siegelman (2010) for a review). This literature focuses on the revealed preference implications of

6 Throughout, we focus on those who “would be rejected”, which corresponds to those whose choice set excludes insurance, not necessarily the same as those who actually apply and are rejected.
private information by looking for a correlation between insurance purchase and subsequent claims. This approach can only identify private information amongst those served by the market. In contrast, our approach can study private information for the entire population, including rejectees. Our results suggest significant amounts of private information for the rejectees, but less for those served by the market. Thus, our results provide a new explanation for why previous studies using the revealed preference approach have not found evidence of significant adverse selection in life insurance (Cawley and Philipson (1999)) and LTC insurance (Finkelstein and McGarry (2006)). The absence of adverse selection may be the insurer's selection.

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

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

1.2 Theory

This section develops a model of private information. Our primary result (Theorem 1.1) is a no-trade condition which provides a theory of how private information can lead insurance companies to not offer any contracts.

1.2.1 Environment

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

\(^7\)By choosing particular distributions \( F(p) \), our environment nests many previous models of insurance. For example, \( \Psi = \{p_L, p_H \} \) yields the classic two-type model considered initially by Rothschild and Stiglitz (1976) and subsequently analyzed by many others. Assuming \( F(p) \) is continuous with \( \Psi = [a, b] \subset (0, 1) \), one obtains an environment similar to Riley (1979). Chade and Schlee (2011) provide arguably the most general treatment to-date of this environment in the existing literature by considering a monopolists problem with an arbitrary \( F \) with bounded support \( \Psi \subset [a,b] \subset (0,1) \).
that we have conditioned on observable information (e.g. \( F(p) = F(p|X = x) \) where \( X \) includes all observable characteristics such as age, gender, and observable health conditions).

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

\[
p u(c_L) + (1 - p) u(c_{NL})
\]

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

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

**Definition 1.1.** An allocation \( A = \{c_L(p), c_{NL}(p)\} \) is implementable if

1. \( A \) is resource feasible:
\[
\int [w - pl - pc_L(p) - (1 - p)c_{NL}(p)] dF(p) > 0
\]

2. \( A \) is incentive compatible:
\[
p u(c_L(p)) + (1 - p) u(c_{NL}(p)) \geq p u(c_L(\bar{p})) + (1 - p) u(c_{NL}(\bar{p})) \quad \forall p, \bar{p} \in \Psi
\]

3. \( A \) is individually rational:
\[
p u(c_L(p)) + (1 - p) u(c_{NL}(p)) \geq p u(w - l) + (1 - p) u(w) \quad \forall p \in \Psi
\]

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

### 1.2.2 The No-Trade condition

Theorem 1.1 characterizes when the endowment is the only implementable allocation.

**Theorem 1.1.** (No Trade). The endowment, \( \{(w - l, w)\} \), is the only implementable allocation if and only if

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

where \( \Psi \setminus \{1\} \) denotes the support of \( F(p) \) excluding the point \( p = 1 \).
Conversely, if (1.1) does not hold, then there exists an implementable allocation which strictly satisfies resource feasibility and individual rationality for a positive mass of types.

Proof. See Appendix 1.A.1.

The left-hand side of equation (1.1), \[ \frac{P}{1-p} \frac{u'(w-l)}{u'(w)} \] is the marginal rate of substitution between consumption in the event of no loss and consumption in the event of a loss, evaluated at the endowment, \((w-l, w)\). It is a type \(p\) agent’s willingness to pay for an infinitesimal amount of additional consumption in the event of a loss, in terms of consumption in the event of no loss. The actuarially fair cost of this transfer to the type \(p\) agent is \(\frac{P}{1-p}\). However, the right hand side of equation (1.1) is the price of providing such a transfer, not at type \(p\)’s own cost of \(\frac{P}{1-p}\), but rather at the average cost if all higher-risk types \(P > p\) also obtained this transfer, \(\frac{E[P|P > p]}{1-E[P|P > p]}\). Intuitively, if no other contracts are offered, then a contract preferred by type \(p\) will also be preferred by all types \(P > p\), rendering the cheapest possible provision of insurance to type \(p\) to be at a price ratio of \(\frac{E[P|P > p]}{1-E[P|P > p]}\). If no agent is willing to pay this cost, the endowment is the only implementable allocation.

Conversely, if equation (1.1) does not hold, there exists an implementable allocation which does not totally exhaust resources and provides strictly higher utility than the endowment for a positive mass of types. So, a monopolist insurer could earn positive profits by facilitating trade. In this sense, the no-trade condition (1.1) characterizes when one would expect trade to occur.

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

Example 1.1. Suppose that \(F(p)\) is uniform, \(F(p) = p\). Then, \(E[P|P > p] = \frac{1+p}{2}\). The no-trade condition 1.1 is given by
\[
\frac{p}{1-p} \frac{u'(w-l)}{u'(w)} \leq \frac{1+p}{2} \frac{1}{1-\frac{1+p}{2}} \quad \forall p \in [0, 1)
\]
which holds if and only if
\[
\frac{u'(w-l)}{u'(w)} \leq 2
\]

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

The no-trade condition has an unraveling intuition similar to that of Akerlof (1970). His model considers a given contract and shows that it will not be traded when its demand curve lies everywhere below its average cost curve, which is in turn a function of those who demand it. Our

---

8While Theorem 1.1 is straightforward, its proof is less trivial because one must show that Condition 1.1 rules out not only single contracts but also any menu of contracts in which different types may receive different allocations.

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

10We discuss this tax rate analogy further in Section 1.3.
model is different. While Akerlof (1970) derives conditions under which a given contract would unravel and result in no trade, our model provides conditions under which any contract or menu of contracts would unravel.

This distinction is important since previous literature has argued that trade must always occur environments similar to ours with no restrictions on the contract space (Riley (1979); Chade and Schlee (2011)). The key difference in our approach is that we do not assume types are bounded away from 1.11 In fact, the no-trade condition requires the highest risk type in the economy have a probability of a loss arbitrarily close to \( p = 1 \). Otherwise the highest risk type, say \( \bar{p} \), would be able to obtain an actuarially fair full insurance allocation, \( c_L(\bar{p}) = c_{NL}(\bar{p}) = w - \bar{p}l \), which would not violate the incentive constraints of any other type.

**Corollary 1.1.** Suppose condition (1.1) holds. Then \( F(p) < 1 \forall p < 1 \).

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

At the same time, the fact that the no-trade condition requires risks arbitrarily close to 1 can be viewed as a technicality. In reality, insurance companies offer a finite set of contracts, presumably because they incur a setup cost for creating each contract. If we require that each allocation other than the endowment must attract a non-trivial fraction of types, then we no longer require risks arbitrarily close to 1, as illustrated in Remark 1.1.

**Remark 1.1.** Suppose each consumption bundle \((c_L, c_{NL})\) other than the endowment must attract a non-trivial fraction \( \alpha > 0 \) of types. More precisely, suppose allocations \( A = \{c_L(p), c_{NL}(p)\}_p \) must have the property that for all \( q \in \Psi \),

\[
\mu(\{p| (c_L(p), c_{NL}(p)) = (c_L(q), c_{NL}(q))\}) \geq \alpha
\]

where \( \mu \) is the measure defined by \( F(p) \). Then, the no-trade condition is given by

\[
\frac{p}{1 - p} \cdot \frac{u'(w - l)}{u'(w)} \leq \frac{E[P|P \geq p]}{1 - E[P|P \geq p]} \forall p \in \Psi_{1-\alpha}
\]

where \( \Psi_{1-\alpha} = [0, F^{-1}(1 - \alpha)] \cap (\Psi \setminus \{1\}) \).13 Therefore, the no-trade condition need only hold for values \( p < F^{-1}(1 - \alpha) \).

In other words, if contracts must attract a nontrivial fraction of types, then no trade can occur even if types are bounded away from \( p = 1 \). Going forward, we retain the benchmark assumption of no such frictions or transactions costs, but return to this discussion in our empirical work in Section 1.7.

---

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

12 Note that we do not require any positive mass at \( p = 1 \), as highlighted in Example 1.1.

13 If \( F^{-1}(1 - \alpha) \) is a set, we take \( F^{-1}(1 - \alpha) \) to be the supremum of this set.
The no-trade condition (1.1) provides a theory of rejections: they occur in market segments where (1.1) holds and insurance is offered in segments where (1.1) does not hold, where market segments are defined by observable information. In order to derive testable implications of this theory, the next section examines properties of distributions, $F(p)$, which make the no-trade condition more likely to hold.

### 1.3 Comparative Statics and Testable Predictions

Qualitatively, Theorem 1.1 suggests a property of distributions which lead to no trade: thick upper tails of risks. The presence of a thicker upper tail increases the value of $E[P|P \geq p]$ at given values of $p$. In this section, we formalize this intuition by constructing precise measures of the barrier to trade imposed by private information which will guide our empirical tests of the theory.

#### 1.3.1 Two Measures of Private Information

We construct two measures of private information. To begin, we multiply the no-trade condition (1.1) 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 the agents’ marginal utilities in the loss versus no loss state, evaluated at the endowment. The right-hand side independent of the utility function, $u$, and is the cost of providing an infinitesimal transfer to type $p$ if the pool of types worse than $p$, $P > p$, also were attracted to the contract. We define this term the pooled price ratio.

**Definition 1.2.** For any $p \in \Psi \setminus \{1\}$, the **pooled price ratio** at $p$, $T(p)$, is given by

$$T(p) = \frac{E[P|P \geq p]}{1 - E[P|P \geq p]} \frac{1-p}{p} \quad (1.2)$$

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

**Corollary 1.2.** *(Quantification of the barrier to trade)* The no-trade condition holds if and only if

$$\frac{u'(w-l)}{u'(w)} \leq \inf_{p \in \Psi \setminus \{1\}} T(p) \quad (1.3)$$

Whether or not there will be trade depends on only two numbers: the agent’s underlying valuation of insurance, $\frac{u'(W-L)}{u'(W)}$, and the cheapest cost of providing an infinitesimal amount of insurance, $\inf_{p \in \Psi \setminus \{1\}} T(p)$. When this cost is above the underlying valuation of insurance, there can be no trade. We call $\inf_{p \in \Psi \setminus \{1\}} T(p)$ the **minimum pooled price ratio**. This number characterizes the barrier to trade imposed by private information.
Equation (1.3) has a simple tax rate interpretation. Suppose for a moment that there were no private information but instead a government levies a sales tax of rate \( t \) on insurance premiums in a competitive insurance market. The value \( \frac{u'(w-l)}{u'(w)} - 1 \) is the highest such tax rate an individual would be willing to pay to purchase any insurance.\(^{14}\) Thus, \( \inf_{p \in \Psi \setminus \{1\}} T(p) - 1 \) is the tax rate equivalent of the barrier to trade imposed by private information. In this sense, it quantifies the magnitude of the barrier to trade imposed by private information.

Equation (1.3) leads to a simple comparative static.

**Corollary 1.3.** (Comparative static in the minimum pooled price ratio) Consider two market segments with pooled price ratios \( T_1(p) \) and \( T_2(p) \) and common vNM preferences \( u \). Suppose

\[
\inf_{p \in \Psi \setminus \{1\}} T_1(p) \leq \inf_{p \in \Psi \setminus \{1\}} T_2(p)
\]

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 1.3 is the key comparative static on the distribution of private information provided by the theory.\(^{15}\)

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

**Definition 1.3.** For any \( p \in \Psi \), define the **magnitude of private information at** \( p \) by \( m(p) \), given by

\[
m(p) = E[P|P \geq p] - p
\]

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 (1.1).

**Corollary 1.4.** (Comparative static in the magnitude of private information) Consider two market segments with magnitudes of private information \( m_1(p) \) and \( m_2(p) \) and common support \( \Psi \) and common vNM preferences \( u \). Suppose

\[
m_1(p) \leq m_2(p) \quad \forall p \in \Psi
\]

\(^{14}\)To clarify, the equivalence is to a tax rate paid only in the state of no loss, so that it can be interpreted as a tax on the insurance premium.

\(^{15}\)Corollaries 1.2 and 1.3 do require that the willingness to pay, \( \frac{u'(w-l)}{u'(w)} \), does not vary with \( p \). This would be violated if agents with different values of \( p \) made different informal insurance decisions, such as larger savings. We consider this case in detail in Appendix 1.A.2. In short, \( \frac{u'(w-l)}{u'(w)} \) may vary with \( p \), motivating a comparative static in \( T(p) \) for all \( p \), as opposed to the minimum pooled price ratio. Our empirical results are quite robust to this more restrictive test.
Then if the no-trade condition holds in segment 1, it must also hold in segment 2.

Higher values of the magnitude of private information are more likely to lead to no trade. Notice that the values of \( m(p) \) must be ordered for all \( p \in \Psi \), and it is thus a less precise statement than the comparative static provided in Corollary 1.3.

1.3.2 High-Risk Distributions

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

Proposition 1.1. Suppose \( P_1 \) and \( P_2 \) are ordered according to the hazard rate ordering. Let \( T_1 \) and \( T_2 \) denote their associated pooled price ratios. Then

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

for any \( \Psi \subset \Psi \setminus \{1\} \). In particular, (1.5) holds for \( \Psi = \Psi \setminus \{1\} \) or \( \Psi = \Psi_{1-\alpha} \) as defined in Remark 1.1.

Proof. Follows immediately from the fact that the hazard rate ordering implies the mean-residual life ordering. See Shaked and Shanthikumar (1994). \( \square \)

When distributions can be ordered according to their hazard rates, the higher mean risk distribution has a larger minimum pooled price ratio.\(^{16}\) Therefore, it satisfies the no-trade condition for a larger set of values of \( \frac{u'(w-l)}{u'(w)} \). In this sense, higher risk distributions are more likely to lead to no trade, which can explain why it is so often those with high (mean) risk characteristics who are rejected.

1.3.3 Moving Towards Data: Testable Hypotheses

Our goal of the rest of the paper is to test the empirical predictions of the theory by estimating properties of the distribution of private information, \( F(p|X) \), for rejectees and non-rejectees. Assuming for the moment that \( F(p|X) \) is observable to the econometrician, our ideal tests are as follows. Qualitatively, we test whether \( F(p|X) \) has a thicker upper tail of high risks for the rejectees. Quantitatively, we estimate the minimum pooled price ratio for each \( X \) and conduct

\(^{16}\)Note that the hazard rate ordering is weaker than the likelihood ratio ordering. So if distributions can be ordered according to their likelihood ratios (e.g. they have the monotone likelihood ratio property, "MLRP"), then higher mean risk distributions lead to larger minimum pooled price ratios.
two types of tests: first, we test the comparative statics given by Corollaries 1.3 and 1.4 of higher values of the minimum pooled price ratio for rejectees versus non-rejectees. Second, we ask whether the minimum pooled price ratio is large (small) enough to explain (the absence of) rejections for plausible values of agents’ willingness to pay, as suggested by Corollary 1.2.\footnote{Our tests do not focus on potential demand side variation across values of X (i.e. how willingness-to-pay, \( u'(w-x) \) varies with X and potentially differs across rejectees and non-rejectees). Finding empirical support for our comparative static tests would only be inconsistent with the theory if the difference in willingness-to-pay for rejectees versus non-rejectees is larger than our estimated differences in the minimum pooled price ratio. In contrast, if rejectees have lower willingness-to-pay than non-rejectees, our tests are too strict: they may lead us to find evidence inconsistent with the theory when in fact the theory is correct.}

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

\section*{1.4 Empirical Methodology}

We develop an empirical methodology to study private information and operationalize the tests in Section 1.3.3. The key feature of our approach is that we utilize information contained in subjective probability assessments to infer properties of the distribution of private information. Let \( L \) denote an event (e.g. dying in the next 10 years) that is commonly insured in some insurance market (e.g. life insurance).\footnote{Of course, individuals face more than a single binary event and insurance generally insures a combination of many different events. Our approach is to focus on one commonly insured event and ask whether the pattern of rejections in that market is consistent with the predictions of our theory about whether insurance could be provided for that binary event.} Let \( Z \) denote an individual's subjective probability elicitation about event \( L \) (i.e. \( Z \) is a response to the question “what do you think is the probability that \( L \) will occur?”). A premise of our approach is that these elicitation are non-verifiable to an insurance company. Therefore, they can be excluded from the set of public information, which we will denote by \( X \), and used to infer properties of the distribution of private information. But while these elicitation are non-verifiable to insurance companies, they are arguably noisy and potentially biased measures of true beliefs.

We develop two complementary approaches for dealing with the potential error in subjective probability elicitation. Our first approach provides a nonparametric lower bound on the average magnitude of private information, \( E[m(P)] \), and tests whether rejection segments have higher values of \( E[m(P)] \). This provides a test in the spirit of the comparative static in \( m(p) \) (Corollary 1.4) while relying on very minimal assumptions on the relationship between agents' beliefs and their probability elicitation. Our second approach adds a parametric structure to the distribution of elicitation error, which allows us to (non-parametrically) identify the distribution of private information (so that the overall approach is semiparametric). We then estimate the pooled price ratio, \( T(p) \), and a close analogue to the minimum pooled price ratio, \( \inf_{p \in \Psi \setminus \{1\}} T(p) \), where we focus on the minimum over a compact set \( \Psi \) which excludes points in the upper quantiles of \( F(p) \)
to avoid problems associated with extreme value estimation. We then test both whether segments facing rejection have larger values of the minimum pooled price ratio (Corollary 1.3) and whether these estimates are large (small) to explain (the absence of) rejections for plausible values of \( u'(w-l) / u'(w) \), as suggested by Corollary 1.2.

In this section, we introduce these empirical approaches. We defer a discussion of the empirical specification and statistical inference to Sections 1.6 and 1.7, after we have discussed our data and settings.

1.4.1 Nonparametric Lower Bound Approach

To begin, we retain the assumption from the theoretical section that agents act as if they have beliefs about the probability of the loss \( L \).\(^{19}\) Moreover, as has heretofore been implicit, we assume these beliefs are correct.

Beliefs \( P \) are correct: \( \Pr\{L|X, P\} = P \)

Assumption 1.4.1 states that if we hypothetically gathered a large group of individuals with the same observable values \( X \) and the same beliefs \( P \) and then observed whether or not they experience the loss \( L \), we would find that, on average, a fraction \( P \) of this group experiences the loss. As an empirical assumption, it is relatively strong, but it provides perhaps the simplest link between the realized loss \( L \) and beliefs.\(^{20}\) Note that we have now introduced public information, \( X \). To most closely match the theory, we assume \( X \) is the set of information that an insurance company would use to price insurance. We discuss this important data requirement further in Section 1.5.

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

\( Z \) contains no additional information than \( P \) about the loss \( L \), so that \( \Pr\{L|X, P, Z\} = \Pr\{L|X, P\} \)

Assumption 1.4.1 is very weak; it would be violated only if people could provide elicitation which are informative about \( L \) even conditional on the true beliefs of those making the reports.\(^{21}\)

---

\(^{19}\)Our approach therefore follows the view of personal probability expressed in the seminal work of Savage (1954): Although agents may not perfectly express their beliefs through survey elicitation, they would behave consistently in response to gambles over \( L \) ("consistently" in the sense of Savage's axioms).

\(^{20}\)This is a common assumption made, either implicitly or explicitly, in existing (revealed preference) approaches to studying private information (e.g. Einav et al. (2010)). We find some motivation for correct beliefs and our treatment of subjective probability elicitation in existing empirical work in the forecasting literature spanning economics, psychology, and engineering. Broadly, this literature suggests survey elicitation suffer significant limitations as measures of beliefs, but implicit forecasts based on behavior, as in prediction markets, tend to be more accurate (for an overview, see Sunstein (2006) and (Arrow et al. 2008)). Examples of the limitations of survey elicitation of beliefs include Gan et al. (2005) who consider the subjective mortality probabilities we use in this paper. Additional examples in psychology and cognitive engineering shows that simple improvements in elicitation methods can substantially improve forecasts by reducing elicitation biases (Miller et al. (2008), Gigerenzer and Hoffrage (1995)).

\(^{21}\)Assumptions 1.4.1 and 1.4.1 are jointly implied by a rational expectations model in which agents know both \( X \)
For the empirical tests, we classify segments $X$ into those in which insurance companies do and do not sell insurance, $X \in \Theta^{NoReject}$ and $X \in \Theta^{Reject}$. We then proceed as follows. First, we form the predicted value of $L$ given the observable variables $X$ and $Z$,

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

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

We then measure the extent to which $Z$ explains $L$ conditional on $X$ using a measure of dispersion inspired by the theory. Recall from Definition 1.3 that $m(p)$ in segment $X$ is given by $m(p) = E[P|P \geq p, X] - p$. We construct an analogue with $P_Z$,

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

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

In the spirit of the comparative statics given by Corollary 1.4, we test whether rejectees have higher values of $E[m_Z(P_Z)|X]$:

$$\Delta_Z = E[ m_Z(P_Z) | X \in \Theta^{Reject} ] - E[ m_Z(P_Z) | X \in \Theta^{NoReject} ] > 0 \quad (1.6)$$

which asks whether segments in which insurance companies have chosen to not sell insurance have higher average magnitudes of private information implied by $Z$ than segments in which they sell insurance. Stated more loosely, equation (1.6) asks whether the subjective probabilities of the rejectees better explain the realized losses than the non-rejectees, where “better explain” is measured using $E[m_Z(P_Z)|X]$. Equation (1.6) is the key empirical test provided by the lower bound approach.

and $Z$ in formulating their beliefs $P$. In this case, our approach views $Z$ as a “garbling” of the agent’s true beliefs in the sense of Blackwell ((1951), (1953)).

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

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

24 In addition to aggregating the data across all $X$ in each rejection classification, we also conduct the test for subgroups (e.g. conditional on age or gender).
Lower Bounds  Our estimable variable, $P_Z$, is not equal to the true beliefs, $P$. Rather we obtain distributional lower bounds, as illustrated in Proposition 1.2.

**Proposition 1.2.** (Lower bound) Suppose assumptions 1 and 2 hold. Then

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

   $$P_Z = E[P|X, Z]$$  \hspace{1cm} (1.7)

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

   $$E[m_Z(P_Z)|X] \leq E[m(P)|X]$$  \hspace{1cm} (1.8)

**Proof.** See Appendix 1.B.1. \hfill \Box

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

Statement (2) highlights that testing $E[m_Z(P_Z)|X] = 0$ provides a nonparametric test for the presence of private information (Note $E[m_Z(P_Z)|X] > 0$ implies $E[m(P)|X] > 0$). Since $E[m_Z(P_Z)|X] = 0$ if and only if $\Pr\{L|X, Z\} = \Pr\{L|X\}$, the test for private information is straightforward: do the subjective probabilities explain the realized loss?\(^25\)

Our approach is nonparametric in the sense that we have made no parametric restrictions on how the elicitations $Z$ relate to the true beliefs $P$.\(^26\) For example, $P_Z$ and $m_Z(p)$ are invariant to monotonic transformations in $Z$: $P_Z = P_{h(Z)}$ and $m_Z(p) = m_{h(Z)}(p)$ for any monotonic function $h$. Thus, we do not require that $Z$ be a probability or have any cardinal interpretation. Respondents could all change their elicitations to $1-Z$ or $100Z$; this would not change the value of $E[m_Z(P_Z)|X]$.

But while the benefit of the lower bound approach is that we make only minimal assumptions on how subjective probabilities relate to true beliefs, the resulting empirical test in equation (1.6) suffers several limitations. First, orderings of lower bounds of $E[m(P)|X]$ across segments do not...\(^25\)Our test for the presence of private information is different from the test used by Finkelstein and McGarry (2006) that was initially proposed in Finkelstein and Poterba (2006). Their approach treats subjective probabilities as “unused observables” that are excluded from the set of variables used by insurance companies for pricing insurance. They infer the presence of asymmetric information if two conditions are satisfied: 1) the subjective probabilities are correlated with the realized loss (conditional on observables) and 2) the subjective probabilities are correlated with insurance purchase (conditional on observables). In contrast, we show that the second requirement is not necessary when using subjective probabilities for identifying private information. Indeed, it would prevent identification of private information amongst rejectees.

\(^26\)In particular, we have not imposed parametric restrictions on the distribution of $Z$ given beliefs $P$, $f_{Z|P}(Z|P)$.\(^26\)
necessarily imply orderings of its true magnitude. Second, orderings of $E[m(P)|X]$ does not imply orderings of $m(p)$ for all $p$, which was the statement of the comparative static in $m(p)$ in Corollary (1.4). Finally, in addition to having limitations as a test of the comparative static, this approach cannot quantify the minimum pooled price ratio. These shortcomings motivate our second approach, which imposes some structure on the relationship between $Z$ and $P$ and allows us to move from lower bounds to point estimates of the distribution of private information.

1.4.2 Semiparametric Approach: Estimation of the Distribution of Private Information

The goal of the second approach is to estimate the distribution of private information and the minimum pooled price ratio. We then examine the distribution for the presence of thicker upper tails for the rejectees relative to the non-rejectees. With the minimum pooled price ratio, we test whether it is larger for rejectees versus non-rejectees (Corollary 1.3) and whether it is large (or small) enough to explain (the absence of) rejections for plausible values of the willingness to pay for insurance (Corollary 1.2). Whereas our nonparametric lower bound approach allowed for an arbitrary relationship between $Z$ and $P$, we now restrict the way in which elicitations relate to beliefs.

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

This restriction limits the extent to which the distribution of $Z$ can vary with $P$. In our particular specification discussed further in Section 1.7.1.1, we will allow $f(Z|P;\theta)$ to capture noise and bias. In addition to Assumption 1.4.2, we retain Assumptions 1.4.1 and 1.4.1 which ensure that $\Pr \{L = 1|X, Z, P\} = P$.

With Assumptions 1.4.1-1.4.2, the joint p.d.f./p.m.f. of the observed variables $L^{29}$ and $Z$

---

27 In Appendix (1.B.1.3), we provide a stylized example of elicitation error which yields conditions under which orderings of our lower bounds do imply orderings of the true magnitude. Loosely, we require the error in the elicitation to be similar between the two segments under comparison.

28 Because $E[m(P)]$ is a measure of dispersion, it is invariant to location shifts in the distribution of $P$ (i.e. if $\hat{P} = P + \eta$, then $E[m(P)] = E[m(\hat{P})]$). So, testing equation (1.6) is distinct from analyzing whether rejectees have higher mean risk, as suggested by Proposition 1.1. Since rejectees have almost universally higher mean risks, testing for higher values of $E[m(P)]$ for the rejectees may a priori be an overly restrictive test of the theory. Since this biases us against finding results consistent with the theory, we do not discuss this interaction in detail. We discuss this further in Appendix 1.B.1, where we show that the minimum pooled price ratio is bounded above (using a Holder inequality) by a term increasing in both the mean, $\Pr \{L|X\}$, and $E[m(P)|X]$.

29 For notational brevity, we let $L$ also denote the binary indicator that the event $L$ occurs, $1\{L\}$.
(conditional on $X = x$), denoted $f_{L,Z}(L,Z)$, is given by

$$f_{L,Z}(L,Z) = \int_0^1 f_{L,Z}(L,Z|P = p) f_P(p) \, dp$$

$$= \int_0^1 (\text{Pr}\{L = 1|Z, P = p\})^L (1 - \text{Pr}\{L = 1|Z, P = p\})^{1-L} f_{Z|P}(Z|p; \theta) f_P(p) \, dp$$

$$= \int_0^1 p^L (1 - p)^{1-L} f_{Z|P}(Z|p; \theta) f_P(p) \, dp$$

where $f_P(p)$ is the unobserved density of the distribution of private information (assumed to be continuous for ease of exposition). The first equality follows by taking the conditional expectation with respect to $P$. The second equality follows by expanding the joint density of $L$ and $Z$ given $P$ and Assumption 1.4.2. The third equality follows from Assumptions 1.4.1 and 1.4.1.

Assumption 1.4.2 allows us to estimate $\theta$, as opposed to an arbitrary two-dimensional continuous function, $f_{Z|P}$. This allows us to estimate both $\theta$ and $f_P$ using the observed joint distribution of the data, $f_{L,Z}(L,Z)$. We discuss identification in general and for our particular functional form choice in Appendix 1.B.2. But the order condition is straightforward. The observed joint distribution, $f_{L,Z}$, contains two continuous functions of $Z$ (one for $L = 1$ and another for $L = 0$). We use one of these functions to identify $f_P$ and another to identify $\theta$. While we have imposed a functional form on $\theta$, we do not impose a functional form on the distribution of private information, $f_P$. In practice, we flexibly approximate $f_P$ and estimate all parameters (both $\theta$ and the approximating parameters for $f_P$) using maximum likelihood.

Given estimates of the distribution of private information, we translate these into measures of the barrier to trade imposed by private information. Recall from Corollary 1.2 that this magnitude is fully characterized by the minimum pooled price ratio, $\inf_{p \in \hat{\Psi} \setminus \{1\}} T(p)$, where $T(p)$ can be calculated at each $p$ using estimates of $E[P | P \geq p]$ derived from the estimated distribution of private information. One remaining limitation is that for values of $p$ in the upper quantiles of $F(p)$, $E[P | P \geq p]$ is an extreme value that is not well-identified, since the expectation is taken with respect to a smaller and smaller effective sample as $p$ increases. However, for a fixed quantile $\tau$, estimates of the minimum pooled price ratio over $\hat{\Psi}_\tau = [0, F^{-1}(\tau)] \cap (\hat{\Psi} \setminus \{1\})$ are continuously differentiable functions of the MLE parameter estimates of $F(p)$ for $p \leq F^{-1}(\tau)$.\footnote{Non-differentiability could hypothetically occur at points where the infimum is attained at distinct values of $p$.} So, derived MLE estimates of $\inf_{p \in \hat{\Psi}_\tau} T(p)$ are consistent and asymptotically normal. Thus, our approach is to construct the minimum pooled price ratio over $\hat{\Psi}_\tau$ for a fixed $\tau < 1$. We then assess robustness to the choice of $\tau$.

