NBER WORKING PAPER SERIES

# INFERRING RISK PERCEPTIONS AND PREFERENCES USING CHOICE FROM INSURANCE MENUS: THEORY AND EVIDENCE

Keith Marzilli Ericson
Philipp Kircher
Johannes Spinnewijn
Amanda Starc

Working Paper 21797
http://www.nber.org/papers/w21797

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2015

We would like to thank Richard Blundell, Ian Crawford, Mark Dean, Geert Dhaene, Liran Einav, Phil Haile, Arthur Lewbell, Bernard Salanie and Frans Spinnewyn for helpful comments and discussions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w21797.ack

Inferring Risk Perceptions and Preferences using Choice from Insurance Menus: Theory and Evidence

Keith Marzilli Ericson, Philipp Kircher, Johannes Spinnewijn, and Amanda Starc

NBER Working Paper No. 21797
December 2015
JEL No. D8,I13

## ABSTRACT

Demand for insurance can be driven by high risk aversion or high risk. We show how to separately identify risk preferences and risk types using only choices from menus of insurance plans. Our revealed preference approach does not rely on rational expectations, nor does it require access to claims data. We show what can be learned non-parametrically from variation in insurance plans, offered separately to random cross-sections or offered as part of the same menu to one cross-section. We prove that our approach allows for full identification in the textbook model with binary risks and extend our results to continuous risks. We illustrate our approach using the Massachusetts Health Insurance Exchange, where choices provide informative bounds on the type distributions, especially for risks, but do not allow us to reject homogeneity in preferences.

Keith Marzilli Ericson
Boston University School of Management
595 Commonwealth Avenue
Boston, MA 02215
and NBER
kericson@bu.edu

Philipp Kircher
Department of Economics
University of Edingburgh
30 Buccleuch Place,
Edinburgh, EH8 9JT, UK
philipp.kircher@gmail.com

Johannes Spinnewijn
Department of Economics
London School of Economics, Houghton Street
London, WC2A 2AE
J.Spinnewijn@lse.ac.uk

Amanda Starc
The Wharton School
University of Pennsylvania
3641 Locust Walk
Room 303 CPC
Philadelphia, PA 19104
and NBER
astarc@wharton.upenn.edu

# Inferring Risk Perceptions and Preferences using Choice from Insurance Menus: Theory and Evidence

Keith Marzilli Ericson
Boston University

Philipp Kircher
University of Edinburgh

Johannes Spinnewijn
London School of Economics

Amanda Starc*
Wharton School

December 5, 2015

**Abstract**

Demand for insurance can be driven by high risk aversion or high risk. We show how to separately identify risk preferences and risk types using only choices from menus of insurance plans. Our revealed preference approach does not rely on rational expectations, nor does it require access to claims data. We show what can be learned non-parametrically from variation in insurance plans, offered separately to random cross-sections or offered as part of the same menu to one cross-section. We prove that our approach allows for full identification in the textbook model with binary risks and extend our results to continuous risks. We illustrate our approach using the Massachusetts Health Insurance Exchange, where choices provide informative bounds on the type distributions, especially for risks, but do not allow us to reject homogeneity in preferences.

## 1 Introduction

Two key sources of variation in the demand for insurance are variation in risk preferences (e.g., degree of risk aversion) and variation in risk types (e.g., probability of making a claim). The distinction between the two is crucial for positive and normative analysis (e.g., Einav, Finkelstein and Cullen 2010, Chetty and Finkelstein 2013). Adverse selection, in which consumers select into insurance plans based on expected expenditure, can lead to market unravelling and inefficiently low coverage. In contrast with heterogeneity in risks, preference heterogeneity cannot cause insurance markets to be adversely selected. In fact, recent empirical work finding advantageous selection (in

*We would like to thank Richard Blundell, Ian Crawford, Mark Dean, Geert Dhaene, Liran Einav, Phil Haile, Arthur Lewbell, Bernard Salanié and Frans Spinnewyn for helpful comments and discussions.

which low risk individuals purchase more generous insurance plans) has been considered evidence for the importance of preference heterogeneity (e.g., Cutler, Finkelstein, and McGarry 2008).

A growing empirical insurance literature estimates heterogeneity in both preferences and in risk types using data on plan choices and insurance claims (see Einav, Finkelstein and Levin 2010, Barseghyan, Molinari, O'Donoghue and Teitelbaum 2015a). This approach is data-demanding compared to standard demand estimation. More importantly, using this type of data, separate identification of preferences and risks requires assumptions about individuals' (rational) expectations over the distribution of claims. Yet in the context of health insurance, for example, consumers may not understand how different health states map into health expenditures due to the opacity of health care prices (Leiber 2014). They may be overconfident about their own health states (Grubb 2015) and underweight small probability events (Johnson et al. 1993). Given these constraints and biases, it seems reasonable to believe that consumers may have distorted perceptions of the risk they face. If consumers do not have rational expectations over the distribution of their future claims, claims data cannot help to separate a low degree of risk aversion from overoptimistic beliefs about risk.

We present an alternative framework, based on revealed preference, that shows how to identify individuals' risk perceptions and preferences from choice data alone. Our approach does not rely on rational expectations, nor does it require claims data, but exploits variation in the plans from which individuals can choose. This new framework allows us to revisit the question how important preference heterogeneity is for the observed variation in insurance choices. It also provides an alternative approach to estimating perceived risks. Our estimates could be used, in combination with claims data, to assess the accuracy of risk perceptions, which is relevant for welfare and policy analysis in insurance markets (Spinnewijn 2012, Handel, Kolstad and Spinnewijn 2015).