While our motivation for restricting attention to $\hat{\Psi}_\tau$ as opposed to $\hat{\Psi}$ is primarily because of statistical limitations, Remark 1.1 in Section 1.2.2 provides an economic rationale for why $\inf_{p \in \hat{\Psi}_\tau} T(p)$ may not only be a suitable substitute for $\inf_{p \in \hat{\Psi} \setminus \{1\}} T(p)$ but also may actually be more relevant if firms face frictions to setting up contracts. If contracts must attract a non-trivial fraction $1 - \tau$ of
the market in order to be viable, then \( \inf_{p \in \mathcal{P}} T(p) \) characterizes the barrier to trade imposed by private information.

In short, the semiparametric approach makes an additional parametric assumption on the statistical relationship between elicitations \( Z \) and beliefs \( P \) which allows us to estimate the distribution of private information and implement the tests outlined in Section (1.3.3).

1.5 Setting and Data

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

1.5.1 The Three Non-Group Market Settings

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

The private disability insurance protects against the lost income resulting from a work-limiting disability. It is primarily sold through group settings, such as one's employer; more than 30% of private workers have group-based disability policies. In contrast, the non-group market is quite small. Only 3% of non-government workers own a private non-group disability policy, most of whom are self-employed or professionals who do not have access to employer-based group policies (ACLI (2010)).

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

Not everyone can purchase insurance in these three non-group markets. As mentioned in the introduction, Murtaugh et al. (1995) estimates that 12-23% of 65 year olds have a health condition which would cause them to be rejected by LTC insurers. In life and disability insurance, we know of no formal studies documenting the prevalence of rejections, but our review of underwriting

---

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

32 In contrast to health insurance where the group market faces significant tax advantages, group disability policies are taxed. Either the premiums are paid with after-tax income, or the benefits are taxed upon receipt.
guidelines and conversations with underwriters in these markets establish a prevalence of rejections based on certain pre-existing conditions that we discuss in more detail in Section 1.5.2.2.

Insurance companies in these markets are not legally prevented from charging higher prices to reflect actuarial differences in risk.\footnote{The Civil Rights Act does prevent purely racial discrimination in pricing.} They do face some regulation. Capital levels must be maintained to prevent policy default. Also, they are limited in the extent to which policy prices can be raised over time after purchase, which is intended to prevent exploitative price increases on those who have already sunk payments into a policy. But no regulation prevents insurance companies from offering risk-adjusted prices to those who are currently rejected in these three market settings.\footnote{Interviews with underwriters in these markets also suggest that fear of regulation is not an issue in preventing charging a higher price to those currently rejected.}

Previous research has found minimal or no evidence of private information using the revealed preference approach in these settings. In life insurance, Cawley and Philipson (1999) find no evidence of adverse selection. He (2009) revisits this with a different sample focusing on new purchasers and does find evidence of small amounts of adverse selection. In long-term care, Finkelstein and McGarry (2006) find direct evidence of private information by showing subjective probabilities are correlated with subsequent nursing home use. However, they find no evidence that this private information leads to adverse selection in the form of a correlation between insurance purchase and subsequent losses in the LTC insurance market.\footnote{They suggest heterogeneous preferences, in which good risks also have a higher valuation of insurance, can explain why private information doesn’t lead to adverse selection.} To our knowledge, there is no previous study of private information in the non-group disability market.

1.5.2 Data

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

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

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

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

1.5.2.1 Loss Variables and Subjective Probability Elicitations

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

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

Figures 1-1(a,b,c) display histograms of these responses (divided by 100 to translate into probabilities). As has been noted in previous literature using these subjective probabilities (Gan et al. (2005); Finkelstein and McGarry (2006)), these histograms highlight why it would be problematic to view these as true beliefs. Many respondents report 0, 50, or 100. Taken literally, responses of 0 or 100 imply an infinite degree of certainty, which is difficult to believe. We find it more likely that respondents who report focal point values are responding on more of an ordinal scale (e.g. high, medium, low) as opposed to having a literal probabilistic interpretation. Our lower bound approach remains agnostic on the way in which focal point responses relate to true beliefs. Our parametric approach will take explicit account of this focal point response bias, discussed further in Section 1.7.1.1.

Corresponding to each subjective probability elicitation, we construct binary indicators of the loss, L. In long-term care, L denotes the event that the respondent enters a nursing home in the subsequent 5 years. In disability, L denotes the event that the respondent reports that their health limits their work activity in the subsequent 10-11 years. In life, L denotes the event that

36 We use the sample selection described in Subsection (1.5.2.4)
37 For our empirical specification, we will include indicators for focal point responses
38 Our loss variable is necessarily defined as 11 years for those in the AHEAD 1993 wave 2 group because the panel does not provide responses exactly 10 years from 1993. Our results are robust to the exclusion of this group.
Figure 1-1: Subjective Probability Histograms

(a) Long-Term Care

(b) Disability

(c) Life
the respondent dies before AGE, where \( \text{AGE} \in \{75, 80, 85, 90, 95, 100\} \) corresponds to the subjective probability elicitation, which is 10-15 years from the survey date.\(^{39}\)

### 1.5.2.2 Rejection Classification

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

To identify conditions that lead to rejection, we obtain underwriting guidelines used by underwriters and provided to insurance agents for use in screening applicants. An insurance company’s underwriting guidelines provide a list of conditions for which underwriters are instructed to not offer insurance at any price. These guidelines are not publicly available, which limits our ability to obtain this information. The extent of our access varies by market: In long-term care, we obtain a set of guidelines used by an insurance broker from 18 of the 27 largest long-term care insurance companies collectively representing over 95% of the US market.\(^{40}\) In disability and life, we obtain several underwriting guidelines and supplement this information with interviews with underwriters at several major insurance companies.

To match these conditions to our dataset, we use the detailed health and demographic information available in the HRS to identify individuals with conditions which would lead them to be rejected. While the HRS contains a relatively comprehensive picture of respondents’ health, sometimes the conditions which would lead to rejection are too precise to be accurately matched in the HRS. For example, individuals with advanced stages of lung disease would be unable to purchase life insurance; however, the HRS only provides information for the presence of a lung disease.

We exercise caution in performing this match by constructing a third classification, "Uncertain", to which we classify those who may be rejected, but for whom data limitations prevent a solid assessment. This allows us to be relatively confident in our classification of rejectees and non-rejectees. We present our lower bound estimates for all three classifications.\(^{41}\)

Table 1.1 presents the list of conditions for the rejection and uncertain classification, along with the frequency of each condition in our sample (using the sample selection outlined below in Section 1.5.2.4). In long-term care, activity of daily living (ADL) restrictions (e.g. needs assistance walking, dressing, using toilet, etc.), any previous stroke, any previous home care, and anyone over the age of 80 would be rejected. In disability, a back condition, obesity (40+ BMI), and doctor-diagnosed psychological conditions such as depression or bi-polar would lead to rejection. In life, individuals

---

\(^{39}\)We construct the corresponding elicitation to be \( 100\% - Z_{\text{live}} \) where \( Z_{\text{live}} \) is the survey elicitation for the probability of living to \( \text{AGE} \).

\(^{40}\)These guidelines display broad consistency in the rejection practices across firms. We thank Amy Finkelstein for making this broker-collected data available.

\(^{41}\)For brevity, we do not present results from our semiparametric approach for the uncertain group.
Table 1.1: Rejection Classification

<table>
<thead>
<tr>
<th>Classification</th>
<th>Long-Term Care</th>
<th>% Sample</th>
<th>Disability</th>
<th>% Sample</th>
<th>Life</th>
<th>% Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Condition</td>
<td></td>
<td>Condition</td>
<td></td>
<td>Condition</td>
<td></td>
</tr>
<tr>
<td>Rejection</td>
<td>Any ADL/IADL Restriction</td>
<td>6.5%</td>
<td>Back Condition</td>
<td>22.7%</td>
<td>Cancer (Current)</td>
<td>13.1%</td>
</tr>
<tr>
<td></td>
<td>Past Stroke</td>
<td>7.8%</td>
<td>Obesity (BMI &gt; 40)</td>
<td>1.7%</td>
<td>Stroke (Ever)</td>
<td>7.3%</td>
</tr>
<tr>
<td></td>
<td>Past Nursing/Home Care</td>
<td>12.4%</td>
<td>Psychological Condition</td>
<td>6.3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Over age 80</td>
<td>18.4%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uncertain</td>
<td>Lung Disease</td>
<td>10.0%</td>
<td>Arthritis</td>
<td>36.9%</td>
<td>Diabetes</td>
<td>13.8%</td>
</tr>
<tr>
<td></td>
<td>Heart Condition</td>
<td>28.4%</td>
<td>Diabetes</td>
<td>7.7%</td>
<td>High Blood Pressure</td>
<td>50.7%</td>
</tr>
<tr>
<td></td>
<td>Cancer (Current)</td>
<td>14.7%</td>
<td>Lung Disease</td>
<td>5.1%</td>
<td>Lung Disease</td>
<td>10.9%</td>
</tr>
<tr>
<td></td>
<td>Hip Fracture</td>
<td>1.3%</td>
<td>High Blood Pressure</td>
<td>35.3%</td>
<td>Cancer (Ever, not current)</td>
<td>12.1%</td>
</tr>
<tr>
<td></td>
<td>Memory Condition¹</td>
<td>0.8%</td>
<td>Heart Condition</td>
<td>6.1%</td>
<td>Heart Condition</td>
<td>26.5%</td>
</tr>
<tr>
<td></td>
<td>Other Major Health Problems²</td>
<td>26.7%</td>
<td>Cancer (Ever Have)</td>
<td>4.6%</td>
<td>Other Major Health Problems²</td>
<td>23.5%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Blue-collar/high-risk Job³</td>
<td>23.3%</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Other Major Health Problems²</td>
<td>16.2%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

²Wording 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?"

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

⁴We exclude minor basel cell cancers

Note that percentages will not add to the total fraction of the population classified as rejection and uncertain because of people with multiple conditions.
with a past stroke or current cancer would be rejected. We classify individuals with these conditions as rejected in their respective markets.

Table 1.1 also lists the conditions leading to an uncertain classification in each market. In addition to specific conditions for which the HRS data is too coarse, we also attempt to capture the presence of rarer conditions not asked in the HRS (e.g. Lupus would lead to rejection in LTC, but is not explicitly reported in the HRS). To do so, we take advantage of a question in the HRS which asks respondents if they have any additional major health problems which were not asked about in the survey. We classify individuals reporting yes to this question as Uncertain.

1.5.2.3 Public Information

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

The HRS contains an extensive set of health, demographic, and occupation information which allows us to approximate the set of information which insurance companies use in pricing insurance.\(^42\) The quality of this approximation varies by market. For long-term care, we replicate the information set of the insurance company quite well. For example, perhaps the most obscure piece of information that is acquired by some LTC insurance companies is an interview in which applicants are asked to perform word recall tasks to assess memory capabilities; the HRS conducts precisely this test with survey respondents. In disability and life, we replicate most of the information used by insurance companies in pricing. One caveat is that insurance companies will sometimes perform tests, such as blood and urine tests, which we will not observe in the HRS. Conversations with underwriters in these markets suggest these tests are primarily to confirm application information, which we can approximate quite well with the HRS. But, we cannot rule out the potential that there is additional information which can be gathered by insurance companies in the disability and life settings.\(^43\)

In addition to our preferred specification which includes variables used in pricing, we also assess the robustness of our estimates to alternative sets of controls.\(^44\) This is for two reasons.

\(^42\) We are not the first to note the ability of the HRS to replicate the information used by insurance companies in pricing; for LTC, see Finkelstein and McGarry (2006) and for Life, see He (2009).

\(^43\) In LTC, insurance companies are legally able to conduct tests, but it is not common industry practice.

\(^44\) 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
First, although we have been careful in constructing the pricing controls, it may not be a perfect representation of the set of information used in pricing. Second, we do not want our conclusions for the amount of private information for the rejectees to depend on an assumption of how insurance companies would hypothetically use information to price their contracts. We therefore perform our analysis for three increasing sets of public information:

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

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

We conduct the lower bound approach for all three sets of controls. For brevity, we focus exclusively on our preferred specification of pricing controls for our semiparametric approach.\(^45\)

The variables used in the pricing and full controls specifications for each market are presented in Table 1.2. In LTC, our preferred specification includes age, age squared, and gender interactions; indicators for various health conditions; ADL restrictions; and performance on a word recall test. Our extended controls specification adds full interactions for age and gender, along with interactions of 5 year age bins with measures of health conditions, indicators for the number of living relatives (up to 3), census region, and income deciles. For disability, our preferred specification includes age, age squared, and gender interactions; indicators for self employment and various health conditions; BMI; and wage decile. Our extended controls specification adds full interactions of age and gender; full interactions of wage decile, part time status indicator, job tenure quartile, and self-employment indicator; interactions between 5 year age bins and various health conditions and BMI; full interactions of job characteristics (e.g. “job requires heavy lifting”); and full interactions of 5 year age bins and census region. For life, our preferred specification includes age, age squared, and gender interactions, smoking status, indicators for the death of a parent before age 60, BMI, income decile, and indicators for a set of health conditions. We also include a set of indicators for the years between the survey date and the \(\text{AGE}\) corresponding to the loss.\(^46\) Our extended controls specification adds full interactions of age and gender; full interactions between age and the \(\text{AGE}\) predictive power of another variable (price, in this case). Of course, conditioning on additional variables \(X'\) which are uncorrelated with \(L\) or \(Z\) has no effect on the population value of \(E[m(P)|X \in \Theta]\).\(^45\)

\(^45\)Because the extended controls specification includes a lot of variables, we risk over-fitting the data. As we discuss in Section 1.6.1, this does not pose an insurmountable problem for the lower bound approach. However, it would pose a problem for the semiparametric approach, and thus provides another reason for our exclusive focus on the preferred pricing specification.

\(^46\)We also include this in our age & gender and extended control specifications for life.
<table>
<thead>
<tr>
<th>Table 1.2: Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Long-Term Care</strong></td>
</tr>
<tr>
<td><strong>Price Controls</strong></td>
</tr>
<tr>
<td>Age, Age(^2), Gender</td>
</tr>
<tr>
<td>Gender(*age)</td>
</tr>
<tr>
<td>Word Recall Performance(^1)</td>
</tr>
<tr>
<td>Indicators for ADL/IADL Restriction</td>
</tr>
<tr>
<td>Psych Condition</td>
</tr>
<tr>
<td>Diabetes</td>
</tr>
<tr>
<td>Lung Disease</td>
</tr>
<tr>
<td>Arthritis</td>
</tr>
<tr>
<td>Heart Disease</td>
</tr>
<tr>
<td>Cancer</td>
</tr>
<tr>
<td>Stroke</td>
</tr>
<tr>
<td>High blood pressure</td>
</tr>
<tr>
<td>Interactions between 5 yr age bins and the presence of: Number of Health Conditions (High bp, diabetes, heart condition, lung disease, arthritis, stroke, obesity, psych condition)</td>
</tr>
<tr>
<td>Number of ADL / IADL Restrictions</td>
</tr>
<tr>
<td>Number of living relatives (&lt;=3)</td>
</tr>
<tr>
<td>Past home care usage</td>
</tr>
<tr>
<td>Census region (1-5)</td>
</tr>
<tr>
<td>Income Decile</td>
</tr>
</tbody>
</table>

\(^1\)Indicator for lowest quartile performance on word recall test
\(^2\)Full indicator variables for number of years to AGE reported in subjective probability question
in the subjective probability question; interactions between 5 year age bins and smoking status, income decile, census region, and various health conditions; \(^{47}\) BMI; and an indicator for death of a parent before age 60.

1.5.2.4 Sample Selection

For each sample, we begin with years 1993-2008 of the HRS. Our selection process varies across each of the three market settings due to data constraints. Table 1.3 presents the summary statistics for each sample.

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

In addition to our primary sample, we will report results for our nonparametric lower bounds using a sample that excludes individuals who own long-term care insurance (roughly 13% of remaining sample) to ensure we estimate private information inherently held by the individual which is not the effect of insurance contract choice on subsequent utilization (a.k.a. "moral hazard").\(^{48}\)

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

**Disability** For disability, we begin with the set of individuals up to age 60 who are currently working and report no presence of work-limiting disabilities. To construct the corresponding loss realization, we limit the sample to individuals who we can observe for a subsequent 10 years (years

\(^{47}\)Although the HRS asks whether respondents have (non-basal cell) cancer, it only asks which organ the cancer occurs in the 2nd wave (1993/1994) of the survey. In the robustness section, we will consider an additional extended controls specification for life insurance which uses data only from these years and includes a full set of cancer organ indicators (50+ indicators).

\(^{48}\)While one might be tempted to control for the purchase of insurance or the contract characteristics, this would be misguided. If agents with different beliefs sort into different contracts, controlling for contract choice could lead to a finding of no private information. Insurance purchase is a potentially endogenous response to the presence of private information and thus should not be included as a control variable.
Table 1.3: Sample Selection

<table>
<thead>
<tr>
<th>Subj. Prob (mean)(^1) (std dev)</th>
<th>Long-Term Care</th>
<th>Disability</th>
<th>Life</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No Reject</td>
<td>Reject</td>
<td>Uncertain</td>
<td>No Reject</td>
<td>Reject</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.111 (0.194)</td>
<td>0.168 (0.249)</td>
<td>0.132 (0.207)</td>
<td>0.292 (0.257)</td>
<td>0.385 (0.264)</td>
</tr>
<tr>
<td>Loss</td>
<td>0.039 (0.195)</td>
<td>0.175 (0.38)</td>
<td>0.054 (0.227)</td>
<td>0.156 (0.363)</td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>71.7 (4.366)</td>
<td>79.4 (6.934)</td>
<td>72.2 (4.303)</td>
<td>54.7 (4)</td>
</tr>
<tr>
<td>Female</td>
<td>0.622 (0.485)</td>
<td>0.631 (0.483)</td>
<td>0.564 (0.496)</td>
<td>0.606 (0.489)</td>
</tr>
<tr>
<td>Health Status Indicators</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arthritis</td>
<td>0.479 (0.5)</td>
<td>0.616 (0.486)</td>
<td>0.552 (0.497)</td>
<td>0.000 (0)</td>
</tr>
<tr>
<td>Diabetes</td>
<td>0.140 (0.347)</td>
<td>0.172 (0.377)</td>
<td>0.147 (0.354)</td>
<td>0.000 (0)</td>
</tr>
<tr>
<td>High Blood Pressure</td>
<td>0.505 (0.5)</td>
<td>0.598 (0.49)</td>
<td>0.535 (0.499)</td>
<td>0.280 (0.449)</td>
</tr>
</tbody>
</table>

Sample Size
- Observations (Ind x wave): 9,051, 10,108, 10,690, 2,540, 2,216, 3,757, 2,689, 2,362, 6,800
- Unique Individuals: 4,418, 3,215, 5,190, 1,480, 1,280, 1,929, 1,720, 1,371, 4,270
- Unique Households: 3,283, 2,620, 3,860, 1,112, 975, 1,540, 1,419, 1,145, 3,545

Fraction Insured\(^2\)
- 13.9% | 10.7% | 14.8% | 65.1% | 63.3% | 64.2%

\(^1\)We transform the life insurance variable to 1-Pr(living to \text{AGE}) to correspond to the loss definition.

\(^2\)Calculated based on full sample prior to excluding individuals who purchased insurance.
2000-2008 are used solely for the construction of the loss indicator). Our final sample consists of 2,540 observations from 1,480 individuals for our no reject classification, 2,216 observations from 1,280 individuals for our reject classification, and 3,757 observations from 1,929 individuals for our uncertain classification.49

Rejectees differ from non-rejectees on many dimensions. They are more likely to have high blood pressure, diabetes, and arthritis,50 and have a higher risk of experiencing a work-limiting disability (44.1% versus 15.6%). But, similar to LTC, not everyone with a rejection condition will experience a work-limiting disability in the subsequent 10 years, which again suggests they face unrealized risk.

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

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

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

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

Third, rejectees differ from non-rejectees on many dimensions; their older, sicker, and have a higher probability of experiencing the loss. This is consistent with Proposition 1.1 which showed that higher risk distributions are more likely to satisfy the no-trade condition.

---

49Ideally, we would also test the robustness of our results using a sample of those who do not own disability insurance, but unfortunately the HRS does not ask about disability insurance ownership.

50Diabetes and arthritis may lead to rejection, so those without a rejection condition but with one of these two conditions are placed in the uncertain classification.
1.5.2.5 Relation to Ideal Data

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

1.6 Lower Bound Estimation

We now turn to the estimation of lower bounds of the average magnitude of private information, $E[m_Z(P_Z)|X]$, outlined in Section 1.4.1.

1.6.1 Empirical Estimation

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

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

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

where $X$ are our control variables and $\Gamma(age, Z)$ captures the effect of $Z$ on $Pr\{L|X, Z\}$.

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

41
form allows the affect of $Z$ to vary with age (note that age is already included in $X$).\footnote{In our LTC Reject Sample, we also include full interactions between $\Gamma$ and an indicator for having a rejection health condition; this allows $\Gamma$ to vary differentially for those over age 80 with no other rejection conditions besides age $\geq 80$.} We approximate $\Gamma (\text{age}, Z)$ using full interactions of functions of $Z$ and functions of $\text{age}$. For $Z$, we use second-order Chebyshev polynomials plus separate indicators for focal point responses at $Z = 0$, 50, and 100. We use a linear function of $\text{age}$. Our approximation of $\Gamma (\text{age}, Z)$ is then given by the full set of these interactions (whose coefficients are to be estimated). All results are robust to the inclusion of additional or fewer polynomials in $Z$ or $\text{age}$, or the use of a linear or logit specification, as opposed to the probit. For our extended controls specification, the high dimensionality of $X$ leads us to use a linear specification, $\Pr \{ L | X, Z \} = \beta X + \Gamma (\text{age}, Z)$. We use the same approximation for $\Gamma$.\footnote{Of course, the switch from the probit to linear specification leads $\Gamma$ to have a different interpretation.}

Estimating $m_Z (p) = E [ P_Z | P_Z \geq p, X ] - p$ requires estimating the entire distribution of $P_Z$ at each possible value of $X$. To make this feasible, we adopt an assumption for how the distribution of $P_Z$ varies with $X$: conditional on ones age and rejection classification, the distribution of residual private information implied by $Z$, $P_Z - E [ P_Z | X ]$, does not vary with $X$. This allows observable variables affect the mean but not the shape of the distribution of $P_Z$ (conditional on age and rejection classification).\footnote{This assumption is only required to arrive at a point estimate for $E [ m_Z (P_Z) | X \in \Theta ]$, and is not required to test for the presence of private information (i.e. whether $\Gamma = 0$). Also, our results for $E [ m_Z (P_Z) | X \in \Theta ]$ are robust to alternative assumptions, such as assuming the residual distribution does not vary with $X$ within 5 year age bins.} We then estimate the conditional expectation, $m_Z (p) = E [ P_Z | P_Z \geq p, X ] - p$ using the empirical distribution of $P_Z - E [ P_Z | X ]$ within each age grouping.

After estimating $m_Z (p)$, we construct its average using the empirical distribution of $P_Z$, yielding $E [ m_Z (P_Z) | X \in \Theta ]$, where $\Theta$ is a given sample (e.g. LTC rejectees). For each market, we then construct the difference between the reject and no reject estimates,

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

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

### 1.6.2 Statistical Inference

Statistical inference for $E [ m_Z (P_Z) | X \in \Theta ]$ for a given sample $\Theta$ and for $\Delta_Z$ is straightforward, but requires a bit of care to cover the possibility of no private information. In any finite sample, our estimates of $E [ m_Z (P_Z) | X \in \Theta ]$ 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 amongst those with \( X \in \Theta \)), then the bootstrap distribution is not asymptotically normal and does not provide adequate finite-sample inference.\(^{55}\) We therefore supplement the bootstrap with a Wald test which restricts \( \Gamma(\text{age}, Z) = 0 \).\(^{56}\) This tests for the presence of private information. We report results from both the Wald test and the bootstrap.

We conduct inference on \( \Delta_Z \) in a similar manner. To test the null hypothesis that \( \Delta_Z \leq 0 \), we construct conservative p-values by taking the maximum p-value from two tests: 1) a Wald test of no private information held by the rejectees, \( E\{m_Z(P_Z) | X \in \Theta^{\text{Reject}} \} = 0 \), and 2) the p-value from the bootstrapped event of less private information held by the rejectees, \( \Delta \leq 0 \).\(^{57}\)

### 1.6.3 Graphical Results for \( P_Z - E\{P_Z|X\} \)

We begin with graphical evidence of the predictive power of subjective probability elicitations. Figure 1-2(a,b,c) plots the estimated distribution of \( P_Z - E\{P_Z|X\} \) aggregated by rejection classification for the rejectees and non-rejectees, using our preferred pricing control specification.\(^{58}\)

Consistent with the hypothesis that rejectees are better informed about whether or not they would experience the loss, the distribution of \( P_Z - E\{P_Z|X\} \) is more dispersed for the rejectees relative to those served by the market in all three market settings we consider. As we now show, this translates into higher estimates of our lower bounds on the average magnitude of private information.

### 1.6.4 Lower Bound Results

The top row of Table 1.4 provides the estimates of \( \Delta_Z \). Across all specifications and all market settings, we estimate larger lower bounds on the average magnitude of private information for the rejectees relative to those served by the market. These differences are all statistically significant at the 1% level (third row of Table 1.4). Consistent with the theory, this suggests private information imposes a greater barrier to trade for rejectees relative to those served by the market.

The lower sets of rows in Table 1.4 report the estimated magnitude of private information for each classification (No Reject, Reject, and Uncertain), along with their standard errors and p-values for the presence of private information. We discuss these details by market.

---

\(^{55}\)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).

\(^{56}\)The event \( \Gamma(\text{age}, Z) = 0 \) in sample \( \Theta \) is equivalent to both the event \( \Pr\{L|X, Z\} = \Pr\{L|X\} \) for all \( X \in \Theta \) and the event \( E\{m_Z(P_Z) | X \in \Theta^{\text{Reject}} \} = 0 \).

\(^{57}\)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, we construct conservative p-values robust to any weight in \([0, 1]\).

\(^{58}\)Subtracting \( E\{P_Z|X\} = \Pr\{L|X\} \) allows for simple aggregation across \( X \) within each sample.
Figure 1-2: Distribution of Residual Private Information

(a) LTC

(b) Disability

(c) Life
Table 1.4: Magnitude of Private Information (Lower Bound)

<table>
<thead>
<tr>
<th></th>
<th>LTC</th>
<th>Disability</th>
<th>Life</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age &amp; Gender</td>
<td>Price Controls</td>
<td>Extended Controls</td>
</tr>
<tr>
<td>Difference: $\Delta z$</td>
<td>0.0234***</td>
<td>0.0245***</td>
<td>0.0213***</td>
</tr>
<tr>
<td>Bootstrap s.e.</td>
<td>0.0041 (0.004)</td>
<td>0.0000 (0.0004)</td>
<td>0.0000 (0.004)</td>
</tr>
<tr>
<td>p-value$^2$</td>
<td>0.0041 0.0041</td>
<td>0.0040 0.0040</td>
<td>0.0282***</td>
</tr>
<tr>
<td>No Reject</td>
<td>0.3880 0.4330</td>
<td>0.2156 0.0656</td>
<td>0.0009 0.0056</td>
</tr>
<tr>
<td>Reject</td>
<td>0.0274***</td>
<td>0.0286***</td>
<td>0.0253***</td>
</tr>
<tr>
<td>Bootstrap s.e.</td>
<td>0.0038 (0.0036)</td>
<td>0.0030 (0.0034)</td>
<td>0.0098 (0.0089)</td>
</tr>
<tr>
<td>Wald test p-value$^3$</td>
<td>0.0000 0.0000</td>
<td>0.0000 0.0000</td>
<td>0.0000 0.0000</td>
</tr>
<tr>
<td>Uncertain</td>
<td>0.0058**</td>
<td>0.0056**</td>
<td>0.0053**</td>
</tr>
<tr>
<td>Bootstrap s.e.</td>
<td>0.0002 (0.0018)</td>
<td>0.0019 (0.0019)</td>
<td>0.0072 (0.0066)</td>
</tr>
<tr>
<td>Wald test p-value</td>
<td>0.0121 0.0472</td>
<td>0.0147 0.0000</td>
<td>0.0000 0.0000</td>
</tr>
</tbody>
</table>

$^1$Bootstrap standard errors computed using block re-sampling at the household level (results shown for N=500 repetitions)

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

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

*** $p<0.01$, ** $p<0.05$, * $p<0.10$
In long-term care, we find significant evidence of private information amongst the rejectees, with estimated magnitudes of 0.0286 ($p < 0.001$) in our preferred specification.\footnote{Because the estimated magnitudes are lower bounds, we do not focus our discussion on their absolute magnitudes. But the interpretation is straightforward: the estimated magnitude of 0.0286 implies that $E[P_z|P_z \geq p]$ differs from $p$ by 0.0286 on average, which implies that, on average, the average predicted probability of a loss (given $Z$) for worse risks differs from ones' own risk by 2.86pp.} In contrast, we find no statistically significant evidence of private information held by the non-rejectees (0.0041, $p = 0.433$ for our preferred specification). The estimates are quite similar for different control specifications: all estimates lie within an estimated standard error (0.004).

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

In disability, we find significant evidence of private information held by both the rejectees and non-rejectees. In our preferred specification, we estimate magnitudes of 0.0512 ($p < 0.01$) for the rejectees and 0.0257 ($p < 0.01$) for non-rejectees, leading to an estimated difference of $\Delta Z = 0.0255$ ($p = 0.006$). The results are robust to the inclusion of additional controls: the extended controls specification yields statistically indistinguishable results from our preferred specification (.0234, $p = .018$). Our age and gender specification leads to slightly higher estimated magnitudes for the rejectees of 0.0737, but not significantly different from our price controls specification.

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

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

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

**Uncertain Classification** The estimated magnitudes for the uncertain classification generally fall between the estimates for the rejection and no rejection groups, as indicated by the bottom set of rows in Table 1.3. In general, our 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. \( E[m_Z(P_Z)|X] \approx m^R \)) and non-rejectees (e.g. \( E[m_Z(P_Z)|X] \approx m^{NR} \)), then linearity of the expectation implies

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

(1.9)

where \( \lambda \) 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 our estimates for the rejectees and non-rejectees, as we find. Nevertheless, we have no theoretical reason to suppose the average magnitude of private information is constant within rejection classification; thus this should be viewed only as a potential method for interpreting the results, not as a robust prediction of the theory.

1.6.5 Subsample Analysis

The results in Table 1.4 aggregate across all observables, \( X \), within each rejection classification. While this aggregation improves statistical power, it is important to also examine the results within subgroups to test whether the rejectees have higher magnitudes of private information conditional on observable variables.\(^{60}\) In this section, we examine age-based subgroups.

Figure 1-3(a, bi, bii, c) plots the estimates of \( E[m_Z(P_Z)|X \in \Theta^{\text{age, rejectclass}}] \) separately for each age and rejection classification.\(^{61}\) In all three settings, we find larger estimates for the rejectees versus non-rejectees conditional on age.

---

\(^{60}\)In addition to analyzing subgroups as a finer test of the theory, one might also worry that aggregation masks other potential drivers of the magnitude of private information aside from the presence of rejection conditions. In particular, the rejectees are generally older than the non-rejectees. If older people naturally, for some reason, have more private information, irrespective of whether or not they would be rejected, then we would estimate \( \Delta_Z > 0 \) in aggregate, even though it may not be the case that \( \Delta_Z > 0 \) conditional on age.

\(^{61}\)We use our preferred specification, which is quite flexible in age and allows \( \Gamma \) to vary with age. Figure 1-3(a, bi, bii, c) provides bootstrapped standard errors, which are consistent as long as \( \Gamma \neq 0 \). In general, we cannot reject the null hypothesis that \( \Gamma = 0 \) on a subsample consisting of one specific age, but our results in Table 1.4 do reject \( \Gamma = 0 \) at all ages for all but the LTC and life no reject samples.
Figure 1-3: Magnitude of Private Information (Lower Bound)

(a) LTC

(b) Disability: Females (Left) and Males (Right)

(c) Life
In long-term care, we can also more closely examine the specific rejection practices based on age. LTC insurers reject applicants above age 80 regardless of health conditions (such as ADL restrictions or a past stroke). Figure 1-3a plots the lower bound estimates at each age, separately reporting estimates for those with and without rejection health conditions above age 80. The results show that the estimates for those without rejection conditions increases at ages nearing 80. Indeed, an individual at age 81 with no other rejection conditions (but who would be rejected based on age) has a very similar magnitude of private information to a 65 year old who would be rejected. In short, the results provide a picture of why insurance companies automatically reject individuals beginning at age 80 as opposed to other age cutoffs.

In life (Figure 1-3c), we find larger estimates for the rejectees across the age spectrum. For disability, we also generally find larger estimates (Figure 1-3b), although the difference between the reject and no reject estimates appears to be increasing in age. In short, we find larger estimates for rejectees conditional on age.

1.6.6 Robustness

1.6.6.1 Insurance Ownership Sample Selection

Our primary results in Table 1.4 do not exclude individuals who own insurance. If insurance choice affects the risk of experiencing the loss, then differential insurance ownership could cause a finding of private information. We test the robustness of our results to this potential bias by restricting to those who do not own insurance in our LTC and Life samples. Table 1.5 presents these results.

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

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

---

62 We present separate results for males and females in Figure 1-3b because of the changing gender composition of the sample over time. Individuals below age 55 are included in the HRS only if they have a spouse above age 55. Thus, we have relatively more females below age 55. But, as shown in these figures, we generally find larger estimates for rejectees conditional on age and gender. We have also examined LTC and Life by age & gender and the results again show larger magnitudes for the rejectees conditional on age and gender.

63 This is consistent with the much smaller sample size leading to a greater (spurious) predictive power of the subjective probabilities.
Table 1.5: Robustness Checks: Sample Selection

<table>
<thead>
<tr>
<th></th>
<th>LTC, Price Controls</th>
<th>Life, Price Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Primary Sample</td>
<td>Excluding Insured</td>
</tr>
<tr>
<td>Difference: $\Delta_z$</td>
<td>0.0245***</td>
<td>0.0257***</td>
</tr>
<tr>
<td>Bootstrap s.e.¹</td>
<td>(0.004)</td>
<td>(0.0043)</td>
</tr>
<tr>
<td>p-value²</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>No Reject</td>
<td>0.0041</td>
<td>0.0033</td>
</tr>
<tr>
<td>Bootstrap s.e.</td>
<td>(0.0017)</td>
<td>(0.0018)</td>
</tr>
<tr>
<td>Wald test p-value³</td>
<td>0.4330</td>
<td>0.8280</td>
</tr>
<tr>
<td>Reject</td>
<td>0.0286***</td>
<td>0.029***</td>
</tr>
<tr>
<td>Bootstrap s.e.</td>
<td>(0.0036)</td>
<td>(0.0039)</td>
</tr>
<tr>
<td>Wald test p-value³</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Uncertain</td>
<td>0.0056**</td>
<td>0.0056</td>
</tr>
<tr>
<td>Bootstrap s.e.</td>
<td>(0.0018)</td>
<td>(0.0019)</td>
</tr>
<tr>
<td>Wald test p-value³</td>
<td>0.0472</td>
<td>0.1352</td>
</tr>
</tbody>
</table>

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

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

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

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

1.6.6.2 Organ Controls for Life Specification

Our specifications for life insurance did not include controls for the affected organ of cancer sufferers in the reject sample. Although later years of the survey do not specify the organ of the cancer, it is provided in the 1993/4 wave of the survey. In Table 1.6, we report the results from our primary specification (all years) and the results from a specification restricted to years 1993/1994 which includes a full set of 54 indicators for the affected organ added to our set of extended controls. Our finding of significant private information amongst rejectees is robust to including these additional controls. We estimate a value for $E \left[ m_Z (P_Z) | X \in \Theta_{Rej} \right]$ of 0.0308 ($p = .018$) including these controls, as compared to 0.0338 ($p < 0.001$) for our primary specification.

1.6.7 Summary

We estimate significantly larger lower bounds for the average magnitude of private information for the rejectees versus non-rejectees. Our estimates are robust to a wide set of controls for public information, are robust to excluding those who own insurance in LTC and life, and are also consistent within age-based subsamples. Consistent with the theory in Section (1.2), our results suggest private information imposes a greater barrier to trade for the rejectees.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta z$</td>
<td>0.0338***</td>
<td>0.0308**</td>
</tr>
<tr>
<td>Bootstrap s.e.$^1$</td>
<td>(0.0109)</td>
<td>(0.0121)</td>
</tr>
<tr>
<td>p-value$^2$</td>
<td>0.0020</td>
<td>0.0140</td>
</tr>
<tr>
<td>No Reject</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bootstrap s.e.</td>
<td>0.0249</td>
<td>0.0218</td>
</tr>
<tr>
<td>Wald test p-value$^3$</td>
<td>(0.0067)</td>
<td>(0.007)</td>
</tr>
<tr>
<td>Reject</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bootstrap s.e.</td>
<td>0.0587***</td>
<td>0.0526***</td>
</tr>
<tr>
<td>Wald test p-value$^3$</td>
<td>(0.0088)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Uncertain</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bootstrap s.e.</td>
<td>0.0294***</td>
<td>0.0342***</td>
</tr>
<tr>
<td>Wald test p-value$^3$</td>
<td>(0.0053)</td>
<td>(0.0081)</td>
</tr>
</tbody>
</table>

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

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

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

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

1.7 Estimation of Distribution of Private Information

While the lower bound approach provides evidence that private information imposes larger barriers to trade for the rejectees, it suffers several limitations. First, we made comparisons using lower bounds, not the levels of $E[m(P) | X \in \Theta]$. Second, we made comparisons using the average magnitude of private information, $E[m(P) | X]$, not $m(p) \forall p$ or inf $T(p)$ as suggested by Corollaries 1.3 and 1.4. Third, we could not quantify the minimum pooled price ratio.

To overcome these limitations, we introduce additional structure to the statistical relationship between elicitations and beliefs, as outlined in Section 1.4.2.

1.7.1 Empirical Specification

1.7.1.1 Elicitation Error Structure

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

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

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

where $\phi$ denotes the standard normal p.d.f. and $\Phi$ the standard normal c.d.f. We estimate four elicitation error parameters: $(\sigma, \lambda, \kappa, \alpha (X))$. $\sigma$ captures the dispersion in the elicitation error, $\lambda$ is the fraction of focal point respondents, $\kappa$ is the focal point window. We allow the elicitation bias term, $\alpha (X)$, to vary with the observable variables, $X$.

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

### 1.7.1.2 Flexible Approximation for the Distribution of Private Information

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

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

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

where we assume $\tilde{F} (p|q)$ is continuous in $q$. This assumes that the distribution of private infor-

---

64 In Appendix 1.B.2.2, we provide Monte Carlo evidence that our estimation of the distribution of private information, $F (p)$, is reasonably robust to relaxing normality by introducing skewness and kurtosis.

65 This allows elicitations to be biased, conditional on $X$; but we maintain the assumption that true beliefs are unbiased.
information is the same for two segments, $X$ and $X'$, that have the same observable loss probability, $\Pr \{L|X\} = \Pr \{L|X'\}$. We will refer to $q$ as an index. This assumption provides empirical tractability while still allowing the shape of $F$ to vary with observables. Also, recall we conduct estimation separately for the rejectees and non-rejectees, so that we only impose the index assumption conditional on rejection classification.

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

$$\hat{F}(p|q) = \sum_i \eta_i \text{Beta}(\mu_i(q), \psi_i)$$

where $\eta_i$ is the weight on each beta distribution, $\mu_i(q)$ is the mean of the $i^{th}$ beta distribution at $q$, and $\psi_i$ is the shape parameter of the $i^{th}$ beta distribution.\(^6\) We allow the shape parameter to vary for each beta distribution and we allow the mean of each beta distribution to vary as a linear function of $q$, $\mu_i(q) = \mu_i^0 + \mu_i^1 q$. Consistent beliefs (Assumption 1.4.1) imposes the restriction $\sum_i \eta_i \mu_i(q) = q$ which provide constraints on $\{\mu_i^0, \mu_i^1\}$, reducing the number of estimated parameters.

Beta distributions are quite flexible and well-suited for approximating arbitrary distributions. In practice, they fit our data quite well with a small mixture; we use two beta distributions for our preferred results in all settings except the no reject sample for LTC where we include an additional term to capture a point-mass at $q$.\(^6\) All of our results are robust to including a 3rd beta distribution.

Bootstrap delivers consistent standard errors provided the true distribution, $\hat{F}(p|q)$, is continuous. This assumption is violated in the event of no private information (in which case $\hat{F}(p|q) = 1 \{p < q\}$ for all $p, q$). However, the Wald tests for the presence of private information (constructed using our lower bound approach) provide a simple test of this event.\(^6\)

\(^6\) The p.d.f. of a beta distribution with parameters $\alpha$ and $\beta$ is given by

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

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

\(^6\) Although the beta distributions can theoretically approximate uninformed (point-mass) distributions quite well, convergence is slow in practice. We speed up our estimation for the no reject sample for LTC by mixing a truncated normal distribution which converts to a point mass distribution for a variance below 0.000025. This allows the estimation to more easily capture uninformed distributions. Including this point mass term in the other samples does not affect our results.

\(^6\) Notice that $F(p|q) = 1 \{p \leq q\}$ for all $p$ and $q$ if and only if $\Pr \{L|X, Z\} = Pr \{L|X\}$ for all $X$, so that our Wald test for the latter equality (from our lower bound approach) continues to provide a valid test for the presence of private information in our semiparametric approach. One could also construct a test for no private information for various sets of $q$ values, however this suffers problems of limited power (and potential multiple testing issues), so we choose to focus on one aggregate test for the presence of private information in each rejection classification. In principle, imposing our restrictions on $f_{Z|P}(Z|P)$ could produce a more powerful test for the presence of private information. However, such a test faces technical hurdles since it involves testing whether $F(p|q)$ lies along a boundary of the set of possible distributions and must account for sample clustering (which makes a likelihood ratio test inappropriate).
1.7.1.3 Pooled Price Ratio (and its Minimum)

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

As described in Section 1.4.2, we estimate the analogue to the minimum pooled price ratio, \( \inf_{p \in Φ_r} T(p) \) for \( T = [0, F^{-1}(τ)] \). Our preferred choice for τ is 0.8, as this ensures at least 20% of the sample (conditional on q) is used to estimate \( E[|P|P ≥ p] \) and produces estimates that are quite robust to changes in the number of approximating beta distributions. For robustness, we also present results for τ = 0.7 and τ = 0.9 along with plots of the pooled price ratio for all p below the estimated 90th quantile, \( F^{-1}(0.9) \). We construct 5/95% confidence intervals for \( \inf_{p \in Φ_r} T(p) \) by combining bootstrapped confidence intervals and extending the 5% boundary to 1 in the event that we cannot reject a null hypothesis of no private information.

1.7.2 Estimation Results

1.7.2.1 Graphs of the Distribution of Private Information

Qualitatively, no trade requires the existence of a “thick upper tail” of high risks who prevent the provision of insurance to lesser risks. We now assess this prediction. Figure 1-4(a-f) and Figure 1-5(a-f) present the estimated p.d.f.s and c.d.f.s of private information \( \tilde{F}(p|q) \). Figure 1-4 plots the distribution for the index equal to the mean loss in each sample, q = Pr \{L|X ∈ Θ\}, and Figure 1-5 plots the distribution for values of the index at the 20th, 50th, and 80th quantiles of the distribution of q in each sample. In all three market settings and across a wide range of values of the index, q, in each sample, we find evidence of a pronounced upper tail of risk for the rejectees. In contrast, we do not find such a significant upper tail for the non-rejectees. Broadly, these thicker upper tails for rejectees are consistent with the qualitative prediction of the theory that private information leads to rejections.

1.7.2.2 Minimum Pooled Price Ratio

We now turn to our quantitative estimates of the minimum pooled price ratio. Table 1.7 presents the estimates of the minimum pooled price ratio evaluated at several values of the index, q: the sample mean (q = Pr \{L|X ∈ Θ\}), the 20th, 50th, and 80th quantiles of the distribution of q within each sample. We let τ = 0.8 and assess robustness to this choice in Table 1.8, discussed below.

Andrews (2001) provides a potential method for constructing an appropriate test, but we leave this for future work.
Figure 1-4: PDF of Private Information

(a) LTC (Mean $q$)  
(b) Disability (Mean $q$)  
(c) Life (Mean $q$)  
(d) LTC (Quantiles of $q$)  
(e) Disability (Quantiles of $q$)  
(f) Life (Quantiles of $q$)

Figure 1-5: CDF of Private Information

(a) LTC (Mean $q$)  
(b) Disability (Mean $q$)  
(c) Life (Mean $q$)  
(d) LTC (Quantiles of $q$)  
(e) Disability (Quantiles of $q$)  
(f) Life (Quantiles of $q$)
Table 1.7: Minimum Pooled Price Ratio

<table>
<thead>
<tr>
<th>Quantile of Index, q</th>
<th>LTC</th>
<th>Disability</th>
<th>Life</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>20%</td>
<td>50%</td>
</tr>
<tr>
<td>Reject</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5%</td>
<td>1.715</td>
<td>1.681</td>
<td>1.711</td>
</tr>
<tr>
<td>95%</td>
<td>1.538</td>
<td>1.525</td>
<td>1.523</td>
</tr>
<tr>
<td>Pr(L</td>
<td>reject)</td>
<td>0.175</td>
<td>0.094</td>
</tr>
<tr>
<td>No Reject</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5%</td>
<td>1.128</td>
<td>1.269</td>
<td>1.161</td>
</tr>
<tr>
<td>95%</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Pr(L</td>
<td>No Reject)</td>
<td>0.039</td>
<td>0.017</td>
</tr>
<tr>
<td>Difference (Reject - No Reject)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5%</td>
<td>0.586</td>
<td>0.412</td>
<td>0.550</td>
</tr>
<tr>
<td>95%</td>
<td>0.406</td>
<td>0.212</td>
<td>0.150</td>
</tr>
<tr>
<td></td>
<td>0.657</td>
<td>0.575</td>
<td>0.637</td>
</tr>
</tbody>
</table>

Note: 5/95% CI computed using bootstrap block re-sampling at the household level (N=250 Reps); 5% level extended to include 1 if p-value of F-test for presence of private information is less than .05; Bootstrap CI is bias corrected using the non-accelerated procedure in Efron (1982).
Table 1.8: Minimum Pooled Price Ratio (Robustness to $\tau$)

<table>
<thead>
<tr>
<th>Quantile Region: $\Psi_r$</th>
<th>LTC 0-70%</th>
<th>LTC 0-80%</th>
<th>LTC 0-90%</th>
<th>Disability 0-70%</th>
<th>Disability 0-80%</th>
<th>Disability 0-90%</th>
<th>Life 0-70%</th>
<th>Life 0-80%</th>
<th>Life 0-90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reject</td>
<td>1.715</td>
<td>1.715</td>
<td>1.715</td>
<td>2.350</td>
<td>1.954</td>
<td>1.727</td>
<td>1.865</td>
<td>1.727</td>
<td>1.572</td>
</tr>
<tr>
<td>5%</td>
<td>1.538</td>
<td>1.538</td>
<td>1.627</td>
<td>2.216</td>
<td>1.890</td>
<td>1.682</td>
<td>1.626</td>
<td>1.500</td>
<td>1.415</td>
</tr>
<tr>
<td>95%</td>
<td>1.784</td>
<td>1.784</td>
<td>1.782</td>
<td>2.549</td>
<td>2.076</td>
<td>1.817</td>
<td>2.577</td>
<td>2.241</td>
<td>2.110</td>
</tr>
<tr>
<td>No Reject</td>
<td>1.128</td>
<td>1.128</td>
<td>1.128</td>
<td>1.611</td>
<td>1.611</td>
<td>1.611</td>
<td>1.444</td>
<td>1.361</td>
<td>1.281</td>
</tr>
<tr>
<td>5%</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.272</td>
<td>1.243</td>
<td>1.247</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>95%</td>
<td>1.227</td>
<td>1.227</td>
<td>1.227</td>
<td>2.346</td>
<td>2.346</td>
<td>2.089</td>
<td>1.457</td>
<td>1.397</td>
<td>1.286</td>
</tr>
<tr>
<td>Difference</td>
<td>0.586</td>
<td>0.586</td>
<td>0.586</td>
<td>0.739</td>
<td>0.343</td>
<td>0.117</td>
<td>0.421</td>
<td>0.366</td>
<td>0.291</td>
</tr>
<tr>
<td>5%</td>
<td>0.406</td>
<td>0.406</td>
<td>0.485</td>
<td>-0.252</td>
<td>-1.181</td>
<td>-0.582</td>
<td>0.170</td>
<td>0.116</td>
<td>0.115</td>
</tr>
<tr>
<td>95%</td>
<td>0.657</td>
<td>0.657</td>
<td>0.661</td>
<td>1.085</td>
<td>0.695</td>
<td>0.443</td>
<td>1.134</td>
<td>0.950</td>
<td>0.817</td>
</tr>
</tbody>
</table>

Note: 5/95% CI computed using bootstrap block re-sampling at the household level (N=250 Reps); 5% level extended to include 1 if p-value of F-test for presence of private information is less than .05; Bootstrap CI is bias corrected using the non-accelerated procedure in Efron (1982).
LTC For the rejectees, the pooled price ratio reaches a minimum of 1.715 (5/95% CI of [1.575,1.779]) at the mean value of the index, $q = 0.175$, as reported in Table 1.7. This implies private information imposes an implicit tax of 71.5%. The estimates are similar for rejectees with other values of the index, $q$, ranging from 1.681 to 1.730. Together, the results are consistent with Corollary 1.2 provided the rejectees are unwilling to pay a 70% tax for insurance.

For those served by the market, we cannot reject the null hypothesis of no private information, $\hat{F}(p|q) = 1\{p < q\}$ for all $q$. We estimate a minimum pooled price ratio of 1.206 (5/95% CI of [1.00-1.484]) at the mean value of the index $q (q = 0.175)$; estimates range from 1.337 to 1.147 as we vary the index between the 20th and 80th percentile of its distribution ($q = 0.017$ to $q = 0.057$).

Our point estimates are consistent with the presence of trade as long as non-rejectees are willing to pay a 14-34% implicit tax. Finally, the bottom rows of Table 1.7 report the estimated difference between the pooled price ratio for the rejectees relative to the non-rejectees, suggesting we can reject a null hypothesis of smaller minimum pooled price ratios for the rejectees relative to the non-rejectees.69

In sum, our results are consistent with the theory that private information causes rejections as long as rejectees are unwilling to pay implicit taxes of 68-72% and non-rejectees are willing to pay implicit taxes of 10-27%.70 To assess whether this is plausible, we perform two analyses. First, Table 1.9 presents calibrated values of $\frac{u'(w-l)}{u'(w)} - 1$ for values of the coefficient of relative risk aversion (henceforth CRRA) of 1, 2 or 3, and the size of the uninsured drop in consumption of 10%, 15%, and 20%.71 For example, if CRRA is 2 and the nursing home entry is equivalent to a drop in consumption of 15%, then agents are would be willing to pay a 38.4% tax on insurance, rationalizing the observed pattern of trade. Second, the calibrated model of Brown and Finkelstein (2008) suggests individuals are willing to pay roughly a 27-62% markup for existing LTC insurance policies, which easily rationalizes our observed pattern of rejections.72

Disability For the rejectees, we estimate a minimum pooled price ratio of 1.954 (5/95% CI of [1.884,2.032]) at the mean index ($q = 0.441$), which implies a tax rate equivalence of 95.4%. The estimates are similar at the 20th and 50th quantile of $q$ (1.900 and 1.937), and higher at the 80th

---

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

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

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

72 These numbers are not provided directly in Brown and Finkelstein (2008), but can be inferred from Figure 1 and Table 2. Figure 1 suggests the break-even point for insurance purchase is at the 60-70th percentile of the wealth distribution. Table 2 shows this corresponds to individuals being willing to pay a tax of 27-62%. Of course, these estimates are only approximations to $\frac{u'(w-l)}{u'(w)}$, since they consider the willingness-to-pay for a given (non-marginal) LTC policy, not the willingness-to-pay for an $e$-sized transfer which would yield an estimate of $\frac{u'(w-l)}{u'(w)}$. 

58
Table 1.9: Willingness to Pay Calibration

<table>
<thead>
<tr>
<th>Consumption drop</th>
<th>Coeff. Rel. Risk Aversion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>10%</td>
<td>11.1%</td>
</tr>
<tr>
<td>15%</td>
<td>17.6%</td>
</tr>
<tr>
<td>20%</td>
<td>25.0%</td>
</tr>
</tbody>
</table>

quantile (2.282) of the index. The results are consistent with Corollary 1.2 provided the rejectees are unwilling to pay a 90-130% tax for insurance.

For the non-rejectees, we estimate a minimum pooled price ratio of 1.611 (5/95% CI of [1.272,2.391]), implying that the barrier to trade faced by a person with an average observable loss probability among the rejectees is equivalent to a 61% tax on insurance premiums. This ranges from 1.703 to 1.572 as we vary the index \( q \) from its 20th to 80th quantile of its distribution \((q = 0.109 \text{ to } q = 0.197)\). This suggests individuals must be willing to pay a 55-70% tax on insurance premiums in order to facilitate trade in this market. Finally, the estimated differences between the rejectees and non-rejectees are all positive, yet statistically insignificantly different from zero, arguably a result of the imprecise estimation for the non-rejectees.

In sum, our results are consistent with the theory that private information leads to rejections as long as rejectees are unwilling to pay a 90-130% tax for disability insurance and non-rejectees are willing to pay a 55-70% tax. The implied willingness to pay for insurance of 70-90% is consistent with Bound et al. (2004), which calibrates the marginal willingness to pay for an additional unit of disability to be roughly 46-109%. The observed pattern of rejections can also be rationalized if disability yields a loss size of roughly 15-20% of consumption and agents have a coefficient of relative risk aversion of 3.

**Life** For the rejectees, we estimate a minimum pooled price ratio of 1.727 (5/95% CI of [1.527,2.193]) at the mean index, \( q = 0.572 \), indicating a tax rate equivalence of 72.7%. The estimates for other values of the index range from 1.642 at the 20th quantile of \( q \) \((q = 0.572)\) to 2.269 at the 80th quantile of \( q \) \((q = 0.791)\). The results are consistent with Corollary 1.2 as long as the rejectees are unwilling to pay a 65-130% tax for insurance.\(^7\)

\(^7\)See column 6 of Table 1.2. 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 our analysis.

\(^7\)The mapping to a willingness to pay in terms of CRRA preferences and a consumption drop is a bit more abstract and perhaps less useful in our life insurance setting; but if death is equivalent to a 15% consumption drop and CRRA is 3, then individuals would be willing to pay a 62.8% tax, insufficient to sustain trade and consistent with Corollary 1.2.
For those served by the market, we cannot reject the null hypothesis of no private information, as in LTC. We estimate a minimum pooled price ratio of 1.361 (5/95% CI of [1.00, 1.421]), which implies a tax rate of 36.1% for a rejectee with an average observable loss probability. Our estimates for other values of the index, \( q \), range from 1.640 at the 20th quantile (\( q = 0.273 \)) to 1.345 at the 80th quantile (\( q = 0.458 \)). Although some of these point estimates are large, we cannot reject the null hypothesis of a zero tax rate faced by the non-rejectees. The estimated differences between the rejectees and non-rejectees are generally significant at the 5% level, aside from the comparisons involving the point estimate of 1.640 for the 20th percentile of the index for the non-rejectees.

**Choice of \( \tau \)** Table 1.8 presents results for \( \tau = 0.7, 0.8, 0.9 \) in each sample using the mean index value, \( q \), in each sample. Also, Figure 1-6(a-f) plots the estimated pooled price ratios, \( T(p) \), for values of \( p \) less than the estimated 90th quantile of the distribution of private information at varying values of the index, \( q \), in each sample.

For LTC, the minimum of the pooled price ratio occurs at an interior point of the distribution, both for the rejectees and non-rejectees. Thus, our results are not sensitive to the choice of \( \tau \) (in the range where \( \tau \leq 0.9 \)). For disability, the minimum of the pooled price ratio occurs at an interior point for the non-rejectees, but is on the boundary for the rejectees, so the minimum pooled price ratio for the rejectees drops as we increase \( \tau \), as reported in Table 1.8 and shown in Figure 1-6b. Although the estimates for the rejectees fall to 1.727 for \( \tau = 0.9 \) and rise to 2.350 for \( \tau = 0.7 \), they remain quite large across these choices of \( \tau \). For life, the minimum of the pooled price ratio lies
at the boundary for both the rejectees and non-rejectees. For the rejectees, the minimum pooled price ratio falls from 1.727 to 1.572 at \( \tau = 0.9 \) and rises to 1.865 at \( \tau = 0.7 \). For the non-rejectees, our estimates rise to 1.444 at \( \tau = 0.7 \) and fall to 1.281 at \( \tau = 0.9 \). As indicated by the bottom rows of Table 1.8, we can still reject the null hypothesis of a lower minimum pooled price ratio for rejectees relative to non-rejectees at each value of \( \tau \).