The key challenge in inferring risk perceptions and preferences from insurance choices is that both high risk and risk aversion increase the willingness to buy insurance. To overcome this challenge, we use variation in insurance plan characteristics that differentially attract individuals along the risk and preference dimension. The identification argument in our model exploits the fact that marginal willingness to buy insurance is more rapidly decreasing in coverage for individuals with high risk aversion (but low risk) than for individuals with low risk aversion (but high risk). As a consequence, two plans that differ in their coverage level and premiums can differentially attract individuals along the risk and preference dimension. Our approach can provide informative bounds on preference and risk types using cross-sectional data on individuals choosing from a single menu of plans. The approach is more powerful when applied to choice data from similar populations facing different menus of plans. Such variation in insurance options for otherwise identical populations might be driven by differences in the regulatory environment, by differences in costs of insurance provision

(across states or time), or by differences in market power of insurance providers.[1]

We convey the key intuition for identification in a simple model with binary risks and binary choices with cross-sectional variation in the choice sets. In this binary choice setting, data on insurance choices from a single menu is insufficient to reject homogeneity in risks or preference (even if heterogeneity was substantial). The observation of cost realizations cannot help either to reject preference homogeneity if the market is adversely selected. But with similar individuals facing different menus, the difference in plan shares under the menus allow us to put bounds on the distribution of both risk types and risk preferences. In order to achieve this partial identification, individuals must face a menu in which plans that provide more coverage charge a higher price, but are cheaper per unit of coverage. As a result, such plans will attract high-risk, but low risk-aversion types. In the absence of such variation, it is impossible to reject homogeneity in preferences, even if claims data is observed and expectations are rational. Data on plan choices allow us to overcome another challenge when working with claims data of inferring heterogeneity in (ex ante) risk types from (ex post) risk realizations. Without additional structure on the heterogeneity in risk types, such inference can only be made from plan choices.

Remaining with binary risks, we demonstrate the potential of plan variation for identification in the standard textbook insurance model. Here, individuals decide how much coverage to buy at a constant price per unit of coverage. This can be represented as choices among binary sets with high vs. low coverage, but with a large number of such choice sets this conveniently reduces to the textbook model where individuals may choose any amount of coverage at the specified unit price. As risk aversion determines the gradient of the marginal willingness to pay with respect to coverage, it also determines the *change* in preferred coverage when the unit price of coverage changes, while both an individual's risk and risk aversion determine the agent's preferred coverage *level.* We show how the joint distribution of binary risks and CARA preferences can be non-parametrically identified exploiting price variation in the textbook model. Full identification would require infinite price variation. Limited price variation already suffices to identify key moments capturing the heterogeneity in both dimensions.

We then extend the model beyond binary risks and choice sets to settings that more closely resemble actual health insurance coverage. In practice, health insurance plans are characterized by deductibles, co-insurance rates and out-of-pocket maxima. Individuals of different preference and risk types care differentially about covering different types of expenses. For example, the decreasing returns to coverage imply that individuals with high risk aversion care more about reducing high out-of-pocket expenses (e.g., a decrease in the out-of-pocket maximum) than the reduction of out-of-pocket expenses

[1] Standard revealed preference arguments are based on the same individual choosing consumption bundles at different prices. In insurance markets we rarely have such data: insurance options for individuals often change when the characteristics of the individual changed and individuals' responses may not reflect their preference ranking due to inertia (Handel 2013).

that are already low (e.g., a decrease in the deductible). We also illustrate how the same plan variation drives identification when all plans are offered within one menu to a single cross-section of individuals. The key intuition is the same as in the case with cross-sectional variation in binary choice sets, as different plans need to differentially attract types along the different dimensions. The obtained bounds are different though. *Within-menu* plan variation naturally arises in many practical settings, which is also what we exploit in our empirical analysis.

We apply our method to choice data from the Massachusetts Health Insurance Exchange. We find informative bounds on the distribution of preferences and risks exploiting variation in the features of the contracts offered. Interestingly, we cannot reject homogeneity in preferences as the different plan choices can be rationalized with only heterogeneity in risks. However, we do reject homogeneity in risks. The required variance in risks increases as we restrict the analysis to reasonable preference parameters. We then compare our bounds to estimates from the existing literature. Our application shows what can be learned from choice data alone and highlights the strengths of the revealed preference approach.

Our paper is motivated by a recent, but prominent literature analyzing heterogeneity in preferences and risks empirically, reviewed in Einav, Finkelstein and Levin (2010) and Barseghyan et al. (2015a). This literature started with empirical tests for asymmetric information in insurance markets, often finding a weak relationship between risk type and insurance choice (see Chiappori and Salanié, 2014 and Cohen and Siegelman, 2010). This inspired a new series of papers estimating the heterogeneity in risk preferences jointly with the heterogeneity in risk types in different markets. Examples are auto insurance (Cohen and Einav 2007), annuities (Einav, Finkelstein and Schrimpf 2010) and health insurance (Bundorf, Levin and Mahoney 2012). These studies use choice and claims data to estimate a structural model of heterogeneity and find significant preference heterogeneity. Our work starts from a similar model of consumer choice in which individuals choose insurance plans that maximize their expected utility given their specific risk and preference parameters. Our approach, however, does not require the additional structure on heterogeneity and relaxes the assumption of rational expectations.

A separate empirical literature documents deviations from the standard model of insurance demand. For example, Abaluck and Gruber (2011) find that individuals buying Medicare Part D are over-responsive to prices relative to expected coverage and are not responsive to the variance in out-of-pocket expenses. Sydnor (2010) demonstrates that small distorted probabilities could explain deductible choices in auto and house insurance, while with rational expectations extreme risk aversion would be needed. The identification challenges in the absence of rational expectations have also been addressed using survey data about beliefs (Manski 2004). Handel and Kolstad (2015) complement insurance data with surveys and demonstrate the importance of informa-

tion frictions more generally in explaining the choice of health insurance plans. Most similar in spirit to our paper is Barseghyan et al. (2013), who use insurance choices to separate probability distortions and risk preferences. They extend their benchmark analysis to incorporate unobserved heterogeneity, but rely on parametric assumptions about the joint distribution. In a contemporaneous review of the literature, Barseghyan et al. (2015a) highlight the use of decreasing returns to coverage to separate risk perception and preference and state how "to date, point identification of multidimensional heterogeneity in risk preferences has relied upon parametric assumptions about their joint distribution. It remains a question for future research, to find a field setting and the proper set of assumptions to obtain nonparametric identification."

Following the Revealed Preference (RP) paradigm, our work uses only choices and relies on price/plan variation for identification. We aim to recover preferences, but also risk perceptions. Our methodology is different from standard empirical RP techniques (see Crawford and De Rock, 2014) and our focus is to uncover heterogeneity in types and does not require multiple observations of the same individuals.[2,3] Closely related to our work, Chiappori et al. (2009) analyze the non-parametric identification of preferences under risk from discrete choices, using variation in menus, but again relying on data on risk realizations.[4] Choi et al. (2007) find important preference heterogeneity for choices under risk using experimental variation of prices, but relying on risks with known probabilities. Barseghyan et al. (2015b) use insurance choices of the same individual across different domains to partially identify both risk preferences and risk distortions. Our paper illustrates how both plan variation across menus (offered to multiple cross-sections) and within a menu (offered to one cross-section) can be used for identification. We provide conditions for both partial and full identification.

The paper is organized as follows. Section 2 sets up our choice model and defines our object of interest for identification. Section 3 analyzes the identification of type heterogeneity in a stylized model with binary risk and binary choices. We briefly extend these insights beyond our stylized model in Section 4 and apply them using insurance choices on the Massachusetts Health Exchange in Section 5. We discuss key steps of our proofs in the main text, and provide the formal proofs in the Appendix.

## 2 Setup

We consider a model in which agents decide which plan to buy to provide insurance against a particular risk. Agents differ in two dimensions: the risks they face and the

[2] Exceptions in the RP literature are Crawford and Pendakur (2013), who study the minimum number of types necessary to explain observed choices in cross-sectional data, and Dean and Martin (2015), who study the largest subset of the data which is consistent with homogeneous preferences.

[3] Recent examples in the RP literature that allow for deviations from rational demand are Crawford (2010), Adams et al. (2014) and Caplin and Dean (2015).

[4] Relatedly, Chiappori et al. (2012) and Gandhi and Serrano-Padial (2014) use aggregate data on horse races to either estimate heterogeneity in preferences or heterogeneity in beliefs.

preferences they have. The goal is to identify the population heterogeneity in both dimensions. The identification challenge is that differences in plan choices may be driven by differences in risks, preferences or both. This section describes the most general model, while subsequent sections concentrate on prominent special cases.

**Risk and Preference.** Each agent faces an uncertain cost $k \in \mathbb{R}$ to which he assigns cumulative distribution $F(k|\pi)$. The agent's perceived risk type is denoted by $\pi$, which ranks agents by their riskiness as defined by first-order stochastic dominance. This implies that for two types $\pi_1 > \pi_2$, $F(k|\pi_1) \leq F(k|\pi_2)$ for all $k$. Let $\Pi$ denote the domain of possible risk types.

An insurance plan $X$ has a premium $P$ and maps each cost $k$ into an out-of-pocket expense $x(k) \leq k$. The agent's preferences are represented by expected utility with differentiable Bernoulli-utility function $u(-x|\sigma)$ for an out-of-pocket expense $x$. The agent's preference type is denoted by $\sigma$, which ranks individuals by their preference for risk following Pratt (1964). This is naturally the case for CARA preferences with $u(-x|\sigma) = -\exp(\sigma x)/\sigma$. That is, for any two individuals, if $\sigma_1 > \sigma_2$, individual 1 is more risk-averse than individual 2. We re-scale the preference type $\sigma$ such that for a risk-neutral agent $\sigma = 0$ and types with $\sigma \to \infty$ are infinitely risk averse, so that the domain of possible preference types $\Sigma$ coincides with $\mathbb{R}_+$.[5] The expected utility of a plan $X$ for an agent of risk-preference-type $(\pi, \sigma)$ is

$$U(X|\pi,\sigma) \equiv \int u(-P - x(k)|\sigma)\, dF(k|\pi). \tag{1}$$

**Choice from Menu** We want to infer types from observed *choices* by different individuals. Let $C^i$ denote the choice of individual $i$. The choice is meaningful only relative to the *menu* of options that the individual can choose from. If individuals are offered the choice between insurance plans $X_1, X_2, ..., X_N$, we denote this by the choice menu $\mathcal{M} = \{X_1, X_2, ..., X_N\}$. If different individuals are offered different menus, let $\mathcal{M}^i = \{X_1^i, X_2^i, ..., X_N^i\}$ denote the menu offered to individual $i$. We assume either that different individuals face the same menu as in most of the existing insurance literature, or we assume that different plans are offered randomly to individuals so that the pool of individuals that face each menu has the same underlying distribution of preferences and risks. We drop the agent indicator $i$ when the context is obvious.

Each agent is offered a menu $\mathcal{M}^i$ before her cost is realized. The agent's choice $C^i$ identifies the plan within the menu that gives her the highest expected utility:

$$C^i \in \arg\max_{X \in \mathcal{M}^i} U(X|\pi,\sigma). \tag{2}$$

[5] Convergence to infinite risk aversion means that for any two gambles where the lowest possible outcome in the first gamble is higher than the lowest possible outcome in the second, individuals with high enough risk preference strictly prefer the former.