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

1.7.2.3 Results for Elicitation Error Distribution

Table 1.10 presents our estimated results for the elicitation error distribution. In general, we find considerable support for the maintained hypothesis that subjective probabilities are noisy and potentially biased measures of agents beliefs. Estimates of the standard deviation of the elicitation error are primarily around 0.3-0.4, with the exception of an estimate of 0.1 for the non-rejectees in disability. Also, we estimate a sizable fraction of focal point respondents in each sample (35-50%).\(^7\)

1.7.2.4 Annuities

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

Our estimated distribution of private information about mortality reveals that, although some agents know that they have a relatively higher than average mortality risk, few agents know that they have an exceptionally lower than average mortality risk. As shown in Figure 1-4c, there are relatively few people, rejected or otherwise, with probabilities below the large mass around 0.15-0.2. Repeating our estimation of the pooled price ratio for \( 1 - P \) (probability of living 10-15 years), Table 1.11 reports a minimum of 1.177 for the life non-rejectees sample (for \( \tau = 0.8 \) and mean index \( q \)), which occurs around 0.2 and is insignificantly different from a zero tax of 1.0. Because there are few people with rejection conditions that have significantly lower probabilities of dying, providing an annuity to the large mass of relatively healthy people does not require preventing the sick from being able to purchase it. By reversing the direction of the incentive constraints, rejections no longer occur.

\(^7\) We do find significant evidence of bias, \( \alpha(X) \), which varies with \( X \). The mean bias by sample is given by the difference of the first two rows of Table 1.3. Also, as shown in Table 1.10, we estimate focal windows around 0.2 in LTC (both rejectees and non-rejectees) and the non-rejectees in Life. This suggests focal responses of 0 correspond to non-focal response ranges of \([0, 0.2]\), responses of 50 correspond to non-focal responses of \([0.2, 0.8]\) and responses of 1 correspond to non-focal responses of \([0.8, 1]\). For rejectees in life and for both rejectees and non-rejectees in disability, we find estimates of the focal window close to zero. Estimates of a focal window of 0 have the simple interpretation that the focal point responders report 50% regardless of their private information.
Table 1.10: Elicitation Error Parameters

<table>
<thead>
<tr>
<th></th>
<th>LTC</th>
<th>Disability</th>
<th>Life</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>No Reject</td>
<td>Reject</td>
<td>No Reject</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.287 (0.035)</td>
<td>0.320 (0.011)</td>
<td>0.305 (0.031)</td>
</tr>
<tr>
<td></td>
<td>0.384 (0.016)</td>
<td>0.427 (0.012)</td>
<td></td>
</tr>
<tr>
<td>s.e.</td>
<td>(0.012)</td>
<td>(0.014)</td>
<td></td>
</tr>
<tr>
<td>Fraction Focal Respondents</td>
<td>0.372 (0.055)</td>
<td>0.496 (0.017)</td>
<td>0.343 (0.023)</td>
</tr>
<tr>
<td></td>
<td>0.386 (0.015)</td>
<td>0.392 (0.014)</td>
<td></td>
</tr>
<tr>
<td>s.e.</td>
<td>(0.024)</td>
<td>(0.018)</td>
<td>(0.028)</td>
</tr>
<tr>
<td>Focal Window</td>
<td>0.179 (0.024)</td>
<td>0.238 (0.018)</td>
<td>0.002 (0.028)</td>
</tr>
<tr>
<td></td>
<td>0.033 (0.028)</td>
<td>0.031 (0.011)</td>
<td></td>
</tr>
<tr>
<td>s.e.</td>
<td>(0.012)</td>
<td>(0.014)</td>
<td></td>
</tr>
</tbody>
</table>

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

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

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

Table 1.11: Minimum Pooled Price Ratio (Annuities)

<table>
<thead>
<tr>
<th>Quantile Region: Ψ,</th>
<th>0-70%</th>
<th>0-80%</th>
<th>0-90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Reject</td>
<td>1.2227</td>
<td>1.1770</td>
<td>1.1523</td>
</tr>
<tr>
<td>5%</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>95%</td>
<td>1.4045</td>
<td>1.2651</td>
<td>1.2651</td>
</tr>
<tr>
<td>Reject</td>
<td>1.405</td>
<td>1.334</td>
<td>1.268</td>
</tr>
<tr>
<td>5%</td>
<td>1.248</td>
<td>1.221</td>
<td>1.227</td>
</tr>
<tr>
<td>95%</td>
<td>1.736</td>
<td>1.720</td>
<td>1.720</td>
</tr>
</tbody>
</table>

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

1.A.1 Proof of No-Trade Theorem

We prove the no-trade theorem in several steps. First, we translate the problem to a maximization problem in utility space. Second, we prove the converse of the theorem directly by constructing an implementable allocation other than the endowment when Condition 1.1 does not hold. Third, we prove the no trade theorem for a finite type distribution. Fourth, we show finite type distributions can approximate solutions to arbitrary distributions, proving the no trade theorem for a general type distribution.

Most of these steps are straightforward. In our opinion, the key theoretical contribution comes in step 3 (Lemma (1.A.5)), where we show that Condition 1.1 implies that a separating allocation cannot improve over a full pooling allocation. Indeed, the ability for insurance companies to offer separating contracts is an important ingredient in previous models of this environment (Spence, 1979; Riley, 1979; Chade and Schlee, 2011).

1.A.1.1 Utility Space

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

Given a utility allocation \( A_U = \{u_L(p), u_{NL}(p)\}_p \), let us denote the slack in the resource constraint by

\[
\Pi(A_U) = \int [w - pl - pc(u_L(p)) - (1-p)c_{NL}(p)] dF(p)
\]

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

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

\[
P1: \max_{\{u_L(p), u_{NL}(p)\}_p} \int [w - pl - pc(u_L(p)) - (1-p)c(u_{NL}(p))] dF(p)
\]

\[\text{ s.t. } \quad pu_L(p) + (1-p)u_{NL}(p) \geq pu_L(\hat{p}) + (1-p)u_L(\hat{p}) \quad \forall p, \hat{p} \in \Psi
\]

\[\quad pu_L(p) + (1-p)u_{NL}(p) \geq pu(w-l) + (1-p)u(w) \quad \forall p \in \Psi
\]

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

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

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

1.A.1.2 Converse

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

**Lemma 1.A.2 (Converse).** Suppose Condition 1.1 does not hold so that there exists $\hat{p} \in \Psi \backslash \{1\}$ such that $\frac{\hat{p}}{1 - \hat{p}} \frac{u'(w - l)}{u'(w)} > \frac{E[P|P \geq \hat{p}]}{1 - E[P|P \geq \hat{p}]}$. Then, there exists an allocation $A^U = \{(\hat{u}_L(p), \hat{u}_{NL}(p))\}$, and a positive measure of types, $\Psi \subset \Psi$, such that

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

and

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

**Proof.** The proof follows by constructing an allocation which is preferable to all types $p \geq \hat{p}$ and showing that the violation of Condition 1.1 at $\hat{p}$ ensures its profitability. Given $\hat{p} \in \Psi$, either $P = \hat{p}$ occurs with positive probability, or any open set containing $\hat{p}$ has positive probability. In the case that $\hat{p}$ occurs with positive probability, let $\hat{\Psi} = \{\hat{p}\}$. In the latter case, note that the function $E[P|P \geq p]$ is locally continuous in $p$ at $\hat{p}$ so that WLOG the no-trade condition does not hold for a positive mass of types. WLOG, we assume $\hat{p}$ has been chosen so that there exists a positive mass of types $\hat{\Psi}$ such that $p \in \hat{\Psi}$ implies $p \geq \hat{p}$. Then, for all $p \in \hat{\Psi}$, we have $\hat{\Psi} \subset \Psi$ such that

$$\frac{p}{1 - p} \frac{u'(w - l)}{u'(w)} > \frac{E[P|P \geq \hat{p}]}{1 - E[P|P \geq \hat{p}]} \quad \forall p \in \hat{\Psi}$$
Now, for $\varepsilon, \eta > 0$, consider the augmented allocation to types $p \in \hat{\Psi}$:

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

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

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

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

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

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

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

which is strictly positive if and only if

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

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

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

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

This shows that Condition 1.1 is necessary for the endowment to be the only implementable allocation. \qed
1.A.1.3 Lemmas

Here, we prove two useful lemmas. First, we show that if Condition 1.1 holds, then the MRS is bounded by the pooled price ratio in the relevant quadrant of allocations.

**Lemma 1.A.3.** Suppose Condition 1.1 holds. Then for all \( c_L, c_{NL} \in [w - l, l] \), we have

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

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

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

**Proof.** Since \( u'(c) \) is decreasing in \( c \), we have \( \frac{u'(c_L)}{u'(c_{NL})} \leq \frac{u'(w-l)}{u'(w)} \). Therefore, the result follows immediately from Condition 1.1. The strict inequality follows from strict concavity of \( u(c) \). \( \square \)

**Lemma 1.A.4.** In any solution to \( P1 \), we have \( c_L(p) \geq w - l \) and \( c_{NL}(p) \leq w \).

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

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

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

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

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

We now prove the theorem in two steps. First, we prove the result for a finite type distribution. We then pass to the limit to cover the case of arbitrary distributions.
1.A.1.4 Finite Types

To begin, suppose that $\Psi = \{p_1, \ldots, p_N\}$. We first show that condition 1.1 implies that the solution to $P1$ is a pooling allocation which provides the same allocation to all types.

**Lemma 1.A.5.** Suppose $\Psi = \{p_1, \ldots, p_N\}$ and that condition 1.1 holds (note that this requires $p_N = 1$). Then, the solution to $P1$ is a full pooling allocation: there exists $u_L, u_{NL}$ such that $(u_L(p), u_{NL}(p)) = (\bar{u}_L, \bar{u}_{NL})$ for all $p \in \Psi \setminus \{0, 1\}$, $u_L(1) = \bar{u}_L$, $u_{NL}(0) = \bar{u}_{NL}$.

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

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

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

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

Now, we construct the utility allocation $A^U_i$ by

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

Since the IR constraint binds for type $\hat{p}$, we know that there exists $\hat{t}$ such that $A^U_i = A^U$. By Lemma 1.A.4, $\hat{t} > 0$ and $A^U_i$ satisfies IC and IR for any $t \in [0, \hat{t} + \eta]$ for some $\eta > 0$. Since profits are maximized at $t = \hat{t}$ and since the objective function is strictly concave, it must be the case that

$$\frac{d\Pi(A^U_i)}{dt}\bigg|_{t=\hat{t}} = 0$$

68
where
\[
\frac{d\Pi(A^U)}{dt}_{|t=t} = \int_{P\in J} \left[ pc'(u^*_L(p)) - (1-p) c'(u^*_NL(p)) \frac{\hat{p}}{1-\hat{p}} \right] dF(p)
\]
Re-arranging and combining these two equations, we have
\[
\frac{\hat{p}}{1-\hat{p}} \frac{u'(c(u^*_L(\hat{p})))}{u'(c(u^*_NL(\hat{p})))} = \frac{E[P|P \geq \hat{p}]}{1 - E[P|P \geq \hat{p}]}
\]
which, by strict concavity of \(u\), implies
\[
\frac{\hat{p}}{1-\hat{p}} \frac{u'(w-l)}{u'(w)} > \frac{E[P|P \geq \hat{p}]}{1 - E[P|P \geq \hat{p}]}
\]
which contradicts Condition 1.1.

Now, suppose that the IR constraint does not bind for any member of \(J\). Then, clearly the IC constraint for type \(\hat{p}\) must bind, otherwise profit could be increased by lowering the utility provided to members of \(J\). So, construct the utility allocation \(B^U_{\epsilon}\) to be
\[
(u^*_L(p), u^*_NL(p)) = \begin{cases} (u^*_L(\hat{p}) - \varepsilon, u^*_NL(\hat{p}) + \frac{\hat{p}}{1-\hat{p}} \varepsilon) & \text{if } p \geq \hat{p} \\ (u^*_L(p), u^*_NL(p)) & \text{if } p < \hat{p} \end{cases}
\]
so that \(B^U_{\epsilon}\) consists of allocations equivalent to \(A^U\) except for \(p \in J\). By construction, \(B^U_{\epsilon}\) is IR for any \(\epsilon\). Moreover, because of single crossing and because types are separated (finite types), \(B^U_{\epsilon}\) continues to be IC and IR for \(\epsilon \in (-\eta, \eta)\) for some \(\eta > 0\) sufficiently small. Therefore, we must have \(\frac{d\Pi(B^U_{\epsilon})}{d\epsilon}_{|\epsilon=0} = 0\), which implies
\[
\frac{d\Pi(B^U_{\epsilon})}{d\epsilon}_{|\epsilon=0} = \int_{P\in J} \left[ pc'(u^*_L(\hat{p})) - (1-p) c'(u^*_NL(\hat{p})) \frac{\hat{p}}{1-\hat{p}} \right] dF(p)
\]
\[
= \Pr\{p \in J\} \left[ \frac{E[P|P \geq \hat{p}]}{u'(c(u^*_L(\hat{p})))} - \frac{1}{1 - E[P|P \geq \hat{p}]} \right] \frac{1}{u'(c(u^*_NL(\hat{p})))} \frac{\hat{p}}{1-\hat{p}}
\]
\[
= \Pr\{p \in J\} \frac{(1 - E[P|P \geq \hat{p}])}{u'(c(u^*_L(\hat{p})))} \left[ \frac{E[P|P \geq \hat{p}]}{1 - E[P|P \geq \hat{p}]} \right] \frac{1}{u'(c(u^*_NL(\hat{p})))} \frac{\hat{p}}{1-\hat{p}}
\]
\[
= 0
\]
which implies
\[
\frac{\hat{p}}{1-\hat{p}} \frac{u'(c(u^*_L(\hat{p})))}{u'(c(u^*_NL(\hat{p})))} = \frac{E[P|P \geq \hat{p}]}{1 - E[P|P \geq \hat{p}]}
\]
which, by strict concavity of \(u\), implies
\[
\frac{\hat{p}}{1-\hat{p}} \frac{u'(w-l)}{u'(w)} > \frac{E[P|P \geq \hat{p}]}{1 - E[P|P \geq \hat{p}]}
\]
which contradicts Condition 1.1. Therefore, if Condition 1.1 holds, the only possible solution to
P1 is a full pooling allocation.

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

**Lemma 1.A.6.** Suppose Condition 1.1 holds. Then, the only possible full-pooling solution to P1 is $E^U$.

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

$$(u^*_L, u^*_N) = (u^*_L + (1 - t)(u(w - l) - u^*_L), u^*_N + (1 - t)(u(w) - u^*_N))$$

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

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

where $\frac{p_1}{1 - p_1}$ comes from the fact that we can parameterize the iso-utility curve of the $p_1$ type by $u_L - \tau, u_N + \tau, u^*_L$. But re-arranging the equation, we have

$$\frac{d\Pi(C^U_t)}{dt}|_{t=0} = -E[P|P \geq p_1] \frac{1}{u'(w - l)} + (1 - E[P|P \geq p_1]) \frac{1}{u'(w)} \frac{p_1}{1 - p_1}$$

$$= \frac{1 - E[P|P \geq p_1]}{u'(W - L)} \left( -E[P|P \geq p_1] \frac{1}{u'(w - l)} + E[P|P \geq p_1] \frac{p_1}{1 - p_1} \right) < 0$$

which yields a contradiction of Condition 1.1 at $p = p_1$. Therefore, we have shown that if $\Psi$ is finite, then if Condition 1.1 holds, the only possible allocation is the endowment. It only remains to show that this property holds when $\Psi$ is not finite.

1.A.1.5 Arbitrary Distribution

If $F(p)$ is continuous or mixed and satisfies the no-trade condition, we first show that $F$ can be approximated uniformly by a sequence $F_n$ of finite support distributions on $[0, 1]$, each of which satisfy the no-trade condition.
Lemma 1.A.7. Let $P$ be any random variable on $[0,1]$ with c.d.f. $F(p)$. Then, there exists a sequence of random variables, $P_N$, with c.d.f. $F^N(p)$, such that $F^N \rightarrow F$ uniformly and

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

Proof. Since $F$ is increasing, it has at most a countable number of discontinuities on $[0,1]$. Let $D = \{\delta_i\}$ denote the set of discontinuities and WLOG order these points so that $\lim_{i \to +0} F(\delta_i) - \lim_{i \to -0} F(\delta_i)$ is decreasing in $i$ (so that $\delta_1$ is the point of largest discontinuity). Then, the distribution $F$ is continuous on $\Psi \setminus D$. For any $N$, let $\omega_N$ denote a partition of $[0,1]$ given by $2^N + \min\{N,|D|\} + 1$ points equal to $\frac{j}{2^N}$ for $j = 0,\ldots,2^N$ and $\{\delta_i|i \leq N\}$. We write $\omega_N = \{p_j^N\}_{j=1}^{2^N+\min\{N,|D|\}+1}$. Now, define $\hat{F}^N : \omega_N \rightarrow [0,1]$ by

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

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

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

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

$$\lambda_N = \max\{\lambda|E[P^N_\lambda|P^N_\lambda \geq p] \geq E[P|P \geq p] \quad \forall p\}$$

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

$$F^N(p) = \lambda_N \hat{F}^N(p)$$

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

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

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

71
Moreover, because \( \lambda^*_N \) is bounded, it has a convergent subsequence, \( \lambda_{n_k} \to \lambda^* \). Therefore,

\[
E \left[ P_{\lambda^*_k}^N | P_{\lambda^*_k}^N \geq q \right] \to E \left[ P_{\lambda^*}^\cdot | P_{\lambda^*}^\cdot \geq q \right]
\]

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

\[
E \left[ P_{\lambda_{n_k}}^N | P_{\lambda_{n_k}}^N \geq q \right] \to E \left[ P_{\lambda^*}^\cdot | P_{\lambda^*}^\cdot \geq q \right]
\]

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

\[
E \left[ P_{\lambda_{n_k}}^N | P_{\lambda_{n_k}}^N \geq \bar{p}_N \right] \to E \left[ P | P \geq \bar{p}_N \right]
\]

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

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

\[
E \left[ P^N_k | P^N_k \geq p \right] \geq E \left[ P | P \geq p \right] \quad \forall p
\]

and \( F^N_k(p) \) converges uniformly to \( F(p) \).

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

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

\[
\Pi(\hat{A}_N|F_N) \leq \Pi(E|F_N) = 0 \quad \forall N
\]

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

\[
\Pi(\hat{A}|F) \leq 0
\]

Which yields a contradiction that \( \hat{A} \) was the optimal solution (which required \( \Pi(\hat{A}|F) > 0 \)) and concludes the proof.
1.A.2 Informal Insurance

Consider the following modification to the standard model which allows agents’ privately known types to affect their demand for informal insurance, such as savings. Suppose agents live for two periods. They are endowed with non-stochastic wealth \( w \) in both periods. In the second period, they face a potential loss of size \( l \) which occurs with privately known probability \( p \). There is no discounting and agents can store resources in period 1 for use in period 2. Agents know \( p \) before making a decision about how much to save in period 1. Utility is given by

\[
u(c_1(p)) + pu(c_L(p)) + (1 - p) u(c_{NL}(p))
\]

where \( c_1 \) is consumption in period 1, \( c_L \) (\( c_{NL} \)) is consumption in period 2 in the event of (no) loss. Insurance companies can offer menus contracts \( \{ \tau_1(p), \tau_2(p), b(p) \} \) in which agents pay \( \tau_1 \) in the first period, \( \tau_2 \) in the second period if they do not experience the loss, and receive benefits \( b \) if they do not experience the loss in the second period. Contracts are offered in the first period prior to the agents’ saving decision. Thus, the constraints facing the agent after choosing a contract intended for type \( p \) are given by

\[
\begin{align*}
c_1 + s + \tau_1(p) & \leq w \\
c_L + b(p) & \leq w + s \\
c_{NL} + \tau_2(p) & \leq w + s
\end{align*}
\]

**No Insurance** Suppose there are no insurance contracts. Then an agent of type \( p \) chooses \( s(p) \) to maximize utility, so that \( s(p) \) is defined from the first order condition

\[
u'(w - s(p)) = p u'(w - l + s(p)) + (1 - p) u'(w + s(p))
\]

where it is easy to verify that \( s(p) \) is increasing in \( p \), \( s(1) = \frac{1}{2} \) and \( s(0) = 0 \). Therefore, the ratio

\[
\frac{u'(w - l + s(p))}{u'(w + s(p))}
\]

is decreasing in \( p \) (and equals 1 when \( p = 1 \)).

**No-Trade Condition** Consider an insurance company that attempts to provide insurance. The most valuable infinitesimal transfer is to move consumption in the event of not experiencing the loss in the second period to consumption in the event of experiencing the loss. The valuation of this first unit of infinitesimal transfer is given by the MRS

\[
\frac{p}{1 - p} \frac{u'(w - l + s(p))}{u'(w + s(p))}
\]
If no other insurance contract is offered, then any contract offered to type $p$ would be preferred by all types $P \geq p$. Thus, the cheapest cost of providing an infinitesimal transfer to which type $p$ is indifferent is given by \( \frac{E[P|P \geq p]}{1 - E[P|P \geq p]} \). Provided the economy remains well-behaved (which needs proof, but should be true), the no trade condition is

\[
\frac{p}{1 - p} \frac{u'(w - l + s(p))}{u'(w + s(p))} \leq \frac{E[P|P \geq p]}{1 - E[P|P \geq p]} \quad \forall p < 1
\]

or

\[
\frac{u'(w - l + s(p))}{u'(w + s(p))} \leq T(p) \quad \forall p < 1
\] (1.11)

In contrast to Corollary 1.2, demand for an infinitesimal amount of insurance (given by \( \frac{u'(w - l + s(p))}{u'(w + s(p))} \)) is now a function of type, $p$. The ability to save distorts the demand for insurance differently for high versus low risk types. Savings, which transfers resources to all states of the world in the second period, is a better substitute for insurance for people who have a high probability of experiencing the loss. For people with a low probability of a loss, savings more often transfers resources to the no-loss state of the world in the second period, which has relatively low marginal utility.

**Empirical Tests Robust to Endogenous Savings** Equation (1.11) motivates testing a comparative static in the pooled price ratio, $T(p)$, at each $p$, as opposed to testing a comparative static in the minimum pooled price ratio. These tests are shown in Figures 1-6(a-f), which plot the pooled price ratio, $T(p)$, at each $p$ for the rejectees and non-rejectees. Figure 1-6(a-c) focuses on the pooled price ratio at the mean index, $q$, in each sample, and Figure 1-6(d-f) report results for the 20th, 50th and 80th quantile of the index, $q$, in each sample. We truncate the graphs at the 90th quantile of the distribution of $P$ to avoid extremal value issues discussed in Section 1.4.2.

Consistent with the hypothesis that private information leads to rejections, we generally find larger values of $T(p)$ for the rejectees at each value of $p$ and for most values of the index, $q$. 
1.B Empirical Methodology Appendix

1.B.1 Properties of the Lower Bound Estimator

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


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

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

$$Q_P(\alpha) = \inf_q \{q| Pr\{P < q\} \geq \alpha\}$$

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

$$Q_{P_Z}(\alpha) = \inf_q \{q| Pr\{P_Z < q\} \geq \alpha\}$$

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

$$e(\alpha) = Q_P(\alpha) - Q_{P_Z}(\alpha)$$

(1.12)

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

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

**Proposition.** The following conditions hold

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

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

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

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

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

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

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

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

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

The fifth condition relates $E[m(P)]$ to the quantity that characterizes the barrier to trade, $\inf_p T(p)$. Using a Holder inequality, this condition shows that $E[m(P)]$ is monotonically related to an upper bound on $\inf_p T(p)$. Notice that the RHS of the expression in the fifth condition is increasing in both $E[m(P)]$ and $\Pr\{L\}$, provided $E[P(1 - P)]$, and $E[(P - \Pr\{L\})m(P)]$ remain roughly constant. Thus, smaller values of $E[m(P)]$ lead to smaller upper bounds on the minimum pooled price ratio. But also, smaller values of the mean risk, $\Pr\{L\}$, lead to smaller upper bounds on the minimum pooled price ratio. Thus, since rejectees have larger values of $\Pr\{L\}$, it could very well be the case that this lower bound is smaller for non-rejectees even if $E[m(P)]$ is larger for non-rejectees. In this sense, our lower bound test is a potentially overly restrictive test of the implications of the theory. Since we nonetheless find larger values of $E[m(P)]$ for rejectees, we do not discuss this potential bias in detail in the text; it only renders our empirical findings to be even greater support for the theory that private information leads to rejections.
1.B.1.1 Proof of Proposition

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

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

so that we have the expansion

$$E[m(P)] = \int_0^1 [E_\alpha [Q_P(\tilde{\alpha}) - Q_P(\alpha)|\tilde{\alpha} \geq \alpha]] d\alpha$$

$$= \int_0^1 \frac{1}{1-\alpha} \left[ \int_{\tilde{\alpha} \geq \alpha} [Q_P(\tilde{\alpha}) - Q_P(\alpha) d\tilde{\alpha}] \right] d\alpha$$

$$= \int_0^1 \int_{\tilde{\alpha} \geq \alpha} Q_P(\alpha) d\tilde{\alpha} d\alpha - E[P]$$

$$= \int_0^1 Q_P(\tilde{\alpha}) \int_{\tilde{\alpha}}^1 \frac{1}{1-\alpha} d\tilde{\alpha} d\alpha - E[P]$$

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

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

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

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

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

Now, using part (1), we can write

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

$$= \int_0^1 [Q_P(\alpha) - Q_{P_Z}(\alpha)] \int_0^\alpha \frac{1}{1-\alpha} d\alpha d\alpha$$

$$= \int_0^1 \int_0^\alpha [Q_P(\alpha) - Q_{P_Z}(\alpha)] \frac{1}{1-\alpha} d\alpha d\alpha$$

$$= \int_0^1 \left( \int_0^\alpha [Q_P(\alpha) - Q_{P_Z}(\alpha)] d\alpha \right) \frac{1}{1-\alpha} d\alpha$$

$$\geq 0$$

where the last inequality follows from the fact that $\int_0^1 [Q_P(\alpha) - Q_{P_Z}(\alpha)] d\alpha \geq 0$ for all $\tilde{\alpha}$ because $P$ is a mean-preserving spread of $P_Z$. 

77
Part (4) follows from part (1) and the definition of \( e(\alpha) \).

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

\[
T(p) = \frac{p + m(p)}{1 - p - m(p)} \frac{1 - p}{p}
\]

which can be re-written as

\[
m(p) \frac{1}{\text{t}(p)} = p (1 - p) - pm(p)
\]

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

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

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

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

So that

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

so that

\[
\inf \frac{t(p)}{E \left[ P (1 - P) \right] - E \left[ (P - E[P]) m(P) \right] - E[m(P)] * E[P]} \leq E[m(P)]
\]

and thus

\[
\inf \frac{T(p)}{E \left[ P (1 - P) \right] - E \left[ (P - E[P]) m(P) \right] - E[m(P)] * E[P]} \leq 1 + \frac{E[m(P)]}{E \left[ P (1 - P) \right] - E \left[ (P - E[P]) m(P) \right] - E[m(P)] * E[P]}
\]

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

1.B.1.2 "Tight" lower bound

Here, we show that the lower bound is “tight” in the sense that there exists a joint distribution of \( L, P, \) and \( Z \) satisfying assumptions 1 and 2 such that \( P = P_Z \). This follows relatively trivially. For any elicitation \( Z \), assume that \( P = \Pr \{L|X, Z\} \). Then let \( e = Z - \Pr \{L|X, Z\} \). Then agents’ report \( Z = P + e \) but have beliefs given by \( \Pr \{L|X, Z\} \).
1.B.1.3 Measurement Error Example

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

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

$$E[m_z(P_z)] = \lambda E[m(P)]$$

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

1.B.2 Semiparametric Identification

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

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

- $L$ is a binary random variable (realizations in $\{0, 1\}$, indicating the event of experiencing a loss

- The joint density of $Z$ and $L$ is observed and given by the p.d.f./p.m.f. $f_{L,Z}(l,z)$ with conditional distributions $f_{L|Z}$, $f_{Z|L}$, and marginal distribution $f_z \in D_Z$ for some domain $D_Z$ We assume $Z$ is continuously distributed over $[0, 1]$.

- The variable $P$ is unobserved and continuously distributed with density $f_P(p) \in D_P$ for some domain $D_P$, where we denote the true value by $f^*_P(p)$. We assume $D_P$ is closed under multiplication by $p$, so that $f_P(p) \in D_P$ implies $p f_P(p) \in D_P$.

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

\[ Pr \{ L | Z, P \} = Pr \{ L | P \} = P \]

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

\[
\begin{align*}
  f_{L,Z}(L, Z) &= \int_0^1 f_{L,Z|P}(L, Z | P = p) f_P(p) \, dp \\
  &= \int_0^1 (\Pr \{ L | Z, P = p \})^L (1 - \Pr \{ L | Z, P = p \})^{1-L} f_{Z|P}(Z | P = p; \theta^*) f_P^*(p) \, dp \\
  &= \int_0^1 p^L (1 - p)^{1-L} f_{Z|P}(Z | P = p; \theta^*) f_P^*(p) \, dp
\end{align*}
\]

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

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

\[
\begin{align*}
  [H_\theta(f_P)](z) &= \int_0^1 f_P(p) f_{Z|P}(z | P = p; \theta) \, dp \\
  H_\theta \text{ is linear in } f_P. \text{ Therefore, we can impose standard invertibility conditions on } f_{Z|P}.
\end{align*}
\]

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

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

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

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

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

$$f_{P|L}(p|L = 1) \Pr \{L = 1\} = f_{P,L}(p,1) = \Pr \{L = 1|P = p\} f_P(p) = pf_P(p)$$

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

$$E \left[ P^N | L = 1 \right] \Pr \{ L \} = E \left[ P^{N+1} \right] \quad \forall N \geq 0 \tag{1.13}$$

which provides a simple relationship between the moments of $P$ given $L = 1$ and the unconditional distribution of $P$. Equation 1.13 provides an infinite set of moment conditions which aid in identification of $\theta$ and $f_P$.