We know that a particular choice $C$ from a particular menu $\mathcal{M}$ must be made by a consumer with risk-preference-type that makes this choice optimal. Let $\mathcal{B}(C|\mathcal{M})$ be the set of such consumer types for which we can rationalize this plan choice, that is

$$\mathcal{B}(C|\mathcal{M}) = \left\{ (\pi, \sigma) \, | C \in \arg\max_{X \in \mathcal{M}} U(X|\pi, \sigma) \right\}.$$

**Data and Identification** Let $H(\pi, \sigma)$ denote the (cumulative) distribution of preference and risk types in the population, i.e., the fraction of people with a risk type below $\pi$ and a preference type below $\sigma$. This is the object we seek to identify. An observation in our data set is a tuple $z^i = (C^i, \mathcal{M}^i)$ where $\mathcal{M}^i$ denotes the menu of plans from which individual $i$ could choose and $C^i$ denotes his or her actual choice. We use the notation $D(\cdot)$ to refer to empirically observed fractions. $D(C, \mathcal{M})$ is the fraction of individuals observed in the population facing menu $\mathcal{M}$ and choosing contract $C$. $D(C|\mathcal{M})$ denotes the conditional distribution of choices given a particular menu. In environments where we see multiple menus, $\mathcal{M}^i \neq \mathcal{M}^j$, we assume that these are randomly distributed in the population. As discussed in the introduction, this arises naturally in field experimental settings where otherwise identical people are presented with different menus, but might also arise in other settings where regulations and therefore insurance plans might differ for otherwise identical populations.[6] We sometimes suppress the menu when there is no ambiguity. When we discuss utilizing data on cost realizations, we write $z^i = (k^i, C^i, \mathcal{M}^i)$, with slight abuse of notation, where $k^i$ denotes the cost that actually materialized for this individual. Throughout, we assume that we observe a continuum of individuals, so that the law of large number convention holds. We outline the implication of these assumptions next.

We first consider a single cross-section of observed insurance choices in which all individuals face the same menu $\mathcal{M}$. This follows most of the literature, but we focus on the case where only choices are observed. Following the law of large number convention, we say that the observed distribution $D$ *can be generated under type distribution* $H$ if the number of observed choices does not exceed the number of types that find it optimal to make such choices. Formally, we require for any subset $\mathcal{M}' \subseteq \mathcal{M}$ that

$$\int_{C \in \mathcal{M}'} D(C|\mathcal{M}) dC \leq \int_{(\pi,\sigma) \in \mathcal{U}} dH(\pi, \sigma), \tag{3}$$

where $\mathcal{U} = \cup_{C \in \mathcal{M}'} \mathcal{B}(C|\mathcal{M})$. Simply, the fraction of agents who choose an insurance plan

[6] We do not rely on multiple choice observations per individual as this seems especially unreliable in insurance settings (see also Footnote 1). Usual revealed preference approaches discussed in the introduction consider consumption situations where individuals choose more than once at different relative prices, and have not changed from one choice to the next. This is problematic in insurance settings for two reasons: first, people might update about their risk type depending on whether they had a high cost realization or not (while in most standard consumption settings there is no uncertainty) and second, there is evidence that individuals display inertia. Therefore, we only rely on variation in choices across individuals.

within subset $\mathcal{M}'$ (left-hand side) cannot be larger than the fraction of agents in the set of types that find it optimal to choose an insurance plan in $\mathcal{M}'$ (right-hand side).[7] We say that $D$ *is permissable* if it can be generated under some type distribution. Note that no data on cost realizations is used here, so our identification results from insurance choices identify *perceived* risk type, even if it differs from their true risk. When we discuss utilizing data on cost realizations, we use the assumption that perceived risks coincide with the true risks, and say that distribution $D$ is permissable if both the choices and the empirical frequency of costs can be generated by some distribution $H$, assuming that the risk type $\pi^i$ for each individual $i$ is the true risk that generates the risk realization $k^i$.[8]

We also consider the possibility that otherwise identical populations are asked to choose from different menus. In this case, we observe $z^i = \{C^i, \mathcal{M}^i\}$ for multiple contracts $\mathcal{M}_j$ that are randomly assigned to individuals. That is, we observe a set of multiple cross-sections of choices $\{D(C|\mathcal{M}_j)\}$, one for each different menu $\mathcal{M}_j$. We say that these observations can be generated under $H$ if $D(C|\mathcal{M}_j)$ can be generated under $H$ for each menu $\mathcal{M}_j$. In other words, a single type distribution rationalizes the choices across the multiple cross-sections. The overall set of observations is permissable if it can be generated by some type distribution.

As mentioned earlier, our object of interest is the population heterogeneity captured by the distribution $H$. We evaluate whether full or partial identification is possible depending on the type of data available. The ideal data allows us to point identify the entire joint distribution $H$. We introduce the following definition, where the meaning of a set of observations depends on which of the above settings we are in.

**Definition 1 (Full Identification)** *We say that there is full identification if for each permissable observed distribution $D$ there is a unique type distribution $H$ that can generate it.*

While identifying the full distribution and correlations between different dimensions of heterogeneity is clearly important, it is also challenging. A key result already would be to establish along which dimensions type heterogeneity actually matters . We are thus interested in the marginal CDFs of risk types and preference types, denoted by $H_\pi$

[7] If all agents have a unique optimal choice (i.e., $\arg\max_{X\in\mathcal{M}} U(X|\pi,\sigma)$ is a singleton for any $(\pi,\sigma)$ combination) it can be shown that condition (3) has to hold with equality everywhere by combining it with the fact that both $D$ and $H$ integrate to one. In general, agents can have multiple optimal elements in their menu, so the number of agents who find a particular combination optimal can be larger than their actual choices.

[8] Formally, we say that the conditional distribution of costs and choices $D(k, C|\mathcal{M})$ can be generated under $H$ if there exists a probability distribution $r(C|\mathcal{M}, \pi, \sigma)$ over the set of optimal choices $\arg\max_{C\in\mathcal{M}} U(C|\pi,\sigma)$ such that $D(k, C|\mathcal{M}) = \int f(k|\pi) r(C|\pi, \sigma, \mathcal{M}) dH$, where $f(k|\pi)$ is the probability of outcome $k$ for risk type $\pi$. Note that this implies that the marginal distribution of $D$ over $C$ satisfies (3) in this case, so it is permissable also when we simply disregard the risk realizations $k^i$. We make this assumption here to be able to relate our results to the earlier insurance literature. Most of our work will not use risk realizations, and can therefore back out beliefs $\pi_i$ even if they were biased relative to the true risk.

and $H_\sigma$. We explore whether we can impose any bounds on these marginal distributions and, in particular, whether we can exclude that the variation in contracts can be rationalized for some homogeneous risk-type $\pi$ or preference-type $\sigma$, with all variation arising because of heterogeneity in the other dimension. We introduce the following two definitions:

**Definition 2 (Identification of Bounds on Type Heterogeneity)** *We say that we can identify bounds on preference (respectively, risk) heterogeneity for a permissable distribution of observations $D$ if we can find $\alpha > 0$ such that any distribution $H$ that generates $D$ satisfies $H_\sigma(\bar{\sigma}) \geq \alpha$ or $1 - H_\sigma(\bar{\sigma}) \geq \alpha$ for some $\bar{\sigma} \in \mathbb{R}_+$ (respectively, $H_\pi(\bar{\pi}) \geq \alpha$ or $1 - H_\pi(\bar{\pi}) \geq \alpha$ for some $\bar{\pi} \in \Pi$).*

This definition ensures that we can place either lower or upper bounds on the CDF of preference types (resp. risk types). The lower bound $H_\sigma(\bar{\sigma}) \geq \alpha$ implies that at least a share $\alpha$ of agents have a preference type weakly smaller than $\bar{\sigma}$. The upper bound $H_\sigma(\bar{\sigma}) \leq 1 - \alpha$ implies that at least a share $\alpha$ of agents have a preference type strictly larger than $\bar{\sigma}$. The identification of bounds on the CDF in the preference dimension alone is not sufficient to reject homogeneity in that dimension. For that, we would require both an (informative) lower and upper bound on the CDF for some preference $\bar{\sigma}$. These bounds prevent us from fitting the CDF of a degenerate distribution with all mass concentrated in one specific type. The same holds of course for risks.

**Definition 3 (Inability to Reject Homogeneity)** *We say that one cannot reject preference (respectively, risk) homogeneity if for every permissable distribution of observations $D$ there exists a type distribution $H$ that can generate it and the support of its marginal distribution $H_\sigma$ (respectively, $H_\pi$) is a singleton.*

This definition refers to settings where no distribution of available data could possibly reject homogeneity in preferences (respectively risk).

# 3 Identification in a Stylized Model

We consider a stylized model in which individuals face binary risks and a binary choice. This stylized model helps us to demonstrate the potential for non-parametric identification of type heterogeneity using only choice data, but exploiting plan variation. We briefly contrast this with the identification challenges using the standard claims-based approach. In the next section, we then extend the model beyond binary risks and choice sets to settings that more closely resemble actual health insurance coverage choices to show the practical implementability of our choice-based approach. In particular, we will demonstrate how the same plan variation drives identification when plans are offered jointly as part of the same menu.

**Binary Risk and Choice Set** Any individual (ex ante) faces a binary cost distribution $k \in \{0, L\}$, either losing $L$ or nothing at all. For instance, the individual could become sick and require costly treatment, but faces no medical costs when healthy. The risk type $\pi^i$ of agent $i$ is simply his probability of incurring the cost $L$.

Any individual chooses from a binary menu $\mathcal{M} = \{X_1, X_2\}$ that offers the choice between two insurance options $X_1$ and $X_2$. We focus on the simplest case where individuals can either choose no insurance ($X_1 = \emptyset$) or some insurance ($X_2 = X$): the choice set is simply $\mathcal{M} = \{\emptyset, X\}$ for all individuals.

Since the risk is binary, a plan is fully determined by the premium $P$ and the coverage $q$ paid in case of loss, where we restrict attention to $P < q$ (as no plan with $P \geq q$ will ever be chosen). The expected utility of a plan $X = (P, q)$ simplifies to

$$U(X|\pi, \sigma) = (1 - \pi)\, u(-P|\sigma) + \pi u(-P - [L - q]\,|\sigma),$$

while remaining uninsured gives utility

$$U(\emptyset|\pi, \sigma) = (1 - \pi)\, u(0|\sigma) + \pi u(-L|\sigma).$$

An individual prefers plan $X$ over remaining uninsured if and only if

$$\frac{\pi}{1 - \pi} \frac{u(-P - [L - q]\,|\sigma) - u(-L|\sigma)}{u(0|\sigma) - u(-P|\sigma)} \geq 1. \tag{4}$$

The insurance plan entails a utility gain due to the coverage provided when the bad state realizes (with probability $\pi$), but entails a utility loss due to the premium paid, even when the good state realizes. The ratio of the utility gain relative to the utility loss is increasing in the individual's risk aversion (Pratt 1964). As a consequence, an individual's willingness to buy the plan is not only increasing in the risk type $\pi$, but also in her preference type $\sigma$.

For this binary choice environment, the main tool for analysis is the type frontier $\mathcal{T}(\emptyset, X)$ which groups together all types that are indifferent between buying the plan $X$ and remaining uninsured, i.e.,

$$\begin{aligned} \mathcal{T}(\emptyset, X) &= \{(\pi, \sigma)\,|U(X|\pi, \sigma) = U(\emptyset|\pi, \sigma)\} \\ &= \mathcal{B}(\emptyset|\mathcal{M}) \cap \mathcal{B}(X|\mathcal{M}). \end{aligned}$$

Represented in $(\pi, \sigma)$-space, the type frontier is monotonically decreasing as shown in Figure 1. A risk-neutral individual ($\sigma = 0$) is only willing to buy the plan if her loss probability exceeds the price per unit of coverage, i.e., $\pi \geq P/q$. If the loss probability converges to 0, an individual must become infinitely risk-averse to be willing to buy the insurance plan. We can then state the following result.