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

$$E \left[ \left( \hat{H}_{\theta^*}^{-1}(Z) \right)^N | L = 1 \right] \Pr \{ L = 1 \} = E \left[ \left( \hat{H}_{\theta^*}^{-1}(Z) \right)^{N+1} \right] \quad \forall N \geq 0 \tag{1.14}$$

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

Because $\theta$ is finite-dimensional, the model is generally over-identified. Intuitively, equation 1.14 for $N = 0$ provides identification of the mean of the elicitation error

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

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

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

so that, intuitively, the RHS of the equation varies with the dispersion in the error, holding the mean of the error constant. We recognize that this intuition is fairly abstract because it relies on properties of the operator $H_{\theta}^{-1}$, which is a difficult object to know a priori. We proceed on two fronts. First, we provide a formal proof that a close analogue to our specification in Section 1.7 for which the inverse operator has well-known properties and is identified using only equations $N = 0$ and $N = 1$ of equation 1.14 (so that the moments $N > 1$ provide a theoretical over-identification test). Second, since our specification in Section 1.7 does not have such a well-known inverse operator, we provide Monte Carlo tests of our specification. This allows us not only to confirm identification, but also assess the robustness of our results to various possible mis-specifications of the true elicitation error distribution, $f(Z|P)$.  

81
1. B. 2. 1 Identification without censoring

Here, we consider an analogue to our model in which non-focal elicitation are not censored on [0, 1]. For this specification, the nonlinear inverse problem has well-known properties and \( \theta^* \) and \( f_p^* \) are identified only the observed density, \( f_Z \) and equation 1.14 for \( N = 0 \) and \( N = 1 \), leaving equations \( N > 1 \) as over-identifying conditions.

In particular, suppose \( Z \) is distributed \((1 - \lambda) N (P + \alpha, \sigma^2) + \lambda OP (P + \alpha, \sigma^2, \kappa)\), where \( OP (P + \alpha, \sigma^2, \kappa) \) is an ordered probit with variance \( \sigma^2 \), latent mean \( P + \alpha \), and cutoff regions \([0, \kappa], (\kappa, 1 - \kappa)\), and \([1 - \kappa, 1]\) corresponding to values \( Z = 0, 0.5, 1 \). Note that our specification in Section 1.7 is similar but assumes non-focal values follow a censored normal, \( CN (P + \alpha, \sigma^2) \), as opposed to a normal, \( N (P + \alpha, \sigma^2) \), which captures the elicitations lie in [0, 1].

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

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

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

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

So, for \( N = 0 \), we have the equation

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

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

For \( N = 1 \), we have the equation

\[
E \left[ \left( \hat{P} (Z; \alpha^*, \sigma^*) \right) | L = 1 \right] \Pr \{L = 1\} = E \left[ \left( \hat{P} (Z; \alpha^*, \sigma^*) \right)^2 \right]
\]
Now, notice that the LHS does not vary with $\sigma$. Moreover, the RHS is monotonically decreasing in $\sigma$, since $\sigma_1 < \sigma_2$ implies that $\hat{P}(Z; \alpha, \sigma_1)$ is a mean preserving spread of $\hat{P}(Z; \alpha, \sigma_2)$. Thus, the equation for $N = 1$ identifies $\sigma^*$.

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

This example with uncensored non-focal point values is identical to our specification in Section 1.7, except that we use a censored normal, as opposed to normal distribution, to take into account the fact that elicitations are restricted to the interval, $[0,1]$. The impact of such censoring on the quality of our estimation is difficult to assess theoretically; we thus turn to Monte Carlo evidence to verify the performance of our estimation strategy.

1.B.2.2 Monte Carlo Results

This section presents Monte Carlo analysis of our estimator for the distribution of private information. First, we verify that our approach works well under correct model specification for the elicitation error parameters. Second, we assess the impact of mis-specification of the distribution of elicitation error. Throughout this section, we assume that $F(p)$ is a censored normal distribution with mean 0.3 and standard deviation 0.1. Our first simulation assumes $Z$ is drawn from a censored normal distribution with mean $P + 0.03$ and standard deviation 0.2. With probability 0.4, $Z$ is a focal point value of 0, 0.5, or 1, drawn from an ordered probit distribution with mean $P + 0.03$, standard deviation 0.2, and cutoff regions $[0, 0.3]$, $(0.3, 0.7)$, $[0.7, 1]$. In Figure 1-7a, we present the true c.d.f. of private information, along with the median, 5%, and 95% estimates from $N = 100$ Monte Carlo simulations (of a sample size of 2,000) where we estimate the distribution using a mixture of 2 beta distributions, as in our empirical analysis above. As the figures show, our estimation approach yields unbiased estimates.

Now we consider the impact of mis-specification of the elicitation error. To do so, we assume
that the latent $Z$ is drawn from a mixture of two normals, allowing us to incorporate skewness and kurtosis in the distribution. First, we assume with probability 0.4 the latent $Z$ is drawn from a normal with mean $P + 0.03$ and standard deviation 0.2 (as before). But with probability 0.6, the latent $Z$ is drawn from a normal with mean $P - 0.05$ and standard deviation 0.4. Focal values (again we assume 40% focal values) are then generated from this latent $Z$ with the same cutoff regions, $[0, 0.3], (0.3, 0.7), [0.7, 1]$, and non-focal values are generated as the censored portion of this distribution on $[0, 1]$. The Monte Carlo results are presented in Figure 1-7b. As we can see, our estimation perhaps introduces a slight median bias towards a less dispersed distribution of private information, but performs quite well given this substantial mis-specification.

Finally, we assess the robustness to excess kurtosis in the distribution of $Z$. We assume $Z$ is again drawn from a mixture of normals, but this time assume these normals have the same mean of 0.03. With probability 0.5, the standard deviation is 0.2 and with probability 0.5 the standard deviation is 0.05. We assume 40% focal responses with the same cutoffs regions of $[0, 0.3], (0.3, 0.7), [0.7, 1]$. Figure 1-7c presents the Monte Carlo results. As we can see, our estimation performs quite well (better than the skewed estimation) despite the mis-specification. In short, our estimation procedure appears robust to alternative specifications for $f(Z|P)$ which relax normality by including skewness and kurtosis.

### 1.C Rejections Appendix: Selected Pages from Genworth Financial Underwriting Guidelines

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

PROVIDED BY THE GENWORTH UNDERWRITING DEPARTMENT

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

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

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

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

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

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

We value our relationship with you and look forward to providing high quality service and underwriting for you and your clients.
UNINSURABLE CONDITIONS
Acquired Immune Deficiency Syndrome (AIDS)
ADL limitation, present
AIDS Related Complex (ARC)
Alzheimer’s Disease
Amputation due to disease, e.g., diabetes or atherosclerosis
Amyotrophic Lateral Sclerosis (ALS) , Lou Gehrig’s Disease
Ascites present
Ataxia, Cerebellar
Autonomic Insufficiency (Shy-Drager Syndrome)
Autonomic Neuropathy (excluding impotence)
Behçet’s Disease
Binswanger’s Disease
Bladder incontinence requiring assistance
Blindness due to disease or with ADL/IADL limitations
Bowel incontinence requiring assistance
Buerger’s Disease (thromboangiitis obliterans)
Cerebral Vascular Accident (CVA)
Chorea
Chronic Memory Loss
Cognitive Testing, failed
Cystic Fibrosis
Dementia
Diabetes treated with insulin
Dialysis, Kidney (Renal)
Ehlers-Danlos Syndrome
Forgetfulness (frequent or persistent)
Gangrene due to diabetes or peripheral vascular disease
Hemiplegia
Hoyer Lift
Huntington’s or other forms of Chorea
Immune Deficiency Syndrome
Korsakoff’s Psychosis
Leukemia-except for Chronic Lymphocytic Leukemia (CLL) and Hairy Cell Leukemia (HCL)
Marfan’s Syndrome
Medications:
  Antabuse (disulfiram)
  Aricept (donepezil HCl)
  Campral (acamprosate calcium)
  Cognex ( tacrine)
  Depade (naltrexone)
  Exelon (rivastigmine)
  Hydergine (ergoloid mesylate)
  Namenda (memantine)
  Razadyne (galantamine hydrobromide)
  Reminyl (galantamine hydrobromide)
  ReVia (naltrexone)
  Vivitrol (naltrexone)
Memory Loss, chronic
Mesothelioma
Multiple Sclerosis (MS)
Muscular Dystrophy (MD)
Myelofibrosis
Organ Transplants, except kidney transplants
Organic Brain Syndrome (OBS)
Oxygen use except if used for headaches or sleep apnea
Paralysis/Paraplegia
Parkinson's Disease
Pneumocystis Pneumonia
Polyarteritis Nodosa
Postero-Lateral Sclerosis
Quad Cane use
Quadriplegia
Senility
Spinal Cord Injury with ADL/IADL limitations
Stroke (CVA)
Surgery scheduled or anticipated (except cataract surgery under local anesthesia)
Takayasu's Arteritis
Thalassemia Major
Total Parenteral Nutrition (TPN) for regular or supplementary feeding or administration of medication
Waldenstrom's Macroglobulinemia
Walker use
Wegener's Granulomatosis
Wernicke-Korsakoff Syndrome
Wheelchair use
Wilson's Disease
Chapter 2

Private Information and Unemployment Insurance

2.1 Introduction

Why is there not a thriving private market for unemployment insurance (UI) in the US? The onset of unemployment leads to drops in consumption and significant welfare losses (Gruber (1997), Browning and Crossley (2001), Chetty (2008)). The government provides some unemployment benefits, but recent literature suggests individuals would prefer to purchase additional insurance (Chetty (2008)). So why isn’t there a private UI market just like there exists other insurance markets, such as private health, disability, and life insurance?¹

This paper argues that private information about future unemployment incidence prevents the existence of a market for private unemployment insurance. We make two related claims. First, we argue that private information prevents the existence of a market for additional insurance beyond what is currently provided by the government. Second, under additional assumptions we provide evidence that private information would prevent the existence of a private market even if the government stopped providing benefits.

We develop and test these hypotheses using a model of unemployment risk with two key features: people may have private information about their future unemployment incidence and people may suffer a moral hazard problem, so that insurance can increase their likelihood of unemployment. We characterize when companies would be willing to sell insurance in this environment. We show that a private market cannot exist unless someone is willing to pay the pooled cost of those with higher probabilities of unemployment in order to obtain a small amount of insurance. This pooled cost is derived from the distribution of probabilities that result from the agents’ choices when faced with their endowment (which may incorporate existing sources of informal or government insurance).

¹There are no regulations preventing the sale of private UI. Over the years, many companies have tried to sell private unemployment insurance, but all have failed. Currently, IncomeAssure is the latest company and only to attempt it, starting sales in mid-2011.
When this no-trade condition holds, an insurance company cannot profitably sell unemployment insurance at any price because it would be too heavily adversely selected to deliver a positive profit.

We then test whether this no-trade condition can explain why there does not exist a private unemployment insurance market in the US for additional insurance, beyond what is provided through the government. We derive an empirical test of the no trade condition that depends on the distribution of beliefs about future unemployment incidence and estimates of the markup individuals are willing to pay for a small amount of additional unemployment insurance. We identify private information and the distribution of beliefs using subjective probability elicitation about future unemployment incidence, following an approach developed in Hendren (2011). We then show existing estimates from the literature on optimal UI (e.g. Gruber (1997), Chetty (2008)) provide estimates of the markup individuals would be willing to pay for additional unemployment insurance.

Our empirical results suggest private information prevents the existence of a private unemployment insurance market in the US. We first show that agents’ subjective probability elicitation are predictive of future unemployment spells, conditional on a rich set of demographic, health, and employment characteristics. To the best of our knowledge, this provides the first empirical evidence that private information would affect a private unemployment insurance market. Second, we quantify the distribution of beliefs about future unemployment incidence. We show that individuals must be willing to pay an or implicit tax in excess of 160% over actuarially fair premiums, in order for a market to exist. We contrast this with existing estimates of this willingness to pay for additional unemployment insurance. These estimates range from 30%-50% (Gruber (1997)) to 60% (Chetty (2008)). Thus, we conclude that private information can explain the absence of a market for additional unemployment insurance beyond what is currently provided by the government.

Finally, we address whether a private market would arise if the government were to remove or reduce UI benefits. Here, we use estimates how government UI generosity affects the magnitude of consumption drops upon unemployment, provided in Gruber (1997). Extrapolating to a world with no government UI provision suggests the markup individuals would be willing to pay for private UI would roughly double, reaching 50-90% depending on the coefficient of relative risk aversion. Under a benchmark assumption that the implicit tax imposed by private information would not change as a result of the reduction in government UI, the demand estimates continue to fall well below estimates of the markup people would need to be willing to pay of at least 160%. Thus, our estimates suggest that if the government stopped providing UI benefits, a private market would not arise.

Our paper is related to several strands of literatures. There is a long literature in public economics analyzing the “optimal” level of government-provided unemployment insurance (Baily

---

2 The impact of reducing benefits on the implicit tax imposed by private information is theoretically ambiguous; an important direction for future work is to provide an empirical methodology to estimate how it varies with the generosity of government insurance.
Yet there is no existing explanation for what market failures, if any, provide a role for the government provision of UI in the first place. We show private information about unemployment risk provides a wedge between the allocations that can be implemented by a government (which has access to taxes and potentially mandates to mitigate the selection problem) versus private markets which cannot force people to buy insurance. This provides a potential rationale for government intervention in UI. 3

Our paper also relates to a large theoretical literature studying the impact of private information on the workings of insurance markets. Our no trade result is similar to the notion of market unraveling in Akerlof (1970), who shows that a market unravels when its demand curve lies everywhere below its average cost curve. Hendren (2011) extends this result to the case when insurers can offer an endogenous menu of contracts and derives a no trade condition almost identical to the one derived in this paper. The crucial difference is that, in this paper, we allow agents to have a moral hazard problem. As a result, we show that moral hazard alone cannot shut down a market. An insurance company could always design a (less than full insurance) contract that could earn positive profits. The intuition is that the first dollar of insurance has first order benefit from insurance but a second-order welfare loss from moral hazard (a result of the envelope theorem). In contrast, privately known heterogeneity in unemployment probabilities can shut down a market, even in the absence of any moral hazard problem. 4

Conceptually, our paper is closely related to that of Hendren (2011) which argues private information can explain insurance rejections, the practice of insurance companies choosing to not sell insurance to people with certain observable characteristics. That paper focuses on a comparative static implication of the no trade condition of more private information for those who would be rejected relative to those who can purchase insurance. In contrast, this paper tests the no-trade condition directly by leveraging estimates of the willingness to pay from existing optimal UI literature. Therefore, while Hendren (2011) shows that private information shuts down segments of three major insurance markets (LTC, Life, Disability), this paper shows private information shuts down the entire market for unemployment insurance.

The rest of this paper proceeds as follows. Section 2 presents the theoretical model and the no-trade condition. Section 3 relates the no-trade condition to the existing literature on optimal UI benefits and provides our key empirical tests. Section 4 discusses the data. Section 5 presents evidence that agents have private information about their unemployment risk. Section 6 quant-

---

3 To be clear, our results suggest the absence of a private UI market is a constrained efficient outcome. But, government intervention can be justified in many ways, such as utilitarian welfare maximization or ex-ante welfare grounds (e.g. implementing the optimal insurance contract before agents' realize their types).

4 In this sense, our paper contrasts with Chiu and Kari (1998), which argues the 'interaction of private information regarding employees' preferences for work with the unobservable level of effort exerted on the job may explain the absence of private unemployment insurance'. Our model captures their two-type model as a special case and shows that the role of moral hazard (i.e. unobservable effort exerted on the job) is actually not relevant for evaluating the existence of a private unemployment insurance market. Instead, their no-trade result was entirely driven by an assumption that the high risk type knew with certainty that she would become unemployed.
tifies this magnitude and compares it to the estimates of the willingness to pay provided by the existing literature on optimal UI benefits. Section 7 discusses government crowd-out and Section 8 concludes.

2.2 Theory

We consider a theoretical model of unemployment risk that captures both moral hazard (insurance increases unemployment) and privately known heterogeneity (people differ in their chance of becoming unemployed). The key result, given in Theorem 2.1, characterizes when agents can obtain any insurance beyond what is currently provided in their endowment through informal and formal systems.

2.2.1 Setup

There exists a unit mass of agents who are currently employed. With probability $p$ agents lose their job and become unemployed and with probability $1 - p$ they remain employed. Agents choose this probability by exerting (unobservable) effort, which incurs an additively separable utility cost of $\Psi (1 - p; \theta)$ which is increasing and convex in $1 - p$ for each $\theta$. Agents are heterogeneous in their cost of remaining employed, captured by the parameter $\theta$, distributed in the population according to a c.d.f. $F(\theta | X)$, where $X$ is a set of observable characteristics that could be used by insurance companies to price insurance contracts. Insurance companies cannot observe $p$ or $\theta$, but know $F(\theta | X)$ and all other aspects of the environment (utility function, effort function, etc.). For simplicity in this theoretical section, we condition on a particular observable characteristic, $X = x$, and let $F(\theta)$ denote the c.d.f. for some particular observable characteristic, $X = x$.\footnote{Note that this specification nests the possibility of no moral hazard by choosing $\Psi(p; \theta) = \gamma 1 \{p < \theta \}$ for sufficiently large $\gamma$. This induces the distribution $F(\theta)$ as the distribution of probabilities of remaining employed.} We let $\Theta$ denote the set of types, $\theta$.

Unemployed agents have an endowment of $c_e^u$ units of consumption; employed agents have an endowment of $c_e^e \geq c_u^u$. The endowment, $(c_e^e, c_e^u)$, incorporates existing formal and informal sources of insurance, such as savings, government insurance, firm severance, and informal insurance arrangements. Agents obtain utility $v(c_e)$ from consumption when employed and $u(c_u)$ when unemployed.\footnote{For now we allow for arbitrary complementarity between consumption and labor, but we will impose further assumptions in the empirical implementation.} We assume $v(c)$ and $u(c)$ are twice continuously differentiable, with $v', u' > 0$ and $v'', u'' < 0$. Although agents may have some sources of insurance, we assume $\frac{u'(c_e)}{v'(c_e)} > 1$, so that they are not fully insured against the occurrence of unemployment.

Agents choose $p$ to maximize expected utility:

$$ U(c_e, c_u; \theta) = \max_p \{(1 - p) v(c_e) + pu(c_u) - \Psi (1 - p; \theta)\} $$
We assume \( \Psi(p; \theta) \) is convex in \( p \) so that the effort choice is unique and given by the first order condition
\[
v(c_e) - u(c_u) = \Psi'(1 - p; \theta)
\]
where \( \Psi'(1 - p; \theta) \) denotes the first derivative of \( \Psi \) with respect to \( 1 - p \). Intuitively, the marginal cost of effort is equated to the benefit, given by the difference in utilities between employment and unemployment. The assumption \( \frac{\partial^2 \Psi}{\partial (1-p) \partial \theta} > 0 \) implies that agents of higher types \( \theta \) that face the same consumption bundle will have a higher probability of being unemployed.

We define an allocation in this economy to be a set \( A = \{c_u(\theta), c_e(\theta), p(\theta)\}_{\theta \in \Theta} \) of consumption bundles and probabilities of unemployment for each type \( \theta \).

### 2.2.2 Implementable Allocations

We seek conditions under which an insurer or market of insurers, which observes \( X \) but not \( p \), can provide agents the opportunity to consume a bundle other than this endowment. To ask this question, we consider the set of implementable allocations.

**Definition 2.1.** An allocation \( A = \{c_u(\theta), c_e(\theta), p(\theta)\}_{\theta \in \Theta} \) is **implementable** if

1. \( A \) is resource feasible:
\[
\int [p(\theta)c_e^\theta + (1-p(\theta))c_u^\theta - p(\theta)c_u(\theta) - (1-p(\theta))c_e(\theta)] dF(\theta) \geq 0
\]

2. \( A \) is incentive compatible:
\[
\begin{align*}
U(c_e(\theta), c_u(\theta); \theta) &\geq U(c_e(\tilde{\theta}), c_u(\tilde{\theta}); \tilde{\theta}) \quad \forall \theta, \tilde{\theta} \in \Theta \\
\Psi'(1 - p(\theta); \theta) &= v(c_e(\theta)) - u(c_u(\theta)) \quad \forall \theta \in \Theta
\end{align*}
\]

3. \( A \) is individually rational:
\[
U(c_e(\theta), c_u(\theta); \theta) \geq U(c_e^\theta, c_u^\theta; \theta) \quad \forall \theta \in \Theta
\]

Implementable allocations must not use more resources than are available in the economy and must satisfy the incentive and participation constraints imposed by private information. There are two incentive constraints. The first constraint requires agents choose their prescribed consumption bundle relative to other bundles. The second requires that the probability of unemployment, \( p(\theta) \), is consistent with agents’ effort incentives.\(^7\)

It is easy to verify that most models of market behavior (e.g. competition and monopoly) lead to allocations that must be implementable. We therefore ask under what conditions there exists
implementable allocations that differ from the endowment. By doing so, we provide a no-trade result that is not dependent on any choice of market structure.

2.2.3 Concavity Assumption

Before turning to the no trade result, we make an additional assumption to ensure that the moral hazard problem (i.e. choice of \( p \)) does not induce non-convexities which could cause our local variational analysis to be insufficient. To express this assumption, we need to introduce some additional notation. Let \( \Delta \) denote the difference in utilities between being employed and unemployed, so that lower values of \( \Delta \) correspond to greater amounts of insurance. Define \( \hat{p}(\Delta; \theta) \) to be the induced probability of unemployment for type \( \theta \), which solves

\[
\Psi'(1 - \hat{p}(\Delta; \theta); \theta) = \Delta
\]

It is straightforward to show that \( \hat{p} \) is decreasing in the size of the incentives to work, \( \Delta \). Now, define the cost functions,

\[
\begin{align*}
C_u(x) &= u^{-1}(x) \\
C_v(x) &= v^{-1}(x)
\end{align*}
\]

\( C_u(x) \) measures the amount of consumption required to provide \( x \) units of utility when unemployed; similarly, \( C_v(x) \) measures the amount of consumption required to provide \( x \) units of utility when employed.

Now, let \( \pi(\Delta, \mu; \theta) \) denote the profit obtained from type \( \theta \) if she is provided with total utility \( \mu \) and difference in utilities \( \Delta \),

\[
\pi(\Delta, \mu; \theta) = (1 - \hat{p}(\Delta; \theta)) (c^e_v - C_v(\mu - \Psi(1 - \hat{p}(\Delta; \theta)))) + \hat{p}(\Delta; \theta) (c^e_u - C_u(\mu - \Delta - \Psi(1 - \hat{p}(\Delta; \theta))))
\]

To guarantee the validity of our variational analysis for characterizing when the endowment is the only implementable allocation, it will be sufficient to require that \( \pi(\Delta, \mu; \theta) \) is concave in \((\Delta, \mu)\).

\( \pi(\Delta, \mu; \theta) \) is concave in \((\Delta, \mu)\) for each \( \theta \)

This assumption requires the marginal profitability of insurance to decline in the amount of insurance provided. If the agents choice of \( p \) is given exogenously (i.e. does not vary with \( \Delta \)), then concavity of the utility functions, \( u \) and \( v \), imply concavity of \( \pi(\Delta, \mu; \theta) \). However, allowing agents to choose \( p \) has the potential to create regions in which the marginal profitability of insurance actually increases in the amount of insurance. Yet, these non-concavities are ruled out by reasonable parameter restrictions. For example, in Appendix 2.A.2 we show profits are globally concave for any utility function and effort function satisfying \( \Psi'' \geq 0 \) and \( \frac{u'(c^e_v)}{v'(c^e_v)} \leq 2 \).

---

*The assumption \( \frac{u'(c^e_v)}{v'(c^e_v)} \leq 2 \) holds in our empirical application. The assumption \( \Psi'' \geq 0 \) holds if the probability*
2.2.4 No Trade Condition

Let \( P^e \) be the random variable denoting the probabilities of unemployment that occur when agents consume their endowment

\[
P^e(\theta) = \hat{p}(v(c^e_u) - u(c^e_e); \theta)
\]

where the random variable is generated from the realizations of \( \theta \) drawn with c.d.f. \( F(\theta) \). Let \( \Gamma^e \) denote the support of \( P^e \).

Now, consider an insurance company that tries to sell an insurance contract that provides a dollar to type \( \theta \) in the event she becomes unemployed. Let \( p = \hat{p}(v(c^e_u) - u(c^e_e)) \) denote the probability of unemployment for type \( \theta \) if she consumes her endowment. If she prefers this insurance contract relative to her endowment, then the incentive constraints imply that all of the higher risk types will also prefer this insurance contract relative to their endowment. Therefore, the average probability of those selecting this small insurance contract will be given by \( E[P^e|P^e \geq p] \). The no trade theorem says that unless someone in the economy is willing to pay the average cost of risks worse than them in order to obtain some insurance, there can be no trade.

**Theorem 2.1.** The endowment, \( \{(c^e_u, c^e_e)\}_{\theta \in \Theta} \), is the only implementable allocation if and only if

\[
\frac{p}{1 - p} \frac{u'(c^e_u)}{v'(c^e_e)} \leq \frac{E[P^e|P^e \geq p]}{1 - E[P^e|P^e \geq p]} \quad \forall p \in \Gamma^e \setminus \{1\}
\]

where \( \Gamma^e \setminus \{1\} \) denotes the support of \( P^e \) excluding the point \( p = 1 \).

Conversely, if (2.1) does not hold, then there exists an implementable allocation which strictly satisfies resource feasibility and individual rationality for a positive mass of types.

The LHS of condition (2.1) captures the willingness to pay for a small additional transfer of consumption from the employed to unemployed state given the existing state of insurance (e.g. informal and government insurance) leading to \( c^e_u \) and \( c^e_e \). The RHS is the cost of providing this transfer if the pool of worse risks are also attracted to the insurance contract. Unless someone is willing to pay this pooled cost, there can be no trade. Any contract or menu of contracts offered by an insurance company would be so heavily adversely selected that they would not deliver positive profits at any price. This provides a theoretical explanation for the absence of a private unemployment insurance market: Equation (2.1) holds for all observable characteristics, \( X \).

**Moral Hazard versus Private Information** Although private information about ones’ probability of becoming unemployed can shut down the market for insurance, moral hazard alone cannot. In the absence of unobserved heterogeneity, \( P^e \) is a degenerate distribution equal to a mass point at \( P^e = p \) for some \( p \). Therefore, the no trade condition reduces to \( \frac{u'(c^e_u)}{v'(c^e_u)} \leq 1 \). Therefore, as long

\( 1 - \hat{p}(\Delta) \), is concave in the size of the agents’ incentives, \( \Delta \).
as agents have some insurance value, \( u'(c_u^g) > v'(c_e^g) \), moral hazard alone cannot shut down the market.\(^9\)

**Moral Hazard and the Fiscal Externality** If the consumption endowment, \((c_e^g, c_u^g)\), is the result of other insurance arrangements, such as government UI, then a provision of additional insurance by a private insurer imposes an externality on the government by increasing the probability that the insured will become unemployed and file a claim. As long as the private market insurer is not required to compensate the government for the moral hazard impact of their claims, equation (2.1) continues to characterize when a private market can exist. If a third-party insurer must compensate the original provider of insurance, then the no trade condition becomes more restrictive; condition (2.1) remains sufficient, but not necessary, to ensure no trade.\(^10\)

### 2.2.5 Government Insurance and Crowd-out

If the government reduced or eliminated UI benefits, would a private market arise? We model government benefits as affecting the agents’ endowments and incentives to work. Let the endowment now be given by \((c_e^g(g), c_u^g(g))\) where \(g\) represents the generosity of government UI benefits. We assume \(c'_e < 0\) and \(c'_u > 0\) so that higher values of \(g\) lead to lower consumption when employed but higher consumption when unemployed. Now, let \(p_g(g; \theta)\) denote the unemployment probability under government benefit levels \(g\), which solves

\[
\Psi'(1 - p_g(g; \theta)) = v(c_e^g(g)) - u(c_u^g(g))
\]

and let \(P(g)\) denote the random variable induced from \(p(g; \theta)\) by taking realizations of \(\theta\) from the c.d.f. \(F(\theta)\). In other words, \(P(g)\) is the distribution of unemployment probabilities if the government provides benefits \(g\) and all agents consume their endowment, \((c_e^g(g), c_u^g(g))\).

**Corollary 2.1.** Suppose the government provides benefits, \(g\). Then, the endowment \(\{(c_e^g(g), c_u^g(g))\}_{g \in \Theta}\) is the only implementable allocation if and only if

\[
\frac{p}{1 - p} \frac{u'(c_u^g(g))}{v'(c_e^g(g))} \leq \frac{E[P(g) | P(g) \geq p]}{1 - E[P(g) | P(g) \geq p]} \quad \forall p \in \Gamma(g) \setminus \{1\}
\]

where \(\Gamma(g) \setminus \{1\}\) denotes the support of \(P(g)\) excluding the point \(p = 1\).