**Proposition 1** *Consider a binary costs $k \in \{0, L\}$ and a binary choice set $\mathcal{M} = \{\emptyset, X\}$.*

*With a single cross-section of choices and associated observations $z^i = \{C^i, \mathcal{M}\}$, we cannot put any bounds on the preference heterogeneity. We can neither reject preference homogeneity nor risk homogeneity.*

***Proof.*** *See appendix.* ■

The intuition is straightforward. The share of individuals buying insurance, $\alpha = D(X|\mathcal{M})$, corresponds to the mass of types that lie above the type frontier in the left panel of Figure 1. We cannot exclude that the variation in the choice to buy the plan is driven by heterogeneity in risk types only or by heterogeneity in preference types only. Fix the fraction $\alpha$ of individuals who buy the plan.[9] If agents have preference type $\bar{\sigma}$ but differ in risks so that exactly $1-\alpha$ of them have a type below $\bar{\pi}$, exactly $1-\alpha$ would not buy insurance which would clearly rationalize the observed choices. This case is illustrated by the dashed density above the horizontal gray line, and the shaded area indicates the mass of individuals with risk type below $\bar{\pi}$ that would not buy insurance. Alternatively we could have assumed that all agents have the same risk type $\bar{\pi}$ but are heterogeneous in preferences such that exactly $1-\alpha$ of them have types below $\bar{\sigma}$. Again, such a type distribution would rationalize the observed choices, which is indicated by the dashed-dotted density above the vertical gray line, where again the gray area indicates those types that would not buy insurance. Therefore, we can rule out neither preference nor risk heterogeneity.

Only very weak results can be obtained in this setting. Since individuals are risk-averse, only types with $\pi \leq P/q$ would be willing to buy insurance. The share of uninsured individuals $1-\alpha$ places a lower bound on the share of individuals with loss probability lower than $P/q$, i.e., $H_\pi(P/q) \geq 1-\alpha$.

While the construction suggests that one could discipline this exercise more by observing the costs that individuals end up incurring, we show next that it remains impossible to disentangle the two sources of heterogeneity. We then analyze how identification becomes possible with variation in the offered plans.

### 3.1 Standard Approach using Claims Data

We consider the case in which we observe the (ex post) realization $k \in \{0, L\}$ of the risk the individual decided to insure or not. In order to use this additional data in a meaningful way we need the assumption that individuals perceive their risk accurately. We again consider only one cross-section of individuals facing the same menu.

Inferring (ex ante) risk types from (ex post) realizations is challenging. When individuals face a binary risk and the loss value $L$ is the same for all individuals, the population distribution of cost realizations is binary as well and identified by the average loss probability. Let $\pi_a \in (0,1)$ denote the average probability of a loss in the

[9] In the right panel of Figure 1 type $(\bar{\sigma}, \bar{\pi})$ is chosen as an arbitrary point on the type frontier, implying that this type is indifferent between buying the contract or not.

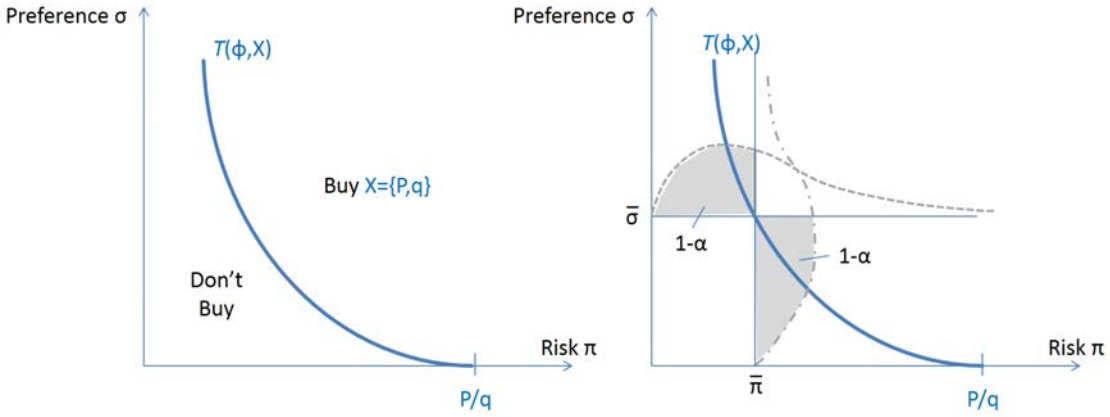


Figure 1: The left panel shows the type frontier for a binary menu $\mathcal{C} = \{\emptyset, X\}$ in $(\pi, \sigma)$-space. Types above the frontier buy the plan, while types below the fontier remain uninsured. The right panel illustrates an indifferent type $(\bar{\sigma}, \bar{\pi})$. If all other individuals have the same risk type $\bar{\pi}$ but a density of preferences as indicated by the dashed line, choices can be rationalized. Alternatively, all individuals could have same preference type $\bar{\sigma}$, but differ in risks as in the dashed-dotted density, and again choices can be rationalized.

population. Without further information on people's insurance choices, the average loss probability is not helpful in identifying risk heterogeneity. In particular, individuals could all have the same risk type (i.e., $\pi^i = \pi_a$ for all $i$), all be certain to face the loss or not (i.e., $\pi^i = 1$ for share $\pi_a$ of individuals and $\pi^i = 0$ for the remaining share $1 - \pi_a$ of individuals), or anything between as long as the average loss probability equals $\pi_a$. Thus, observing costs alone does not establish any bounds on the distribution of risks.

If both cost realizations and plan choices are observed, it is possible to bound the distribution of risks. Let $\pi_\emptyset = D(L|\emptyset)$ denote the average probability of a loss amongst individuals who do not buy insurance and let $\pi_X = D(L|X)$ denote the average probability among individuals who buy a contract. If these probabilities are not the same, the population who buys insurance faces a different risk on average than those who do not. While we can reject homogeneity in risks, we cannot bound the risk distribution much more, as we cannot identify the risk heterogeneity among those making the same choice, who again could all have the same risks (i.e., $\pi = \pi_\emptyset$ for those who don't buy insurance) or might be more heterogeneous with same average.

As long as there is adverse selection ($\pi_X \geq \pi_\emptyset$), we will not be able to rule out preference homogeneity. The reason is simple. Let $\alpha$ be the share of individuals buying the plan, and let $\sigma_X$ and $\sigma_\emptyset$ be the preference types such that a person with either type $(\pi_X, \sigma_X)$ and type $(\pi_\emptyset, \sigma_\emptyset)$ is indifferent to buying insurance. Any individual with preference type $\hat{\sigma} \in (\sigma_X, \sigma_\emptyset)$ would buy insurance when having risk type $\pi^i = \pi_X$, but not with risk type $\pi^i = \pi_\emptyset$. Hence, we can rationalize the observed choices and costs for any of these preference types simply assigning the risk type $\pi_X$ to share $\alpha$ of individuals and risk type $\pi_\emptyset$ to the remaining share. This is summarized in the

following proposition:

**Proposition 2** *Consider a binary cost $k \in \{0, L\}$ and a binary choice set $\mathcal{M} = \{\emptyset, X\}$. With a single cross-section of choices and cost realizations $z^i = \{k^i, C^i, \mathcal{M}\}$, we cannot reject preference homogeneity under adverse selection (i.e., $D(L|X) \geq D(L|\emptyset)$).*

We note that this problem of identifying (ex ante) risk types from (ex post) risk realizations, even under rational expectations, is a general one that extends beyond binary risks for any family of distribution functions that is convex in the sense that a convex combination of any two distributions is still in the family. The standard approach in the literature, however, is to rely on parametric assumptions about the type distributions (see Barseghyan et al. 2015a) and thus uses risk realizations to identify the distribution of risk types as the above property is violated under specific functional forms.[10]

### 3.2 Alternative Approach using Plan Variation: Marginal Variation

The standard approach renders it difficult to bound the distribution of preferences and to reject preference homogeneity. Therefore, we explore whether identification is possible using variation in the offered plans. The key question in this context is which types would be willing to pay for more insurance. We explore in this subsection the differential willingness to pay for a little extra coverage, and return in the next subsection to the actual choice between alternative options.

We first characterize the types of individuals attracted by more generous coverage. This can best be analyzed by considering the marginal willingness to pay for coverage or the rate at which an individual is willing to give up consumption in the "good" state for more consumption in the "bad" state. An insurance plan $X = (P, q)$ implies consumption levels in the respective states of $m_g(X) = -P$ and $m_b(X) = -P - (L - q)$ .The marginal willingness to pay (MWtP) for an extra unit of coverage equals

$$MWtP = \frac{dP}{dq}|_U = \left(-\frac{dm_g}{dm_b}|_U\right)\left(1 - \frac{dm_g}{dm_b}|_U\right)^{-1}, \tag{5}$$

which is a strictly increasing function of the marginal rate of substitution (MRS) across states:

$$\begin{aligned} -\frac{dm_g}{dm_b}|_{U(X|\pi,\sigma)} &= \frac{\pi}{1-\pi} \times \frac{u'(m_b(X)|\sigma)}{u'(m_g(X)|\sigma)} \\ &\cong \frac{\pi}{1-\pi} \times \left\{1 - \frac{u''(m_g(X)|\sigma)}{u'(m_g(X)|\sigma)}[m_g(X) - m_b(X)]\right\}, \end{aligned} \tag{6}$$

[10] For example, the convex combination of two normal distributions tends to have two peaks and is no longer normal. In this case the shape of the overall distribution of risks can identify the distribution of underlying types, but this relies very much on the choice of the underlying family of distributions.

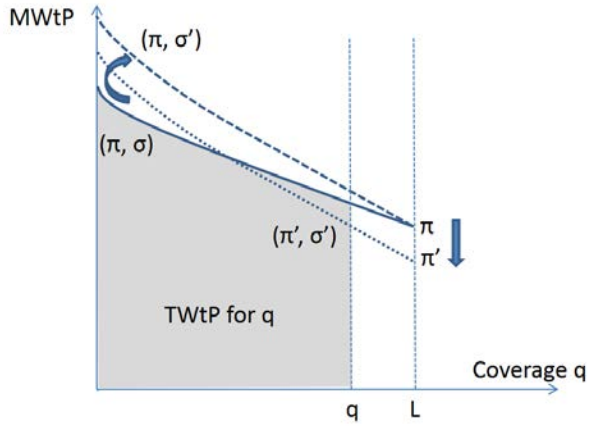


Figure 2: The solid curve depicts the marginal willingness-to-pay for coverage (MWtP) for a type $(\pi, \sigma)$ on the type frontier and the shaded area depicts this type's total willingness to pay. The dashed line shows the MWtP for a type $(\sigma', \pi)$ with higher risk aversion $(\sigma' > \sigma)$. This type has higher total willingness to pay, so is no longer on the type frontier. To be on the type frontier the risk has to be lower so that the MWtP is lowered, as indicated by the dotted line for type $(\sigma', \pi')$ with $\pi' < \pi$.

where the second line uses a Taylor approximation for $u'(m_b(X)|\sigma)$, relying on the third and higher order derivatives of the utility function being negligible. Since the MWtP is fully determined by the MRS, we will often refer to expression (6) even when we are talking about the underlying MWtP.

Expression (6) indicates that both a higher loss probability $\pi$ and higher risk aversion $-u''/u'$ increase the marginal willingness to pay for coverage. However, the relative weight of risk aversion in determining the marginal willingness to pay is smaller the more coverage the plan already provides (i.e., the smaller the consumption wedge, $m_g(X) - m_b(X)$). The marginal value of coverage thus decreases more rapidly for individuals with higher risk aversion. In the extreme case that a plan provides full insurance, the willingness to pay for the last unit of coverage equals the loss probability. The role played by the individual's risk aversion has become of second order.