Corollary 2.1 characterizes when a private market cannot arise as a function of the government benefit level, \(g\). Smaller government benefits raise the value of external insurance by raising \(u'(c_u^g(g)) / v'(c_e^g(g))\).

---

\(^9\) The intuition of this result is similar to that of the standard result in distortionary taxation. The first dollar of insurance has a second-order welfare loss but a first-order welfare benefit resulting from the differences in marginal utilities between being employed and unemployed.

\(^10\) It is quite uncommon for the government to collect such externality payments from insurers in other contexts. Moreover, private insurers which have attempted to sell private UI have not been required to make externality compensation payments to the government.
However, they can also change the distribution of unemployment occurrences in the population, $P(g)$. Clearly, our assumption that $\frac{\partial^2 \phi}{\partial (1-p) \partial g} > 0$ implies that greater unemployment benefits lead to higher unemployment probabilities, on average. However, the impact of government benefits on the shape of the distribution of unemployment risks, and the value of $E[P(g) | P(g) \geq p]$, is an important area for future work.

For the bulk of the paper, we focus on explaining the absence of a private UI market conditional on the existing government benefits; but we return to the issue of government crowd-out in Section 2.7.

### 2.3 Empirical Strategy

This section empirically operationalizes the no trade condition (2.1). Multiplying by $\frac{1-p}{p}$ yields

$$\frac{u'(c^e_0)}{u'(c^e_2)} \leq T^e(p) \quad \forall p \in \Gamma^e \setminus \{1\}$$

where $T^e(p)$ is the pooled price ratio, given by

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

$T^e(p)$ is price ratio imposed on type $p$ if she must pay the average cost of all types $P^e \geq p$ in order to obtain a transfer from the employed to unemployed state. Note that it is a function of the distribution of unemployment probabilities evaluated at the endowment, $P^e$. Taking the infimum, the no-trade condition is given by

$$\frac{u'(c^e_0)}{u'(c^e_2)} \leq \inf_{p \in \Gamma^e \setminus \{1\}} T(p)$$

(2.2)

The value of $\frac{u'(c^e_0)}{u'(c^e_2)}$ is a measure of the willingness to pay for insurance evaluated at the endowment. In particular, $\frac{u'(c^e_0)}{u'(c^e_2)} - 1$ captures the markup individuals are willing to pay for a small transfer from the event of being employed to the event of being unemployed. The RHS of equation (2.2) is the measure of the barrier to trade imposed by private information, also evaluated using the existing probabilities of employment, $P^e$. In particular, $\inf_{p \in \Psi \setminus \{1\}} T(p) - 1$ is the smallest markup, or implicit tax, individuals must be willing to pay for insurance in order for the market to exist.

To estimate $\inf_{p \in \Psi \setminus \{1\}} T(p)$, we obtain an estimate of the distribution of probabilities of employment. For this we use subjective probability elicitations about whether someone will become unemployed in the subsequent year (i.e. a response to the question: “what’s the chance you’re going to lose your job in the next year?”). Following Hendren (2011), at no point do we assume agents report their true beliefs in survey elicitations; rather we allow these elicitations to be noisy and potentially biased measures of true beliefs. We first provide a simple test for the presence of
private information and then, using additional assumptions, quantify \( \inf_{p \in \Psi \backslash \{1\}} T^e(p) \). We provide more specific implementation details in Sections 4 and 5.

To obtain the willingness to pay for insurance, \( \frac{u'(c^e_\alpha)}{u'(c^e_\beta)} \), in equation (2.2), we use two existing sources from the literature on optimal government provided unemployment benefits, which we discuss here.

**WTP #1: Consumption Smoothing** By assuming state-independent utility \( (u = v) \), we can follow Baily (1976) by writing the willingness to pay for insurance as a function of the coefficient of relative risk aversion and the percentage consumption drop upon unemployment:

\[
\frac{u'(c^e_\alpha) - u'(c^e_\beta)}{u'(c^e_\beta)} \approx \frac{u''(c^*)}{u'(c^e_\beta)} (c^e_\alpha - c^e_\beta) \\
\approx \frac{c^e_\beta u''(c^*)}{u'(c^e_\beta)} (c^e_\alpha - c^e_\beta) \\
\approx \frac{\Delta c}{c^e_\beta}
\]

where \( \frac{\Delta c}{c^e_\beta} \) is the percentage consumption drop upon the event of becoming unemployed and \( \sigma \) is the coefficient of relative risk aversion.\(^\text{11}\) Gruber (1997) provides estimates of \( \frac{\Delta c}{c^e_\beta} \) and how it varies with the replacement rate of government UI benefits.

Given the willingness to pay approximation, the no trade condition can be expressed as

\[
\sigma \frac{\Delta c}{c^e_\beta} \leq \inf_{p \in \Gamma^e \backslash \{1\}} T^e(p) - 1 \tag{2.3}
\]

which, given estimates of \( \sigma \frac{\Delta c}{c^e_\beta} \), operationalizes a test of equation (2.2).

**WTP #2: Ratio of Search Elasticities** Chetty (2008) considers a search model conditional on being unemployed and derives a representation of \( \frac{u'(c^e_\alpha)}{u'(c^e_\beta)} \) using ratios of unemployment duration elasticities. The benefit of this approach is that we do not need to assume state independent utility or a coefficient of relative risk aversion. The downside of this approach is that the duration model with homogeneous agents used by Chetty (2008) does not neatly fit into the environment used to provide the no-trade condition above. Nonetheless we present here the intuition for how the ratio of the liquidity to moral hazard elasticities recovers the ratio of marginal utilities.

To begin, we assume unemployed agents choose the probability of becoming re-employed, \( s \), with a separable effort cost, \( \sigma (s) \). Agents choose \( s \) to maximize

\[
s u (c^n_\alpha) + (1 - s) u (c^n_\beta) - \sigma (s)
\]

\(^{11}\)Note that \( \frac{\Delta c}{c^e_\beta} \) is the percentage consumption drop given the observed system of government benefits which generate \( c^e_\alpha \) and \( c^e_\beta \); thus it is identifiable with data on the incidence of unemployment and consumption.
where $c^a_e$ and $c^a_u$ are the consumptions that the agent would receive in the event of becoming re-employed and remaining unemployed, respectively. Assuming $c^a_e \approx c^c_e$ and $c^a_u \approx c^c_u$, the first order condition for $s$ solves

$$\sigma'(s) = u'(c^c_e) - u'(c^c_u)$$

so that the marginal disutility of effort equals the difference in utility levels between being employed and unemployed.

We now analyze how the effort decision varies with the wage and with asset levels. First, consider a small increase in the wage, $w$, which increases consumption in the event of being employed, but has no effect on the event of being employed. Taking the derivative, we have

$$\frac{\partial s}{\partial w} = \frac{u'(c^c_e)}{\sigma''(s)}$$

so that higher wages lead to an increase in search effort in a manner that trades off the utility gain against the search effort costs. In contrast, a small increase in the agents assets upon entering unemployment also provide marginal utility during unemployment, so that

$$\frac{\partial s}{\partial A} = \frac{u'(c^c_e) - u'(c^c_u)}{\sigma''(s)}$$

As long as the marginal utility of consumption is larger when unemployed, this implies $\frac{\partial s}{\partial A} < 0$. Combining these two equations, we have

$$\frac{u'(c^c_e) - u'(c^c_u)}{u'(c^c_e)} = -\frac{\partial s}{\partial A} \frac{\partial s}{\partial w}$$

Given this representation, we can re-write the no-trade condition as

$$\frac{\partial s}{\partial A} \frac{\partial s}{\partial w} \leq \inf_{p \in \Gamma \setminus \{1\}} T(p) - 1$$

which, given estimates of $\frac{\partial s}{\partial A} \frac{\partial s}{\partial w}$ provided in Chetty (2008), operationalizes a test of condition (2.2).

One caveat to this representation is that we have not modeled the search effort decision of the unemployed in the model in Section 2.1.12 Future work could better integrate this search model with the setup and no-trade condition.

---

12 In particular, once people become unemployed in this model, they are all homogeneous and one would potentially expect an insurance market to exist to provide insurance against unemployment duration. Of course, heterogeneity in one's knowledge of unemployment duration could endogenously prevent the existence of the market. But, this yields a multi-dimensional screening problem (heterogeneity in probability of unemployment and duration); we leave the no-trade condition in this setting for future work.
Summary of approach Our empirical approach is to use subjective probability elicitations about future unemployment spells to estimate $\inf_{p \in \Psi \setminus \{1\}} T(p)$. We then compare this to the estimates of the willingness to pay for insurance: the consumption drop estimates from Gruber (1997) and the ratio of elasticities from Chetty (2008).

Because we use estimates of the willingness to pay for insurance and the distribution of unemployment probabilities under the existing regimes of government benefits and informal insurance arrangements, our approach asks whether the no trade condition can explain the absence of a market for additional UI, beyond what is currently provided through the government and other informal arrangements. To be precise, our approach simulates a hypothetical market for an additional dollar of UI. We estimate the implicit tax that would be imposed on this hypothetical market by adverse selection, and then we compare this to the existing estimates of the willingness to pay for additional UI.\textsuperscript{13} Then, in Section 2.7 we discuss extrapolating our results to a world with less or no government UI, as discussed in Section 2.2.5.

2.4 Data

Variables Estimating $\inf_{p \in \Psi \setminus \{1\}} T(p)$ using the methods in Hendren (2011) requires three pieces of data: a subjective probability elicitation about future unemployment, its corresponding indicator for whether or not unemployment occurs, and a set of public information insurance companies could use to price the insurance contracts.

Our data come from the Health and Retirement Study (HRS) spanning years 1993-2008. The HRS is an individual-level panel survey of individuals over 55 and their spouses (included regardless of age). This survey asks respondents: what is the percent chance (0-100) that you will lose your job in the next 12 months? We denote these free-responses by $Z$.

Figure 2.1 presents the histogram of the subjective probability elicitations. As has been noted in previous literature (Gan et al), responses tend to concentrate on focal point values, especially zero. As we will discuss further in Section 4 and 5, our empirical approach to identify whether people have any private information will not require knowing how these elicitations relate to true beliefs. However, to quantify $\inf_{p \in \Psi \setminus \{1\}} T(p)$ we will specify a parametric model for the relationship between elicitations and beliefs and estimate the parameters governing this measurement error.

Using the panel of the survey, we construct the indicator, $U$, denoting the occurrence of unemployment in the subsequent 12 months from the interview. Our definition of unemployment is those who report having lost their job because of their business being closed or because they were “laid off/let go”. Therefore, our definition of unemployment excludes voluntary quits; this implicitly

\textsuperscript{13}Note that we abstract from whether the benefit is paid initially or distributed throughout the course of the unemployment spell. If agents can equate marginal utilities throughout the unemployment spell, as is commonly assumed in the existing literature, then agents valuation of this additional dollar does not depend on when it is paid during the unemployment spell.
assumes that an insurance company could distinguish voluntary quits from involuntary leaves.\textsuperscript{14}

We consider three sets of increasing controls for public information insurance companies would use in pricing, outlined in Table 2.1. Our primary specification includes census region, job industry categories, wage bins, age, gender, and several health status indicators (e.g. diabetes, back condition, obesity, etc.). This set is parsimonious, although it is generally larger than the set of information than has previously been used by insurance companies who have tried to sell unemployment insurance.\textsuperscript{15} In addition to this set, we also assess the robustness of our analysis to a smaller set of controls (age and gender only) and a larger set of controls (which includes additional job characteristics and health). Moreover, we consider an additional specification that excludes anyone who experienced a recent unemployment spell in the past four years. This allows us to test whether insurance companies could avoid adverse selection by selling only to those with stronger employment histories.

**Sample selection** We consider the sample of individuals aged 41-59 who are currently employed but are not self employed. We restrict our sample to those providing a subjective probability elicitation about future unemployment and for whom we can follow for a subsequent year in the panel.

Table 2.2 presents the summary statistics of our sample. Our mean age is 55, in the upper end of the sample range. This is because individuals below age 55 are only found in our sample if they are a spouse of someone over 55. The average wage in our sample is $19/hr. Roughly 4% of our sample become unemployed in the subsequent 12 months. The mean of the subjective elicitations is

\textsuperscript{14}Government UI attempts to exclude voluntary quits and arguably gains enforcement power by requiring employer contributions that are experience-rated. Thus, our approach can be seen as modeling an insurance policy that leverages the existing government UI system but pays an additional $1 to the unemployed on the first day they receive their government UI payment.

\textsuperscript{15}IncomeAssure, the latest attempt to provide private unemployment benefits, prices policies using a coarse industry classification, geographical location (state of residence), and wages.
Table 2.1: Covariate Specification

<table>
<thead>
<tr>
<th>Preferred Controls</th>
<th>Extended Controls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, Age*2, Gender</td>
<td>Full interactions of</td>
</tr>
<tr>
<td>Gender*age</td>
<td>Age</td>
</tr>
<tr>
<td>Gender<em>age</em>2</td>
<td>Gender</td>
</tr>
<tr>
<td>Self employed Indicator</td>
<td>Full interactions of</td>
</tr>
<tr>
<td>Health Variables</td>
<td>wage decile</td>
</tr>
<tr>
<td>Obesity indicator (40+ BMI)</td>
<td>part time indicator</td>
</tr>
<tr>
<td>Psych condition indicator</td>
<td>job tenure quartile</td>
</tr>
<tr>
<td>Back condition indicator</td>
<td>self-employment indicator</td>
</tr>
<tr>
<td>Diabetes indicator</td>
<td>Presence of back condition, psych condition, or obesity</td>
</tr>
<tr>
<td>BMI (linear)</td>
<td></td>
</tr>
<tr>
<td>Job industry dummies</td>
<td>Full interactions of</td>
</tr>
<tr>
<td>Census region indicators</td>
<td>Job requires stooping</td>
</tr>
<tr>
<td></td>
<td>Job requires lifting</td>
</tr>
<tr>
<td></td>
<td>Job requires phys activity</td>
</tr>
<tr>
<td></td>
<td>Presence of back condition, psych condition, or obesity</td>
</tr>
<tr>
<td></td>
<td>Full interactions of</td>
</tr>
<tr>
<td></td>
<td>Census region</td>
</tr>
<tr>
<td></td>
<td>5 year age bins</td>
</tr>
</tbody>
</table>

Table 2.2: Sample Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lose Job (1yr)</td>
<td>0.04</td>
<td>0.20</td>
</tr>
<tr>
<td>Subj Prob Lose Job (1yr)</td>
<td>0.16</td>
<td>0.25</td>
</tr>
<tr>
<td>Age</td>
<td>54.46</td>
<td>3.72</td>
</tr>
<tr>
<td>Female</td>
<td>0.63</td>
<td>0.48</td>
</tr>
<tr>
<td>Wage ($/hr)</td>
<td>18.98</td>
<td>41.72</td>
</tr>
<tr>
<td>Sample Size</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs</td>
<td>12,880</td>
<td></td>
</tr>
<tr>
<td>Households</td>
<td>5,101</td>
<td></td>
</tr>
</tbody>
</table>
16%, indicating significant (mean) bias in elicitations. Such upward bias is perhaps not surprising given the low probability of unemployment and the fact that elicitations are bounded between 0 and 1.

Discussion  Given these variable definitions, how should one think of our empirical approach? We study the potential adverse selection that would occur in a market for a hypothetical contract that provides $1 in the event of losing one's job in the subsequent 12 months from the initial date of contracting and is priced based on the observable characteristics, $X$.\footnote{By changing the set of observable characteristics, $X$, we simulate different underwriting strategies and can assess how variation in the use of observable characteristics potentially mitigates impact of adverse selection.} We ask what markup individuals would need to be willing to pay on this hypothetical contract in order for it to earn nonnegative profits, given by \( \inf_{p \in \Psi \setminus \{1\}} T(p) - 1 \), and we compare this estimate to the willingness to pay estimates implied by consumption smoothing (equation (2.3)) and ratios of search elasticities (equation (2.4)).\footnote{Note that the minimum pooled price ratio and the willingness to pay can both vary with observables, $X$, so that this test can be thought of as being repeated for different values of $X$. We discuss this further in Section 6.}

## 2.5 Presence of private information

Before testing equations (2.3) and (2.4) which require a quantification of \( \inf_{p \in \Psi \setminus \{1\}} T(p) \), we begin with the more straightforward question of whether or not people have private information about their unemployment risk, beyond information contained in the observables, $X$.

### 2.5.1 Identification Assumptions

As in the theoretical section, Let $P$ be the random variable representing agents' beliefs about the occurrence of unemployment in the next 12 months, $U$. Throughout, we assume that agents elicitations, $Z$, may not equal agents true beliefs which govern behavior, $P$. Instead, we assume $P$ is unobserved to the econometrician. To derive a test for the presence of private information, we make two assumptions. First, we assume that agents beliefs are unbiased.

Beliefs are unbiased: $\Pr\{U|X,P\} = P$

This assumption states that if we hypothetically observed someone with true beliefs $P$ then their probability in the data of experiencing unemployment would equal $P$. This assumption is weaker than traditional rational expectations assumptions (we do not require agents to know the probability structure of the environment), but nonetheless is not a trivial assumption. For our purposes, it provides a simple link between the unobserved beliefs and the observed occurrence of unemployment.

In addition to assuming that the unobserved beliefs are unbiased, we assume that the observed elicitations, $Z$, contain no more information about $U$ than would the true beliefs.
No more information: \( \Pr \{ U|X, Z, P \} = \Pr \{ U|X, P \} \)

If we observed both the elicitation and the true beliefs, then the elicitation would not provide any additional forecasting information about \( U \) than does \( P \). This is a relatively weak assumption; it would be difficult for agents to report predictive information that they did not know themselves. Indeed, it allows \( Z \) to be any noisy measure of true beliefs that is independent of \( U \) conditional on \( P \) and \( X \).

Under these two assumptions, the true beliefs are a mean-preserving spread of the distribution of predicted values:

\[
\Pr \{ U|X, Z \} = E \{ P|X, Z \}
\]

In this sense, the distribution of the predicted loss given \( X \) and \( Z \), \( \Pr \{ U|X, Z \} \), are a lower bound for the true distribution of beliefs, \( P \). This motivates a simple test for the presence of private information about \( U \): is \( Z \) is predictive of \( U \) conditional on \( X \)?

### 2.5.2 Specification and Results

We adopt a linear specification\(^{19}\),

\[
U = \beta X + \gamma Z + \epsilon
\]

A finding of \( \gamma \neq 0 \) allows one to reject the null hypothesis of no private information.\(^{20}\) We consider specifications for each of the three sets of controls, \( X \), outlined in Table 2.1.

**Results** Table 2.3 presents the results. Across all three public information specifications (Columns I, II, and III), we reject the null hypothesis of no private information at p-values less than 0.001. Indeed, the results suggest that agents have information beyond what is captured by age, gender, various health characteristics, job industry, and wages. Thus, even if private insurance policies were priced using all of these variables, insurance policies would likely be adversely selected. Without additional assumptions, the coefficient \( \gamma \) does not have an interpretation related to the beliefs, \( P \). However, the magnitude does show that individuals who report a 1 standard deviation higher subjective probability (0.25) have an average probability of unemployment that is 1.5pp higher.

**Historical Screening** Although we find evidence of private information conditional on these sets of controls, insurance companies could potentially also use past unemployment spells to discriminate applicants. For example, insurance companies could require that applicants have no unemployment record in the past several years. To assess the impact of this underwriting procedure, Column IV restricts the sample to those who have not experienced unemployment in the past 2 waves (4 years)

\(^{18}\)More formally, if there exists a \( z \) such that \( \Pr \{ U|X, Z = z \} \neq \Pr \{ U|X \} \), then there exists \( p \) such that \( F(p|X) \neq 1 \{ p \leq \Pr \{ U|X \} \} \).

\(^{19}\)Our results remain robust to other specifications (e.g. probit, logit).

\(^{20}\)Note that the magnitude of the coefficient \( \gamma \) does not have any direct interpretation without additional structure.
Table 2.3: Presence of Private Information

<table>
<thead>
<tr>
<th></th>
<th>I</th>
<th>II</th>
<th>III</th>
<th>IV</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age &amp;</td>
<td>Primary</td>
<td>Extended</td>
<td>Clean</td>
</tr>
<tr>
<td></td>
<td>Gender</td>
<td>Controls</td>
<td>Controls</td>
<td>History Sample</td>
</tr>
<tr>
<td>Subjective Probability</td>
<td>0.061***</td>
<td>0.048***</td>
<td>0.032***</td>
<td>0.035***</td>
</tr>
<tr>
<td>p-value</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.0001</td>
<td>0.0004</td>
</tr>
<tr>
<td>N</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>7,933</td>
</tr>
</tbody>
</table>

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

of the survey. The results suggest the elicitation is still predictive of future unemployment even on the set of people who have not experienced unemployment in the recent past. Thus only selling insurance to those who have not experienced unemployment in the past 4 years would not remove the informational asymmetry.

In short, our results suggest private information would impose constraints on the workings of a private unemployment insurance market. We now turn to quantifying the size of this barrier to trade, \( \inf_{p \in \Psi \setminus \{1\}} T(p) \).

### 2.6 Quantification of Private Information

To estimate \( \inf_{p \in \Psi \setminus \{1\}} T(p) \), we impose additional structure to the relationship between elicitations and beliefs. While the test for the presence of private information does not impose any structure on the relationship between elicitations and true beliefs, we now assume the distribution of elicitations given beliefs can be parameterized by a vector of parameters, \( \theta \).

The distribution of elicitations given beliefs is given by \( f_{Z|P}(Z|P; \theta) \) where \( \theta \) is a finite vector of parameters.

We then estimate \( f_P(p) \) using the identity

\[
f_{Z,U}(Z,U|X) = \int p^U (1-p)^{1-U} f_{Z|P}(Z|X,P,\theta) f_P(p|X) \, dp
\]

where \( f_{Z,U} \) is the observed density of elicitations and unemployment, \( f_{Z|P} \) is the density of elicitations given beliefs (parameterized by \( \theta \)), and \( f_P \) is the density of beliefs. We then flexibly approximate \( f_P \) and estimate both \( f_P \) and \( \theta \) using maximum likelihood.\(^{22}\)

---

\(^{21}\)We use the primary set of controls, although the results remain robust to the extensive set of controls.

\(^{22}\)Hendren (2011) discusses identification using restrictions on \( f_{Z|P} \) in more detail. We show that if \( f(Z|P) \) is normally distributed with mean \( P+\alpha(X) \) then \( f_P \) is non-parametrically identified and there exists an infinite number of over-identifying moments. In practice, we adopt a censored normal distribution for \( f(Z|P) \) which potentially poses identification concerns but in practice performs quite well as indicated by monte carlo results provided in Hendren (2011).
2.6.1 Measurement error distribution

We parameterize the distribution of elicitations given beliefs as a mixture of a censored normal and an ordered probit distribution. The ordered probit captures excess density of elicitations at 0, 50 and 100, as shown in Figure 2-1. More precisely, we assume that the p.d.f./p.m.f. of $Z$ given $P$ is given by

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

where $\phi$ denotes the standard normal p.d.f. and $\Phi$ the standard normal c.d.f. We estimate four elicitation error parameters: $(\sigma, \lambda, \kappa, \alpha(X))$. $\sigma$ captures the dispersion in the elicitation error, $\lambda$ is the fraction of focal point respondents, $\kappa$ is the focal point window. We allow the elicitation bias term, $\alpha(X)$, to vary with the observable variables, $X$.23

2.6.2 Distribution of $f_P$

Ideally, we would flexibly estimate the c.d.f. of $P$ given $X$, $F(p|X)$, and the minimum pooled price ratio, $\inf_{p \in \psi \{1\}} T(p)$, separately for every possible value of $X$. However, the dimensionality of $X$ prevents this in practice. Instead, we adopt an index assumption:

$$F(p|X) = \tilde{F}(p|\Pr \{U|X\})$$

where we assume $\tilde{F}(p|q)$ is continuous in $q$. This assumes that the distribution of private information is the same for two observable values, $X$ and $X'$, that have the same observable unemployment probability, $\Pr \{U|X\} = \Pr \{U|X'\}$. Although one could perform different dimension reduction techniques, controlling for $\Pr \{U|X\}$ is particularly appealing because it nests the null hypothesis of no private information ($F(p|X) = 1 \{p \leq \Pr \{U|X\}\}$). Moreover, it allows us to easily impose unbiased beliefs, so that $\Pr \{U|X\} = E[P|X]$ for all $X$.

We then approximate $\tilde{F}(p|q)$ for $q = \Pr \{U|X\}$ using a mixture of a beta distribution and a point-mass distribution.24

$$\tilde{F}(p|q) = w1 \{ p \leq q - a \} + (1 - w) \sum_i \xi_i \Beta(\mu_i(q), \psi_i)$$

where $w$ is the weight on the point-mass, $q - a$ is the mean of the point-mass, $\Beta(\mu_i(q), \psi_i)$ is the Beta c.d.f. with mean $\mu_i(q)$ and shape parameter $\psi_i$, and $\{\xi_i\}_i$ are the weights on the Beta

---

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

24 The point mass captures the possibility that a fraction of the population has the same information set.
distributions. We assume $\mu_i(q)$ is linear in $q$, $\mu_i(q) = g_i + h_i q$. We use two Beta distributions. Unbiased beliefs imposes the restriction $E[P|q] = q$, which imposes linear restrictions on the mean and weights of each mixture component. We also impose restrictions on the parameter space to ensure the support lies over the interval $[0, 1]$.

2.6.3 Estimation of $\inf T(p)$

Given an estimate of $F(p|q)$, we construct $E[P|P \geq p]$ for each $p$ (and for each value of $q$). As discussed in detail in Hendren (2011), the construction of $E[P|P \geq p]$ suffers an extremal quantile estimation problem for values of $p$ in the upper quantiles of $F(p|X)$. Intuitively, as $p$ increases, the estimate of $E[P|P \geq p]$ relies on a shrinking sample. Thus, we restrict attention to $p \leq F^{-1}(\tau)$ for $\tau = 0.9$ (i.e., values of $p$ less than the 90th quantile of its estimated distribution) and construct $\inf_{p \in [0, F^{-1}(\tau)]} T(p)$. We then assess robustness to the choice of $\tau$.

While the primary motivation for such a restriction is statistical, there is a straightforward economic rational for restricting attention to $p \leq F^{-1}(\tau)$. If insurance companies must attract a non-trivial fraction $1 - \tau > 0$ to any given consumption bundle other than the endowment, then $\inf_{p \in [0, F^{-1}(\tau)]} T(p)$ characterizes the barrier to trade (See Hendren (2011), Remark 1). Indeed, our choice of $\tau = 0.9$ corresponds to a requirement that insurance companies must attract more than just the 10% riskiest fraction of the market in order to sell insurance. That said, we re-iterate that the rationale for this assumption is statistical necessity.

2.6.4 Estimation Results

Figure 2-2 plots the estimated c.d.f., $\tilde{F}(p|q)$ for $q = Pr\{U\} \approx 4\%$ along with its bootstrapped standard errors. The results suggest a large fraction of the population (~70%) have very low probability of unemployment, near zero. The remaining fraction of the population is dispersed throughout an upper tail of higher risks. Translating this into the implied barrier to trade, Table 2.4 presents the estimated values of $\inf_{p \in [0, F^{-1}(0.9)]} T(p)$ for varying values of $q = Pr\{U|X\}$. We estimate a value of 3.506 (95% CI of [2.612, 3.871]) at the mean of the index, $q = 0.04$, and values of 2.652 to 9.571 at the 80th and 20th quantile of the distribution of the index, $q$. Subtracting 1 from 2.652, we conclude that unless people are willing to pay more than a ~165% markup for unemployment insurance, the results are consistent with the absence of a private unemployment insurance market.

---

25The p.d.f. of a beta distribution is given by $f(p; \alpha, \beta) = \frac{p^{\alpha - 1}(1-p)^{\beta - 1}}{B(\alpha, \beta)}$ where $B(\alpha, \beta)$ is the beta function. The mean of the beta distribution is given by $\mu = \frac{\alpha}{\alpha + \beta}$ and the shape parameter is $\psi = \alpha + \beta$. In principle, the point-mass distribution could be replaced with a Beta distribution with a very low variance. In practice, the point-mass performs better since numerical integration of low variance beta distributions is computationally time-consuming.

26For example, we allow the parameter, $\alpha$, to vary with $q$ to ensure that the point mass does not fall below $p = 0$ or above $p = 1$. 

107
Figure 2-2: CDF of Private Information

Table 2.4: Minimum Pooled Price Ratio, inf T(p)

<table>
<thead>
<tr>
<th>Quantile of Index, q</th>
<th>LTC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>inf T</td>
<td>3.506</td>
</tr>
<tr>
<td>5%</td>
<td>2.612</td>
</tr>
<tr>
<td>95%</td>
<td>3.871</td>
</tr>
<tr>
<td>Pr(L</td>
<td>reject)</td>
</tr>
</tbody>
</table>

Note: 5/95% CI computed using bootstrap block re-sampling at the household level (N=150 Reps). Bootstrap CI presents maximum boundaries from a) non-accelerated bias-corrected procedure from Efron (1982) and b) studentized values.

Table 2.5: Robustness to Choice of τ

<table>
<thead>
<tr>
<th>Quantile Region: ψ,</th>
<th>LTC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0-80%</td>
</tr>
<tr>
<td>Reject</td>
<td>3.506</td>
</tr>
<tr>
<td>5%</td>
<td>2.612</td>
</tr>
<tr>
<td>95%</td>
<td>3.871</td>
</tr>
</tbody>
</table>

Note: 5/95% CI computed using bootstrap block re-sampling at the household level (N=150 Reps). Bootstrap CI presents maximum boundaries from a) non-accelerated bias-corrected procedure from Efron (1982) and b) studentized values.
Table 2.6: Willingness to Pay

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Implied WTP</td>
<td></td>
</tr>
<tr>
<td>Δc/c</td>
<td></td>
<td>u'(c)/V(c)^-1</td>
</tr>
<tr>
<td>σ=2</td>
<td>.09-.12</td>
<td>.60</td>
</tr>
<tr>
<td>σ=3</td>
<td>.18-.24</td>
<td></td>
</tr>
<tr>
<td>σ=4</td>
<td>.27-.36</td>
<td></td>
</tr>
<tr>
<td>u'(c)/V(c)^-1</td>
<td>.36-.48</td>
<td></td>
</tr>
</tbody>
</table>

Range shown for 40-50% replacement rates

Robustness to choice of \( \tau \)  Table 2.5 assesses the robustness of the results to lower and higher values of \( \tau \), the upper quantile domain for our estimation of \( \inf_{p \in [0, F^{-1}(\tau)]} T(p) \). We focus on the mean value of the index \( q = 0.04 \). The minimum pooled price ratio is not attained at the upper boundary of \([0, F^{-1}(\tau)]\) for \( \tau \) between the 80th and 90th percentile; thus changes in \( \tau \) in this range do not affect the estimated value of \( \inf_{p \in [0, F^{-1}(\tau)]} T(p) \). However, at the 95th percentile, we do estimate the minimum to be at the boundary. Increasing \( \tau \) to 0.95 leads the minimum pooled price ratio to drop dramatically to 1.136 (95% CI of [1.091, 1.136]). However, this result is arguably driven by functional form. Since the minimum pooled price ratio is not identified as \( \tau \to 1 \), our estimates rely on functional form as we increase towards 1. We have only one beta distribution above the 90th percentile which, we estimate to be relatively concentrated (and thus having a lower value of \( T(p) \)). But, follow-up work could assess the robustness to the inclusion of additional betas in the upper portion of the distribution. In the meantime, this exercise simply highlights the instability of the estimates of \( \inf_{p \in [0, F^{-1}(\tau)]} T(p) \) as \( \tau \to 1 \).

2.6.5 Testing the No Trade Condition

How much of a markup are people willing to pay for unemployment insurance? Table 2.6 presents estimates of the LHS of equation (2.3) from the existing literature on optimal government unemployment insurance. Gruber (1997) yields estimates of \( \Delta c/c \) ranging from 9-12%, depending on the replacement rate of benefits (which generally range between 40 and 50% depending on the state). With a coefficient of relative risk aversion equal to 4, this implies individuals would be willing to pay between a 36% and 48% markup for unemployment insurance, well below the estimated markups individuals would need to be willing to pay to overcome the barriers to trade imposed by private information of at least 160% (the value for \( q = 0.015 \) in Table 2.4).

Chetty (2008) estimates that roughly 60% of the elasticity of unemployment duration with respect to benefits is due to a liquidity effect. This suggests people are willing to pay roughly a 60% markup for insurance, which again falls below the willingness to pay required to overcome the barriers imposed by private information. In short, our results suggest the amount of private information about unemployment risk is large enough to explain an absence of a private unemployment
Table 2.7: Elicitation Error Parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>s.e.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Deviation</td>
<td>0.115</td>
<td>(0.002)</td>
</tr>
<tr>
<td>Fraction Focal Respondents</td>
<td>0.552</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Focal Window</td>
<td>0.229</td>
<td>(0.003)</td>
</tr>
</tbody>
</table>

Note: Bootstrapped standard errors computed using block re-sampling at the household level (results shown for N=150 repetitions)

2.6.6 Measurement error variables

Table 2.7 presents the estimated parameters θ for the distribution of elicitations given beliefs, f(Z|P, X, θ). We find a standard error of 0.115, which suggests people are not able to report their true beliefs perfectly in their elicitations. We also estimate a sizable fraction of focal point respondents of 55.2%, which is consistent with the extremely large response rate of Z = 0. We estimate a focal window of 0.229, which suggests focal point respondents who otherwise would have reported a value of Z less than 0.229 instead collapse their report to Z = 0, generating the excess mass at Z = 0.

2.7 Government Crowd-Out

Private information explains why there is no private unemployment insurance market given the existing set of government benefits. But, if the government were to reduce the generosity of UI benefits, would a private market arise?

Section (2.2.5) shows that an ideal test would be to observe measures of \( \nu'(c^a(g)) \) and an estimate of the distribution of unemployment probabilities, \( P(g) \), for values of \( g \) near zero. Absent such a world, we make two assumptions that allow extrapolation from our estimates to a world with no government-provided unemployment benefits. First, we assume government UI benefits does not significantly affect the distribution of beliefs about the incidence of unemployment. This would be true, for example, if there were no significant moral hazard effects of government UI on the

\footnote{One caveat to this analysis is that our demand estimates are aggregated across values of X. Ideally, we would obtain or construct separate demand estimates for varying values of X and test equations (2.3) and (2.4) for each value of X. Indeed, the samples used by Gruber and Chetty are generally younger than our HRS sample. We leave this adjustment of demand for covariates and a more general test across values of observable characteristics as important future work.}

insurance market.\(^{27}\)
occurrence of unemployment.\textsuperscript{28} With this assumption, \( \inf_{p \in \mathcal{P}\setminus \{1\}} T(p) \) continues to characterize the implicit tax imposed by private information in a world without government UI benefits.

However, we do allow for the provision of government UI to affect the ratio of marginal utilities, \( \frac{\nu'(c_{z}(g))}{\nu'(c_{z}(\theta))} \), consistent with the findings of Gruber (1997). To do so, we adopt the log-linear specification from Gruber (1997), so that

\[
\log(c_{ist}) = a_i + b \cdot U_{it} + c \cdot U_{it} \cdot R_{st} + \gamma_{st} + \epsilon_{ist} \tag{2.6}
\]

where \( c_{ist} \) is consumption of person \( i \) in state \( s \) at time \( t \), \( U_{it} \) is an indicator for unemployment, \( R_{st} \) is the UI replacement rate for state \( s \) at time \( t \), and \( \gamma_{st} \) are state-by-year fixed effects. With this specification, \( b \) captures the consumption drop upon unemployment at the replacement rate of \( R_{st} = 0 \), which corresponds to no government UI. Replacement rates in the data range from 37\% to 54\%, and thus one should keep in mind that this extrapolation is of course out-of-sample.

Estimation results of equation (2.6) from Gruber (1997) suggest that the markup people would be willing to pay for UI would roughly double in the absence of government-provided UI. Gruber (1997) estimates \( b = -0.231 \) and \( c = 0.280 \), which indicates \( \frac{\Delta c}{c} \approx 23\% \) when \( R_{st} = 0 \), roughly twice as large as the estimates of 9–12\% for the consumption drop at existing replacement rates of 40–50\%. Thus, if individuals have a coefficient of relative risk aversion of \( \sigma = 4 \), then they should be willing to pay roughly a 90\% markup for insurance. If individuals have a coefficient of relative risk aversion of \( 2 \), then they would be willing to pay a 46\% tax for insurance. Given our much larger estimates of the barrier to trade imposed by private information, our results suggest that a private market is unlikely to arise in the absence of government provided benefits.

\subsection*{2.8 Conclusion}

This paper studies whether private information explains the absence of a private market for unemployment insurance above and beyond what is currently provided by the government and informal systems. We first document the presence of private information about unemployment risk by showing subjective probability elicitations are predictive of future unemployment, conditional on a rich set of public information that could be used to price such insurance. We then quantify the barrier to trade imposed by private information as equivalent to imposing tax rates in excess of 160\% on insurance premiums on hypothetical insurance contracts. We then show that these estimates are generally greater than estimates of the willingness to pay, found in existing literature on optimal government UI benefits, which generally range from 30–60\%. Thus, the barrier to trade imposed by private information is large enough to explain the absence of a private UI market in the US for insurance beyond what is currently provided by the government. Finally, we argue that, although the absence of government-provided UI would increase the demand for private UI benefits, this

\textsuperscript{28}The impact of moral hazard on \( \inf_{p \in \mathcal{P}\setminus \{1\}} T(p) \) is ambiguous; so assuming no change is perhaps not a bad approximation.
increase would likely not be large enough to overcome the barriers imposed by private information.
2.A Appendix

2.A.1 Proof of Theorem 1

We follow Hendren (2011) and consider the maximization program of a monopolist insurer. Whether there exists any implementable allocations other than the endowment corresponds to whether there exists any allocations other than the endowment which maximize the profit, \( \pi \), subject to the incentive and participation constraints.

The insurer can offer a menu of contracts, \( \{\mu(\theta), \Delta(\theta)\}_{\theta \in \Gamma} \) where \( \mu(\theta) \) specifies a total utility provided to type \( \theta \) and \( \Delta(\theta) \) denotes the difference in utilities if the agent becomes unemployed. Note that \( \mu(\theta) \) implicitly contains the disutility of effort.

For exposition of the proof, we switch focus from the probability of unemployment, \( \hat{p} \), to \( \hat{q} \), which we define to be the probability of employment,

\[
\hat{q}(\Delta; \theta) = 1 - \hat{p}(\Delta; \theta)
\]

so that the agent’s effort cost is \( |q(\Delta; \theta)| \). This way we need not keep track of as many sign changes in derivatives. Note that a type \( \theta \) that accepts a contract containing \( \Delta \) will choose a probability of employment \( \hat{q}(\Delta; \theta) \) consistent with the first order condition \( \Psi'(\hat{q}(\Delta; \theta); \theta) = \Delta \).

Let \( \pi(\Delta, \mu; \theta) \) denote the profits obtained from providing type \( \theta \) with contract terms \( \mu \) and \( \Delta \), given by

\[
\pi(\Delta, \mu; \theta) = \hat{q}(\Delta; \theta) (\mu - \Psi(\Delta; \theta)) + (1 - \hat{q}(\Delta; \theta)) (c_e - C_u(\mu - \Delta - \Psi(\Delta; \theta)))
\]

Note that the profit function takes into account how the agents’ choice of \( p \) varies with \( \Delta \). Assumption 2.2.3 maintains that \( \pi \) is concave in \( (\Delta, \mu) \).

**Preservation of Single Crossing** In the general problem, we allow the monopolist to offer a full menu of contracts. However, Hendren (2011) shows if the incentive constraints satisfy single crossing whereby the higher risk types have higher marginal values of insurance, then we can focus solely on single contract deviations from the endowment. So, consider the utility provided to a type \( \theta \) from contract terms \( (\mu, \Delta) \) when her intended contract was given by \( (\mu(\theta), \Delta(\theta)) \),

\[
\nu(\mu, \Delta|\mu(\theta), \Delta(\theta), \theta) = \mu - [(1 - \hat{q}(\Delta; \theta)) \Delta - (1 - \hat{q}(\Delta(\theta); \theta)) \Delta(\theta)] - [\Psi(\hat{q}(\Delta; \theta); \theta) - \Psi(\hat{q}(\Delta(\theta); \theta); \theta)]
\]

It is easy to verify that

\[
\frac{\partial \nu}{\partial \Delta} = - (1 - \hat{q}(\Delta; \theta))
\]

so that

\[
\frac{\partial^2 \nu}{\partial \Delta \partial \theta} = \frac{\partial q(\Delta; \theta)}{\partial \theta} < 0
\]
which is less than zero by the assumption that higher types $\theta$ have higher marginal costs of effort, \( \frac{\partial^2 \psi}{\partial \Delta \partial \theta} > 0 \). Therefore, the incentive constraints satisfy the single crossing property.

**Reformulation of the Profit Maximization Program** The profit maximization problem for a firm choosing a full menu of contracts, \( \{ \mu(\theta), \Delta(\theta) \}_{\theta \in \Theta} \), is given by

\[
\max_{\mu(\theta), \Delta(\theta)} \int \pi(\Delta(\theta), \mu(\theta); \theta) dF(\theta) \\
\text{s.t.} \quad \mu(\theta) \geq \nu(\mu(\hat{\theta}), \Delta(\hat{\theta}) | \mu(\theta), \Delta(\theta), \theta) \geq U(\theta)
\]

However, as shown in Hendren (2011) (Appendix A, Section 4), when the incentive constraints satisfy single crossing, and $\pi$ is concave, then there exists an allocation other than the endowment that solves this maximization program if and only if there exists a single local contract deviation from the endowment which attracts all types \( \{ \theta \geq \hat{\theta} \} \) to the same consumption bundle. By concavity of the profit function, it suffices to check only local deviations. Therefore, fix any type $\hat{\theta} \in \Theta$. Consider providing this type with a small transfer in the event unemployment occurs. By the envelope theorem, the agent’s marginal rate of substitution captures the agent’s marginal willingness to pay for this transfer

\[
1 - \frac{\hat{q}(\Delta^e; \hat{\theta}) u'(c^e_{\hat{\theta}})}{\hat{q}(\Delta; \hat{\theta}) u'(c^e_{\hat{\theta}})}
\]

Now, if all types $\theta \geq \hat{\theta}$ are also attracted to this contract, the marginal cost of providing the transfer to type $\hat{\theta}$ is given by

\[
1 - \frac{E \left[ \hat{q}(\Delta^e; \theta) | \theta \geq \hat{\theta} \right]}{E \left[ \hat{q}(\Delta; \theta) | \theta \geq \hat{\theta} \right]}
\]

If this marginal cost is greater than the willingness to pay, evaluated at each $\theta$, then there does not exist any local deviations. Thus, the endowment is the only implementable allocation. Replacing $\hat{q} = 1 - \hat{p}$ yields the expression in equation (2.1). Conversely, if the marginal cost is less than this willingness to pay at some $\theta$, then the insurance company could profitably provide this agent with a small transfer; thus the endowment would not be the only implementable allocation. QED.

**Discussion of Moral Hazard** Whether or not trade occurs does not depend on how insurance affects the probability of being employed. Rather, it only depends on the levels of the probabilities evaluated at the endowment, $\hat{q}(\Delta^e; \theta)$. Why is this? To see what goes on perhaps more clearly, consider a simpler setup of the monopolist trying to provide consumption bundles $c_e$ and $c_u$ in the event of being employed and unemployed, respectively, and assume no heterogeneity in effort costs.
Profits are given by:

$$\pi^C (c_e, c_u) = \hat{q} (c_e, c_u) (c_e^e - c_e) + (1 - \hat{q} (c_e, c_u)) (c_u^e - c_u)$$

so that

$$\frac{\partial \pi}{\partial c_e} = -\hat{q} (c_e) + \frac{\partial \hat{q}}{\partial c_e} (c_e^e - c_e)$$

and

$$\frac{\partial \pi}{\partial c_u} = -(1 - \hat{q} (c_u)) - \frac{\partial \hat{q}}{\partial c_u} (c_u^e - c_u)$$

So, at the endowment where \(c_u = c_u^e\) and \(c_e = c_e^e\), the moral hazard terms, \(\frac{\partial \hat{q}}{\partial c_e}\) and \(\frac{\partial \hat{q}}{\partial c_u}\), have no impact on the marginal profitability of providing insurance. Because of this, moral hazard alone cannot cause a complete absence of trade.

2.A.2 Concavity Assumptions

Assumption 2.2.3 maintains that \(\pi\) is globally concave in \((\mu, \Delta)\). Here, we derive sufficient conditions on the primitives of the model that guarantee this concavity. In particular, we show that if \(\Psi''(q; \theta) > 0 \) and \(\frac{w'(c_e^e)}{q(c_e^e)} \leq 2\) then \(\pi\) is globally concave in \((\mu, \Delta)\).

For simplicity, we consider a fixed \(\theta\) and drop reference to it. Profits are given by

$$\pi(\Delta, \mu) = \hat{q} (\Delta) (c_e^e - C_e (\mu - \Psi (\hat{q} (\Delta)))) + (1 - \hat{q} (\Delta)) (c_u^e - C_u (\mu - \Delta - \Psi (\hat{q} (\Delta))))$$

Our goal is to show the Hessian of \(\pi\) is negative semi-definite. We proceed in three steps. First, we derive conditions which guarantee \(\frac{\partial^2 \pi}{\partial \Delta^2} < 0\). Second, we show that, in general, we have \(\frac{\partial^2 \pi}{\partial \mu^2} < 0\). Finally, we show the conditions provided to guarantee \(\frac{\partial^2 \pi}{\partial \Delta^2} < 0\) also imply the determinant of the Hessian is positive, so that both eigenvalues of the Hessian must be negative and thus the matrix is negative semi-definite.

2.A.2.1 Conditions that imply \(\frac{\partial^2 \pi}{\partial \Delta^2} < 0\)

Taking the first derivative with respect to \(\Delta\), we have

$$\frac{\partial \pi}{\partial \Delta} = \frac{\partial \hat{q}}{\partial \Delta} (c_e^e - c_u^e + C_u (\mu - \Delta - \Psi (\hat{q} (\Delta))))$$

$$\quad - (1 - \hat{q} (\Delta)) C'_u (\mu - \Delta - \Psi (\hat{q} (\Delta))) - \hat{q} (\Delta) C'_e (\mu - \Psi (\hat{q} (\Delta)))$$
Taking another derivative with respect to $\Delta$, applying the identity $\Delta = \Psi'(\hat{\Delta})$, and collecting terms yields

\[
\frac{\partial^2 \pi}{\partial \Delta^2} = - \left[ (1 - \hat{\Delta}) (1 + \Delta)^2 C''(\mu - \Delta - \Psi'(\hat{\Delta})) + \hat{\Delta} (\Delta \hat{\Delta}'(\Delta))^2 C''(\mu - \Psi'(\hat{\Delta})) \right] \\
+ \frac{\partial \hat{\Delta}}{\partial \Delta} \left[ (1 - \hat{\Delta}) C'(\mu - \Delta - \Psi'(\hat{\Delta})) + \hat{\Delta} C'(\mu - \Psi'(\hat{\Delta})) \right] \\
+ \frac{\partial^2 \hat{\Delta}}{\partial \Delta^2} \left[ (1 - \hat{\Delta}) \Delta C'(\mu - \Delta - \Psi'(\hat{\Delta})) + \hat{\Delta} C'(\mu - \Psi'(\hat{\Delta})) \right]
\]

We consider these three terms in turn. The first term is always negative because $C'' > 0$. The second term, multiplying $-\hat{\Delta}$, can be shown to be positive if

\[(1 + \hat{\Delta}) C'(\mu - \Delta - \Psi'(\hat{\Delta})) \geq \hat{\Delta} C'(\mu - \Delta)
\]

which is necessarily true whenever

\[
\frac{u'(c_e)}{\lambda'(c_e)} \leq 2
\]

This inequality holds as long as people are willing to pay less than a 100% markup for a small amount of insurance, evaluated at their endowment.

Finally, the third term is positive as long as $\Psi'' > 0$. To see this, one can easily verify that the term multiplying $\frac{\partial^2 \hat{\Delta}}{\partial \Delta^2}$ is necessarily positive. Also, note that $\frac{\partial^2 \hat{\Delta}}{\partial \Delta^2} = -\Psi''$. Therefore, if we assume that $\Psi'' > 0$, the entire last term will necessarily be negative. In sum, it is sufficient to assume $\frac{u'(c_e)}{\lambda'(c_e)} \leq 2$ and $\Psi'' > 0$ to guarantee that $\frac{\partial^2 \pi}{\partial \Delta^2} < 0$.

**2.A.2.2 Conditions that imply $\frac{\partial^2 \pi}{\partial \mu^2} < 0$**

Fortunately, profits are easily seen to be concave in $\mu$. We have

\[
\frac{\partial \pi}{\partial \mu} = - (1 - \hat{\Delta}) C'(\mu - \Delta - \Psi'(\hat{\Delta})) - \hat{\Delta} C'(\mu - \Psi'(\hat{\Delta}))
\]

so that

\[
\frac{\partial^2 \pi}{\partial \mu^2} = - (1 - \hat{\Delta}) C''(\mu - \Delta - \Psi'(\hat{\Delta})) - \hat{\Delta} C''(\mu - \Psi'(\hat{\Delta}))
\]

which is negative because $C'' > 0$. 

116
2.A.2.3 Conditions to imply $\frac{\partial^2 \pi}{\partial \mu^2} - \left( \frac{\partial^2 \pi}{\partial \Delta \partial \mu} \right) > 0$

Finally, we need to ensure that the determinant of the Hessian is positive. To do so, first note that

$$\frac{\partial^2 \pi}{\partial \mu \partial \Delta} = (1 - \hat{q} (\Delta)) C'' (\mu - \Delta - \Psi (\hat{q} (\Delta))) (1 + \Delta \hat{q}' (\Delta)) + \hat{q} (\Delta) C'' (\mu - \Psi (\hat{q} (\Delta))) \Delta \hat{q}' (\Delta)$$

Also, we note that under the assumptions $\Psi'' > 0$ and $\frac{u'(c')}{\nu(c)} \leq 2$, we have the inequality

$$\frac{\partial^2 \pi}{\partial \Delta^2} < - \left[ (1 - \hat{q} (\Delta)) (1 + \Delta)^2 C'' (\mu - \Delta - \Psi (\hat{q} (\Delta))) + \hat{q} (\Delta) (\Delta \hat{q}' (\Delta))^2 C'' (\mu - \Psi (\hat{q} (\Delta))) \right]$$

Therefore, we can ignore the longer terms in the expression for $\frac{\partial^2 \pi}{\partial \Delta^2}$ above. We multiply the RHS of the above equation with the value of $\frac{\partial^2 \pi}{\partial \mu^2}$ and subtract $\left( \frac{\partial^2 \pi}{\partial \Delta \partial \mu} \right)$. Fortunately, many of the terms cancel out, leaving the inequality

$$\frac{\partial^2 \pi}{\partial \Delta^2} \frac{\partial^2 \pi}{\partial \mu^2} - \left( \frac{\partial^2 \pi}{\partial \Delta \partial \mu} \right)^2 \geq (1 - \hat{q} (\Delta)) (1 + \Delta \hat{q}' (\Delta))^2 C'' (\mu - \Delta - \Psi (\hat{q} (\Delta))) C'' (\mu - \Psi (\hat{q} (\Delta)))$$

which reduces to the inequality

$$\frac{\partial^2 \pi}{\partial \Delta^2} \frac{\partial^2 \pi}{\partial \mu^2} - \left( \frac{\partial^2 \pi}{\partial \Delta \partial \mu} \right)^2 \geq \hat{q} (\Delta) (1 - \hat{q} (\Delta)) C'' (\mu - \Delta - \Psi (\hat{q} (\Delta))) C'' (\mu - \Psi (\hat{q} (\Delta))) K (\mu, \Delta)$$

where

$$K (\mu, \Delta) = (1 + \Delta \hat{q}' (\Delta))^2 + (\Delta \hat{q}' (\Delta))^2 - 2 \Delta \hat{q}' (\Delta) - 2 (\Delta \hat{q}' (\Delta))^2$$

$$= 1$$

So, since $C'' > 0$, we have that the determinant must be positive. In particular, we have

$$\frac{\partial^2 \pi}{\partial \Delta^2} \frac{\partial^2 \pi}{\partial \mu^2} - \left( \frac{\partial^2 \pi}{\partial \Delta \partial \mu} \right)^2 \geq \hat{q} (\Delta) (1 - \hat{q} (\Delta)) C'' (\mu - \Delta - \Psi (\hat{q} (\Delta))) C'' (\mu - \Psi (\hat{q} (\Delta)))$$

2.A.2.4 Summary

As long as $\Psi'' > 0$ and $\frac{u'(c')}{\nu(c)} \leq 2$, the profit function is guaranteed to be concave. In practice, Gruber (1997) and Chetty (2008) estimate $\frac{u'(c')}{\nu(c)}$ to be between 1.3 and 1.6. Therefore, our only unsubstantiated assumption for the model is that the convexity of the effort function increases in $p$, $\Psi'' > 0$. An alternative statement of this assumption is that $\frac{\partial^2 \pi}{\partial \Delta^2} < 0$, so that the marginal impact
of work incentives on the employment probability is declining in the size of the work incentives.
Chapter 3

Would the Long-Term Care Insurance Market Survive Adjusted Community Rating Regulation?

3.1 Introduction

Roughly 20% of 65 year olds have a health condition that would lead them to be rejected for long-term care (LTC) insurance (Hendren (2011), Murtaugh et al. (1995)). One potential policy response to this discriminatory practice by insurers is to limit the set of information insurance companies can use to price their policies. So-called “community rating” regulation requires insurance companies to offer contracts uniformly to the population (“pure community rating”) or based on a limited set of characteristics, such as age (“adjusted community rating”). These policies allow those with health conditions to access to the same insurance policies as the healthy.\footnote{Under community rating, insurers are generally required to offer insurance to all who apply.}

Recently, these types of regulations have garnered significant political attention. The Affordable Care Act (ACA), passed in 2010, contains adjusted community rating regulations for the non-group health insurance market.\footnote{This law also contains a mandate requiring individuals to purchase health insurance, which would limit any potential adverse selection induced by the policy.} Indeed, eight states have already enacted community rating policies in their non-group markets (Lo Sasso and Lurie (2009)). Although the goal of community rating is to increase access and equity in insurance markets by preventing the sick from being rejected or charged a higher price, these regulations have the potential to induce significant adverse selection and hinder the workings of the insurance market. Such impacts have been the focus of many previous studies focusing on the health insurance market (Buchmueller and Dinardo (2002), Simon (2005), Herring and Pauly (2006), Lo Sasso and Lurie (2009)).

In long-term care insurance, community rating is also a potential policy option, but has never actually been attempted. The Community Living Assistance Services and Supports program (CLASS
Act), passed as part of the ACA, would have tasked the government with selling long-term care (LTC) insurance on an adjusted-community rating basis. However, this program was eventually abandoned by the Obama administration after actuaries could not establish its long-run budget neutrality. Policy discussions for the future of LTC insurance in the US now focuses on regulatory and tax subsidy solutions that avoid government provision. Community rating is a natural regulation that could provide access to insurance for those currently excluded from the market because of the presence of a pre-existing condition.

This paper evaluates the extent of adverse selection that would occur under an adjusted community rating regulation in the LTC insurance market that forces insurers to ignore health conditions and price only based on age. The advantage of our approach relative to existing literature is that we can conduct the evaluation ex ante: we ask how much adverse selection the market would suffer under the regulation, without actually needing to observe any regulation. Indeed, no such regulation has ever been attempted in LTC.

We use subjective probability elicitations to estimate the distribution of beliefs about future nursing home use conditional on age. We do not require these elicitations to be true measures of beliefs; rather we employ the approach developed in Hendren (2011), which allows these elicitations to be noisy and potentially biased measures of true beliefs. We then use the no-trade theorem of Hendren (2011) to quantify the implicit tax imposed by private information as the markup individuals must be willing to pay for insurance in order for the market to exist. We compare this estimate to existing estimates of the willingness to pay and to the estimates from Hendren (2011) which are argued to lead to a complete absence of trade in LTC.

Our results suggest age-based adjusted community rating regulation would cause the entire LTC insurance market to unravel. We estimate that the implicit tax imposed by private information would be between 50% and 130%, depending on ones age. This generally exceeds existing estimates of the willingness to pay for LTC insurance of 26-62% (Brown and Finkelstein (2008)), and is also of similar magnitude to existing estimates of 65-75% of the implicit tax which is argued to shut down the market for those who are currently rejected in LTC insurance (Hendren (2011)). Thus, our results suggest that if insurance companies could only price based on age, the market for LTC insurance would completely shut down, leaving neither the sick nor the healthy with insurance. Adjusted community rating regulation would lead to a (Pareto) decline in welfare.

Our paper is related to several strands of literatures. Several papers have conducted ex post analyses of community rating regulation in the individual and small group health insurance market (Buchmueller and Dinardo (2002), Simon (2005), Herring and Pauly (2006), Lo Sasso and Lurie (2009)). This literature finds that community rating policies lead to change in the composition of the insured from the healthy to the sick, but do not lead to a complete unraveling of the health insurance market.3

3The one caveat is that Buchmueller and DiNardo find evidence that pure community rating in the small group and individual market is associated with a reduction in the fraction of the population that was insured in New York,
In contrast to those approaches, this paper conducts an *ex ante* evaluation of an adjusted community rating regulation in a market (LTC insurance) that has never attempted such a policy. In this sense, our paper contributes to the literature on ex-ante policy evaluation. Traditionally, this literature involves estimating a structural model and conducting out-of-sample simulation by assuming the stability of certain economic primitives (Wolpin (2007), Todd and Wolpin (2008)). Many recent papers that apply the structural approach to insurance markets and conduct policy analysis (Bundorf et al. (ming), Einav et al. (2010)). For example, Einav et al. (2010) document adverse selection in the UK Annuity market and show that a mandated policy which mandated a longer length guarantee could obtain higher welfare.

However, existing structural approaches cannot readily be used to simulate community rating regulation. The reason is that identification of beliefs in traditional structural approaches rely on revealed preference to identify beliefs. This requires controlling for the set of information insurance companies use to price insurance in order to identify the distribution of beliefs (Chiappori and Salanié (2000)). In contrast, one can use subjective probability elicitation to identify the distribution of beliefs conditional on any chosen set of observable characteristics, and thus conduct policy analysis which changes this set.

We believe this approach to analyzing insurance regulation may have widespread applicability. For example, recent debate over the Affordable Care Act, both in the political sphere and in the Supreme Court, has questioned whether or not the private health insurance market could survive the community rating regulation proposed in the ACA if the mandate were omitted or found to be unconstitutional. The empirical approach taken in this paper is designed precisely to shed light on this type of question. Beyond insurance, many economic settings in labor, credit, and other markets involve a trade-off between redistributive goals and the availability/use of public information. Our approach shows how researchers and policy-makers can evaluate the impact of regulations that depend on the distribution of knowledge in the population.

### 3.2 Insurance Framework

This section discusses the framework developed in Hendren (2011), which uses subjective probability elicitation to estimate properties of what agents' know about their risk. Throughout, we make explicit how the researcher is free to choose the observables used in the empirical analysis and how this allows one to simulate regulations, like community rating, which effectively change the information structure of the economy. Yet in doing so, we do repeat much of the material covered in Hendren (2011); those familiar with this approach could skip to the next section where we apply this framework to assess the impact of adjusted community rating regulation in LTC insurance.

yet they find corresponding declines in Connecticut and Pennsylvania and thus argue that it is not clear to be caused by community rating.
3.2.1 Environment

Agents face the potential of some adverse event, $L$. This event could be an adverse health event, the onset of work-limiting disability, the need for nursing home care, or any other state of the world that has a higher marginal utility of income and thus a demand for insurance. Agents have beliefs about the occurrence of this event, given by the realization of the random variable, $P$. Agents’ beliefs about $L$ (i.e. their realizations of $P$) are unobservable to a potential insurance company, but agents have observable characteristics that could be used to price their insurance. We let $X$ denote the random variable corresponding to this set of public information. With these definitions, our goal is to analyze the impact of regulations affecting the set of observable information, $X$, on the workings of a market for insurance against $L$.

We let $Z$ denote a response to the question: “What is the probability/chance (0-100) that $L$ will occur?”. At no point do we assume people can perfectly report their probabilistic beliefs, $Z = P$. Rather, we conduct two complementary analyses. First, we test for the potential for adverse selection using relatively weak assumptions. We then quantify the impact of the potential adverse selection using assumptions that are stronger but continue to allow the elicitations to be noisy and potentially biased measures of true beliefs.

3.2.2 Identifying the potential for adverse selection

We identify the potential for adverse selection by asking whether agent’s beliefs are predictive of the loss. This approach slightly weakens the assumptions of Hendren (2011), as we do not require agents to have unbiased beliefs in order to identify private information. Instead, we only impose one assumption on the relationship between elicitations and beliefs. We require that agents know $Z$.

Agents know $Z$: $Pr\{L|X, Z, P\} = Pr\{L|X, P\}$

This assumption requires that any information about $L$ that is captured by $Z$ would also have been captured by agents true beliefs, $P$. In other words, if one was trying to forecast the loss, $L$, and knew the agents’ true beliefs, $P$, then also learning $Z$ would not help forecast the occurrence of the loss. Note that we are not assuming that people know how $X$ predicts $L$, nor do we assume beliefs are unbiased; nor do we require that $Z$ is a subjective probability elicitation about $L.$ All that we require is that people can’t report more information (in $Z$) about $L$ than is captured by their true beliefs about $L$.

Given this assumption, we propose a simple test for whether agents have information about $L$ beyond what is captured by a set of observable information, $X$: Is $Z$ predictive of $L$ conditional on

---

4To illustrate this more specifically, we do not require $Pr\{L|X, P\} = Pr\{L|X\}$, which would be the case if people knew $X$. We do not require $Pr\{L|P\} = P$; agents could have biases as in Kahneman and Tversky (1979).
Proposition 3.1. Suppose Assumption 1 holds. Then,

\[ \Pr \{L|X, Z\} \neq \Pr \{L|X\} \implies \Pr \{L|X, P\} \neq \Pr \{L|X\} \]

where \( \Pr \{L|X, Z\} \neq \Pr \{L|X\} \) means there exists a positive mass of realizations of \( Z \), say \( z \), and for \( X \), say \( x \), for which \( \Pr \{L|X = x, Z = z\} \neq \Pr \{L|X = x\} \) and similarly for \( \Pr \{L|X, P\} \neq \Pr \{L|X\} \).

Proof. Follows immediately from the equality:

\[ \Pr \{L|X, Z\} = \mathbb{E}_{P|X,Z} [\Pr \{L|X, P\} | X, Z] \]

Converse. The converse to Proposition 1 is not always true. We need additional assumptions to ensure that if agents have information about \( L \) beyond what is captured in \( X \), then at least a small amount of it is revealed in \( Z \). Although many assumptions could be made to ensure this, we find the following two assumptions quite intuitive, although more restrictive than necessary.

Proposition 3.2. (Partial Converse to Proposition 1) Suppose that (a) \( \Pr \{L|X, P\} \) is weakly increasing in \( P \) and (b) there exists a 1-1 real-valued function \( f(Z) \) such that \( F(P|X, f(Z)) \) is strictly decreasing in \( f(Z) \). Then,

\[ \Pr \{L|X, P\} \neq \Pr \{L|X\} \implies \Pr \{L|X, Z\} \neq \Pr \{L|X\} \]

Proof. We have

\[
\begin{align*}
\Pr \{L|X, Z\} &= \Pr \{L|X, f(Z)\} \\
&= \int \Pr \{L|X, P\} dF_{P|f(Z),X}(P|X, f(Z)) dP \\
&= \Pr \{L|X, P = \tilde{P}\} - \int F_{P|f(Z),X}(P|X, f(Z)) \frac{\partial \Pr \{L|X, P\}}{\partial P} dP
\end{align*}
\]

so that if \( F_{P|X,f(Z)} \) is strictly decreasing in \( f(Z) \) then \( \Pr \{L|X, f(Z)\} \) is strictly increasing in \( f(Z) \).
Assumption (a) requires that those with higher beliefs actually have higher probabilities of experiencing the loss. It would clearly be satisfied if agents had unbiased beliefs, so that $\Pr \{L|X, P\} = P$. But it is weaker. It would be satisfied under the probability transformations used by Kahneman and Tversky (1979), among others. Assumption (b) is a stochastic monotonicity assumption that higher values of $Z$ correspond to higher values of beliefs, $P$. This would be satisfied if agents reported their true beliefs, $Z = P$, but also allows for substantial measurement error in $Z$. In short, under relatively weak conditions, a test for whether or not $Z$ has predictive power for $L$ conditional on $X$ provides a test for the agent having information about $L$ beyond what is captured in $X$. This is our test for the potential for adverse selection.

3.2.3 Quantifying the size of the potential adverse selection problem

We quantify the size of the adverse selection problem in two steps. First, we add additional statistical structure to identify the distribution of beliefs in the population. Second, we add a theoretical structure to translate this distribution into a measure of the implicit tax on insurance premiums individuals must be willing to pay for a market to exist under the regulation.

3.2.3.1 Identifying the Distribution of Beliefs

We make three assumptions to identify the distribution of beliefs conditional on $X$, $F_P (p|X)$. First, we retain Assumption 1 that people know $Z$. Second, we impose unbiased beliefs.

(Unbiased Beliefs) $\Pr \{L|X, P\} = P$

This assumption requires that the empirical probability of the loss occurring, given ones' beliefs, is equal to the beliefs. Implicit in this definition is that agents know $X$, in the sense that the observables, $X$, do not have any additional predictive power for $L$ conditional on the true beliefs, $P$. Although this is a relatively strong assumption, it is quite standard in the literature. Moreover, our assumption does not require that agents know how other observables, outside of what is contained in the chosen set $X$, affect their probability of $L$. For example, if $X$ contains only age, then we do not require that agents know how their medical conditions affect their likelihood of $L$.

Third, we assume that the variable $Z$ is a noisy and potentially biased measure of true beliefs, $P$, where the bias and noise can be parameterized in a parsimonious way. As has been noted in previous literature (Gan et al. (2005), Finkelstein and McGarry (2006)), subjective probability elicitationss tend to concentrate on focal point values of 0, 50, and 100. Therefore, we follow Hendren (2011) and allow the elicitation to have excess density on these focal values. To do so, we parameterize the density of elicitationss given beliefs, $f_{Z|P} (Z|P, X)$ as a mixture of a censored normal and an
ordered probit distribution:

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

where \( \phi \) denotes the standard normal p.d.f. and \( \Phi \) the standard normal c.d.f. We estimate four elicitation error parameters: \( \theta = (\sigma, \lambda, \kappa, \alpha(X)) \). \( \sigma \) captures the dispersion in the elicitation error, \( \lambda \) is the fraction of focal point respondents, \( \kappa \) is the focal point window. We allow the elicitation bias term, \( \alpha(X) \), to vary with the observable variables, \( X \). The censored normal distribution captures the idea that \( Z \) may be a noisy measure of \( P \); the ordered probit allows for additional concentration of mass at 0, 50 and 100 and interprets such concentration as responses of “low”, “medium”, and “high”.

With this statistical assumption on the relationship between elicitations, \( Z \), and true beliefs, \( P \), we estimate the density of beliefs through the equation:

\[
f_{Z,L}(Z, L|X) = \int p^L (1 - p)^{1 - L} f_{Z|P}(Z|P = p, X, \theta) f_P(p|X) \, dp 
\]

where \( f_P(p|X) \) is the p.d.f. of beliefs given \( X \). We then flexibly approximate \( f_P(p|X) \) and estimate it along with the parameters governing the elicitation error, \( \theta \), using MLE. This provides an estimate of the distribution of beliefs given the chosen set of observables, \( X \).

3.2.3.2 Quantification using theory

Given a distribution of beliefs, with c.d.f. \( F_P(p|X) \) and p.d.f. \( f_P(p|X) \), we use theory to quantify the barrier to trade it imposes.

Consider the following environment. Agents are endowed with wealth \( w \) but the occurrence of the loss, \( L \), imposes a consumption loss of \( l \) units. All agents have vNM preferences, \( u \). In the absence of insurance, an agent with beliefs \( p \) has utility given by

\[
p u (w - l) + (1 - p) u (w)
\]

We ask under what conditions (on \( u \) and \( F_P(p|X) \)) agents with observable characteristics, \( X \), can obtain any insurance. The answer, provided in Hendren (2011), is that insurance companies will not sell insurance to anyone if

\[
\frac{u'(w - l)}{u'(w)} \leq \inf_{p \in \Phi_X \{1\}} T_X(p)
\]

(3.2)
where $\Psi_X$ is the support of the distribution of $P$ given $X$, and $T_X(p)$ is given by

$$T_X(p) = \frac{E \{ P \mid X, P \geq p \}}{1 - E \{ P \mid X, P \geq p \}} \frac{1 - p}{p}$$

Equation 3.2 says that trade can only occur if someone is willing to pay the pooled cost of worse risks in order to obtain a small amount of insurance. To see this, note that the value $u'(w-l)/u'(w) - 1$ is the markup individuals would be willing to pay for a small transfer of resources from the event of the loss not occurring to the event of the loss occurring. The value $T_X(p) - 1$ is the markup that would be imposed on a type $p$ if the contract had to cover the cost of all higher risks, $P \geq p$, also selecting the contract. The smallest such markup is then given by $\inf_{p \in \Psi_X \backslash \{1\}} T_X(p) - 1$, which we call the implicit tax imposed by $F_p(p\mid X)$. If no one is willing to pay this implicit tax, the insurance market in segment $X$ unravels. Any contract or menu of contracts would be so heavily adversely selected that they would not deliver positive profits at any price. Thus, the implicit tax imposed by $F_p(p\mid X)$, given by $\inf_{p \in \Psi_X \backslash \{1\}} T(p) - 1$ is a natural measure of the magnitude of the adverse selection problem.

3.2.4 Summary

The approach uses subjective probability elicitation, $Z$, to identify properties of what people in the population know about $L$ conditional on a chosen set of observables, $X$. By matching $X$ to the set corresponding to the proposed regulation, we identify, under weak assumptions, whether people know anything about $L$ conditional on $X$ and, with additional assumptions, quantify the magnitude of the adverse selection as the implicit tax people would need to be willing to pay for the market to exist. We now illustrate our approach with an application to the LTC insurance market.

3.3 Impact of Regulation in LTC

We apply the empirical approach to assess the hypothetical impact of adjusted community rating on the workings of the LTC insurance market.

3.3.1 Data and Sample

Our data come from the Health and Retirement Study (1993-2008), a comprehensive survey of individuals over age 55 and their spouses. Respondents are asked a range of demographic and health questions, along with a battery of subjective probability elicitation. In particular, we let $Z$ denote a response to the question “What is the chance (0-100) that you will enter a nursing home in the next 5 years?”. Corresponding to this elicitation, we let $L$ denote the event that the individual goes to the nursing home in the subsequent 4-5 years\(^5\). We then let $X$ denote an individual’s age,

\(^5\)Specifically, $L$ is an indicator for nursing home entry in the two subsequent waves of the survey.
Table 3.1: Sample Summary Statistics

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Std Dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>NursHome (5y)</td>
<td>0.058</td>
<td></td>
</tr>
<tr>
<td>Subj Prob</td>
<td>0.125</td>
<td>0.207</td>
</tr>
<tr>
<td>Age</td>
<td>71.800</td>
<td>4.126</td>
</tr>
<tr>
<td>Female</td>
<td>0.594</td>
<td></td>
</tr>
<tr>
<td>Arthritis</td>
<td>0.544</td>
<td></td>
</tr>
<tr>
<td>Diabetes</td>
<td>0.161</td>
<td></td>
</tr>
<tr>
<td>High BP</td>
<td>0.544</td>
<td></td>
</tr>
<tr>
<td>Own LTC Ins</td>
<td>0.139</td>
<td></td>
</tr>
<tr>
<td>Sample Size</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Obs</td>
<td>23,292</td>
<td></td>
</tr>
<tr>
<td>Individuals</td>
<td>11,019</td>
<td></td>
</tr>
<tr>
<td>Households</td>
<td>8,385</td>
<td></td>
</tr>
</tbody>
</table>

thereby focusing our estimation on what agents know about $L$ beyond what is captured in their age.

Our sample starts with all individuals aged 65-79 who do not currently reside in a nursing home. The subjective probability elicitation are asked only among those 65 and older. Above age 79, a LTC insurance market does not currently exist. Indeed, Hendren (2011) argues this is the result of private information. Further restricting the set of observables insurance companies use in pricing would presumably not lead to this market opening up. Therefore, we restrict our attention to ages below 80, where the market currently exists for those without health conditions which lead to rejection. We also restrict the sample to those who we can follow for a subsequent 5 years of the survey to construct $L$ and for whom we have data on the subjective probability elicitation, $Z$.6

Table 3.1 presents the sample summary statistics. The sample is 59% female and has an average age of 72. Our sample consists of 23,292 observations from 11,019 unique individuals and 8,385 unique households. We include multiple observations per individual throughout the panel and cluster the standard errors at the household level.7

We include the 14% of our sample who own LTC insurance. Including this group could cause a problem if people who own insurance are more likely to go to a nursing home. In this case, heterogeneous LTC insurance purchase could induce heterogeneity in nursing home use which we could capture as belief heterogeneity. However, multiple studies suggest that moral hazard is not a significant issue in nursing home usage. Those who purchase LTC insurance are no more likely

---

6Our sample is the same as the LTC sample in Hendren (2011) but combines the "reject", "no reject", and "uncertain" categories, but restricts attention to ages below 80.

7Including repeated observations from individuals throughout the panel does not induce any bias in estimation because we observe them with different observables (i.e. different ages) at each point in the panel.
to go to a nursing home (Finkelstein and McGarry (2006)) and those covered under Medicaid expansions are no more likely to go to a nursing home (Grabowski and Gruber (2007)). Thus, we are comfortable including them in our sample.

Figure 3-1 presents the histogram of the subjective probability elicitations, Z. As has been noted in previous literature (Gan et al. (2005), Finkelstein and McGarry (2006)), responses tend to concentrate on focal point values, especially at 0. Our approach to identifying private information does not make any specific assumptions about what generates these elicitations. In contrast, our approach to quantifying the distribution of beliefs assumes that the specific elicitation structure, provided in Section 3.2.3.1, generates this histogram from the underlying distribution of beliefs.

**Continuous versus binary risk** Our empirical approach investigates the properties of what agents know about $L$, which is the binary event of going to a nursing home in the next five years. Of course, nursing home risk extends beyond this binary event. Individuals face risks over the length of stay and whether or not they enter a nursing home in 6, 7, or even 25 years. So how should one think of our empirical approach which focuses on binary losses in a world where the risk for long-term care expenses is continuous?

Our approach *identifies* the presence of asymmetric information for any insurance contract that provides payment in the event $L$ occurs. In this sense, identifying asymmetric information is not limited by only focusing on a binary event. However, our approach *quantifies* the amount of adverse selection in a more specific manner: we estimate the implicit tax imposed by asymmetric information on a hypothetical market that provides an additional dollar in the event that $L$ occurs.\(^8\)

### 3.3.2 Identification of Asymmetric Information

Hendren (2011) shows that individuals with health conditions, such as activity of daily living (ADL) restrictions, have private information about their future nursing home risk. Therefore, one can already conclude that individuals have private information conditional on age. But for completeness,

\(^8\)The implicit tax for related events other than $L$ may be different, and could be estimated with additional elicitations and corresponding loss information.
we present the results of our test for the identification of private information conditional on age. Recall that we wish to know whether the subjective probabilities, $Z$, can explain the realized loss, $L$, conditional on age, $X$. For our primary specification, we consider a probit of $L$ on $Z$:

$$\Pr\{L|X, Z\} = \phi(X\beta + \gamma Z)$$

where $X$ consists a quadratic function of age. We also consider a linear specification where $X$ is a quadratic function of age and a linear specification where $X$ contains a dummy variable for each age.

**Results** Table 3.2 presents the coefficients, $\gamma$, along with their p-values. We reject the null hypothesis that agents have no information beyond what is captured by their age across specifications. The logit rejects with a p-value of less than 0.001, the linear specification with age and age squared rejects with a p-value of 0.0016, and the fully-saturated age model rejects with a p-value of 0.0136. In short, agents know more about their future risk of going to a nursing home than what is captured by their age. Thus, a market which only allowed insurance companies to price based on age would face a potential adverse selection problem.

### 3.3.3 Quantification of Asymmetric Information

To quantify the impact of adverse selection in the regulated market, we now adopt the assumptions outlined in 3.2.3.1. We flexibly approximate $F(p|X)$ and estimate it using equation (3.1) as a likelihood function. The estimated parameters are the elicitation distribution parameters ($\alpha(X), \sigma, \kappa, \lambda$) and the parameters used in the flexible approximation of $F(p|X)$, which we now specify.

---

9 This result is consistent with the findings of Hendren (2011) that those who have health conditions that would prevent them from being able to purchase insurance have private information about their risk.
Approximation for $F(p|X)$  We approximate $F(p|X)$ with a mixture of beta distributions, along with a point-mass distribution.\(^{10}\)

$$F(p|X) = w_1 \{ p \leq \Pr\{L|X\} - a \} + (1 - w) \sum \xi_i \text{Beta}(\mu_i(\Pr\{L|X\}), \psi_i)$$

where $w$ is the weight on the point-mass, $\Pr\{L|X\} - a$ is the mean of the point-mass. The term $\text{Beta}(\mu_i(\Pr\{L|X\}), \psi_i)$ is the Beta c.d.f. with mean $\mu_i(\Pr\{L|X\})$ and shape parameter $\psi_i$, and $\{\xi_i\}_i$ are the weights on the Beta distributions.\(^{11}\) We assume the mean of the beta distributions vary linearly with $\Pr\{L|X\}$, $\mu_i(X) = \gamma_i + \pi_i \Pr\{L|X\}$.\(^{12}\) We use two Beta distributions for our results.

Our specification for $F(p|X)$ is flexible and nests the null hypothesis of no information, $F(p|X) = 1 \{ p < \Pr\{L|X\} \}$. Moreover, we can easily accommodate the restriction $\mathbb{E}[P|q] = q$ imposed by unbiased beliefs. That said, future work could explore an even more flexible specification.

Given an estimate of $F(p|X)$, estimation of the implicit tax must account for the fact that estimating $\mathbb{E}[P|X, P \geq p]$ for values of $p$ in the upper quantiles of $F(p|X)$ suffers an extremal quantile estimation problem. As $p$ increases, there is less effective data with which to construct the average of $P > p$. Thus, we follow Hendren (2011) and restrict attention to values of $p$ less than the $\tau$-th quantile of $F(p|X)$, $p < F^{-1}(\tau|X)$, and estimate $\inf_{p \in [0,F^{-1}(\tau|X)]} T_X(p)$. We then assess the robustness of our results to the choice of $\tau$.\(^{13}\)

Results  Figure 3-2 presents the estimated $F(p|X)$ evaluated at age $X = 72$ (the age which induces mean loss probability in our sample, $\Pr\{L|X\} = 0.058$). The results suggest a significant fraction with relatively homogeneous beliefs, along with the presence of a smaller number of people who have a higher risk of going to a nursing home. Indeed, this is consistent with the findings of Hendren (2011) which shows that those who are able to purchase insurance have no private information (i.e. are represented by a point-mass distribution). But those who would be rejected (roughly 20% of the population) are higher risk and do have private information, generating the presence of an upper tail.

Table 3.3 reports the estimates of the implicit tax, $\inf_{p \in [0,F^{-1}(0.8|X)]} T_X(p) - 1$, for various ages.

---

\(^{10}\)The point mass captures the possibility that a fraction of the population has the same information set.

\(^{11}\)The p.d.f. of a beta distribution is given by $f(p; \alpha, \beta) = \frac{\beta(\alpha, \beta)}{B(\alpha, \beta)} p^{\alpha-1}(1-p)^{\beta-1}$ where $B(\alpha, \beta)$ is the beta function. The mean of the beta distribution is given by $\mu = \frac{\alpha}{\alpha + \beta}$ and the shape parameter is $\psi = \alpha + \beta$. In principle, the point-mass distribution could be replaced with a Beta distribution with a very low variance. In practice, the point-mass performs better since numerical integration of low variance beta distributions is computationally time-consuming.

\(^{12}\)We also impose restrictions on the parameter space to ensure the support lies over the interval $[0, 1]$. For example, we allow the parameter, $a$, to vary with $q$ to ensure that the point mass does not fall below $p = 0$ or above $p = 1$ and we allow the means of the beta distributions, $\mu_i(X)$, to be censored on $[0, 1]$.

\(^{13}\)While the primary motivation for such a restriction is statistical, there is a straightforward economic rational for restricting attention to $p \leq F^{-1}(\tau)$. If insurance companies must attract a non-trivial fraction $1 - \tau > 0$ to any given consumption bundle other than the endowment, then $\inf_{p \in [0,F^{-1}(\tau)]} T(p)$ characterizes the barrier to trade (See Hendren (2011), Remark 1).
Figure 3-2: CDF of Private Information

Table 3.3: Minimum Pooled Price Ratio, inf T(\(p\))

| Age (X) | LTC (Mean Pr(L|X)) | 20th Quantile | Median | 80th Quantile |
|---------|---------------------|---------------|--------|---------------|
| inf T   | 1.673               | 2.288         | 1.847  | 1.479         |
| 5%      | 1.600               | 2.097         | 1.699  | 1.417         |
| 95%     | 1.745               | 2.479         | 1.980  | 1.542         |
| Pr(L|X)   | 0.058               | 0.034         | 0.050  | 0.083         |

Note: 5/95% CI computed using bootstrap block re-sampling at the household level (N=150 Reps); Bootstrap CI presents maximum boundaries from a) non-accelerated bias-corrected procedure from Efron (1982) and b) studentized values.

Table 3.4: Robustness to Choice of \(\tau\)

| Quantile Region: \(\Psi\) | LTC (Mean Pr(L|X)) | 0-70% | 0-80% | 0-90% |
|---------------------------|---------------------|-------|-------|-------|
| Reject                    | 1.673               | 1.673 | 1.673 |
| 5%                        | 1.600               | 1.600 | 1.600 |
| 95%                       | 1.745               | 1.745 | 1.745 |

Note: 5/95% CI computed using bootstrap block re-sampling at the household level (N=150 Reps); Bootstrap CI presents maximum boundaries from a) non-accelerated bias-corrected procedure from Efron (1982) and b) studentized values.
Table 3.5: Elicitation Error Parameters

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Standard Deviation s.e.</td>
<td>0.124 (0.001)</td>
</tr>
<tr>
<td>Fraction Focal Respondents s.e.</td>
<td>0.628 (0.004)</td>
</tr>
<tr>
<td>Focal Window s.e.</td>
<td>0.259 (0.002)</td>
</tr>
</tbody>
</table>

Note: Bootstrapped standard errors computed using block re-sampling at the household level (results shown for N=150 repetitions)

The results suggest private information imposes a barrier to trade equivalent to an implicit tax of 67.3% for those aged 73 (5/95% CI: [60%, 74.5%]) and ranges from 47.9% at age 76 to 128.8% at age 68. Unless people are willing to pay this implicit tax, the impact of the adjusted community rating policy would completely unravel the market.

Comparison to willingness to pay Are people willing to pay this implicit tax? Two pieces of evidence suggest they are not. First, the calibrated model of Brown and Finkelstein (2008) suggest people are willing to pay roughly 26-62% markups on insurance policies, which falls generally below our estimates. Second, Hendren (2011) argues that implicit taxes of 65-75% currently lead to an unraveling of the market for LTC insurance among those who are currently rejected. This suggests, by revealed preference, that the markup people are willing to pay is below 65-75%. Thus, it is likely that the LTC insurance market would completely unravel if insurance companies were only allowed to use age to distinguish amongst applicants.

Robustness to choice of $\tau$ and estimates of $\theta$ The estimates in Table 3.3 focus on the value of $\inf_{P\in[0,F^{-1}(\tau|X)]} T(p) - 1$ over the domain up to $\tau = 0.8$, the 80th quantile of the distribution of $P$ given $X$. Table 3.4 shows that these estimates are quite robust to changes in $\tau$. The estimated minimum of the pooled price ratio does not occur on the upper boundary of the restricted support, so that changes in $\tau$ do not affect the estimated implicit tax.

Table 3.5 presents the estimated parameters $\theta$ for the distribution of elicitations given beliefs, $f(Z|P,X,\theta)$. We find a standard error of 0.124, which suggests people are not able to report their true beliefs perfectly in their elicitations. We also estimate a sizable fraction of focal point respondents of 62.8%, which is consistent with the presence of a large fraction of responses at 0, 50, and 100, as shown in Figure 3-1. We estimate a focal window of 0.259, which suggests focal point respondents who otherwise would have reported a value of $Z$ less than 0.259 instead collapse their report to $Z = 0$, generating the excess mass at $Z = 0$.

---

14Of course, this comparison requires the values of $u'(w-1)/u'(w)$ (e.g. risk aversion) among those currently rejected to be similar to those who can buy insurance.
Limitations There a couple of important limitations to our analysis. First, our sample is limited to those 65 and older. Although 65 is roughly the average purchase age for LTC insurance (CBO (2004), Brown and Finkelstein (2008)), it may be the case that such regulation would not shut down the LTC insurance market for applicants below 65. In the extreme, it may only distort the timing of purchase and not have significant welfare implications. Such concerns with our analysis could easily be alleviated with additional data from those below age 65. Second, our quantification measure is the implicit tax on a hypothetical market for a contract which pays $1 in the event of going to a nursing home in the subsequent 5 years. LTC insurers may be able to redesign contracts, perhaps through waiting periods or payment caps, that counteract the impact of adverse selection. To analyze such policies, we would require additional data to construct the relevant loss, \( L \), and corresponding elicitation, \( Z \).\(^{15}\) However, the estimated implicit tax is similar in size to what Hendren (2011) finds is large enough to shut down the entire market for those with pre-existing conditions; thus one could be skeptical that redesigning the policy with waiting periods or other features could overcome the adverse selection problem. At a minimum, our results suggest great caution in attempting adjusted community rating in the LTC insurance market.

3.4 Conclusion

This paper evaluated the impact of a hypothetical adjusted community rating regulation in the LTC market that would force insurers to provide equal treatment to those of the same age. Our results suggest such a policy would lead to significant adverse selection and would likely lead to a complete unraveling of the LTC insurance market. Thus, we conclude such a policy would lead to a (Pareto) reduction in welfare.

We hope this paper illustrates the usefulness of belief elicitation in assessing the impact of regulation in economies with informational asymmetries. Although our approach is developed in the context of insurance, we believe it could be applied to other settings in which a policy maker is interested not only in choosing price schedules, but also the set of public information available for use in the economy.

\(^{15}\)For example, one could test whether a 1 year waiting period would help if we had elicitation of 1 year entry into a nursing home, in addition to the 5 year entry elicitation. We leave these multiple dimensional design questions as an interesting and perhaps important direction for future work.
Bibliography


135