It is particularly useful to compare the willingness to pay for additional coverage amongst those types who are just indifferent about buying some plan $X$ (i.e., for types $(\pi, \sigma)$ on the type frontier $\mathcal{T}(\emptyset, X)$). If we offered a plan that is more generous, but slightly more expensive, this would attract high risk individuals who are willing to pay more on the margin. By definition, the total willingness to pay for the original plan $X$ with coverage $q$ is the same for all individuals on the type frontier. Moreover, it is the sum of the marginal willingness to pay for increasing coverage from zero to $q$. Consider an individual with type $(\pi, \sigma)$ on the type frontier. His marginal willingness to pay is indicated by the solid line in Figure 2, and his total willingness to pay is indicated by the shaded area. Any other individual on the type frontier must have the same shaded area.

Now consider a more risk averse preference type $\sigma' > \sigma$. If his risk is unchanged, his marginal willingness is always higher, except at plans that offer full insurance. This is depicted as the rotation to the more tilted dashed curve in Figure 2. Since for this type the area under the curve is larger, he would strictly prefer to buy the plan rather than remaining uninsured. For an individual with preference type $\sigma'$ to be indifferent, his risk must be lower. This shift to a lower risk $\pi' < \pi$ reduces the marginal willingness to pay proportionally for any coverage level, indicated by the downward shift to the dotted curve in Figure 2. An initial rotation followed by a downward shift has the property that at coverage level $q$, the dotted curve is below the initial curve. The marginal willingness to pay for any additional coverage beyond what is provided in contract $X$ is lower for the individual with lower risk (and higher risk aversion). This will induce the following property amongst individuals that are indifferent between no insurance and contract $X$, which we will exploit later on: if contract $X$ is replaced by a more generous but more expensive contract $X'$, there will be a cutoff and those with higher risks will strictly prefer to buy the new plan while the others strictly prefer not to (compared to either the old plan or no plan, amongst which they are indifferent).

The previous argument relied on an initial "rotation" of the marginal willingness-to-pay, as risk preferences matter less closer to full insurance. It was followed by a "shift" since risk types matter everywhere, even at full insurance. We focus here on utility functions where the above logic does hold not just close to full insurance, but for any coverage level:

**Assumption 1** *We consider families of utility functions with the following property: Along the type frontier $\mathcal{T}(\emptyset, X)$ the marginal rate of substitution $\frac{\pi}{1-\pi}\frac{u'(m_b(X)|\sigma)}{u'(m_g(X)|\sigma)}$ is increasing in $\pi$, and it converges to zero as $\pi$ goes to zero.*

Since CARA preferences are typically adopted in the empirical insurance literature, we explicitly check this property for this class of preferences:

**Lemma 1** *Assumption 1 holds for CARA preferences.*
***Proof.*** *See appendix.* ■

Under Assumption 1, we can study exactly which types of plan variation lead to crossings of type frontiers which are required to identify bounds on the marginal distributions as we will show shortly. For preferences not satisfying Assumption 1, we would need to study possibly different variation to obtain crossings and thus identification.

### 3.3 Between-Menu Variation in Plans

We now introduce discrete variation in the plans offered. We consider two plans $X_h$ and $X_l$, where plan $X_h$ provides more coverage than plan $X_l$ (i.e., $q_h > q_l$). We continue to analyze binary menus $\mathcal{M}_j = \{\emptyset, X_j\}$, with different plans offered separately to

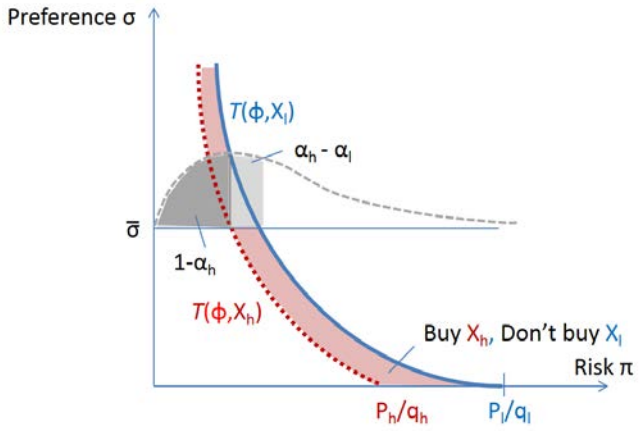


Figure 3: The solid and dotted line show the type frontiers in $(\pi, \sigma)$-space for the binary menu $\mathcal{C} = \{\emptyset, X_h\}$ and $\mathcal{C} = \{\emptyset, X_l\}$ respectively. The type frontiers do not intersect as the high-coverage plan charges the same (or a lower) premium and attracts all types that would also buy the low-coverage plan.

different cross-sections of individuals. The next section shows that the same logic drives identification when the different plans are offered jointly to a single cross-section of individuals.

Consider two randomly selected cross-sections of individuals, where the first cross-section is offered the menu $\mathcal{M}_h = \{\emptyset, X_h\}$ and the second cross-section is offered the menu $\mathcal{M}_l = \{\emptyset, X_l\}$. The share of individuals buying insurance when each plan is offered separately equals $\alpha_h = D(X_h|\mathcal{M}_h)$ and $\alpha_l = D(X_l|\mathcal{M}_l)$ respectively.

If the high-coverage plan charges the same (or a lower) premium, it dominates the low-coverage plan. All types who would buy insurance when offered the low-coverage plan also buy insurance when offered the high-coverage plan (i.e., $\mathcal{B}(X_l|\mathcal{M}_l) \subset \mathcal{B}(X_h|\mathcal{M}_h)$). The high-coverage type frontier $\mathcal{T}(\emptyset, X_h)$ is illustrated by the solid line in Figure 3. The type frontier lies below the low-coverage type frontier $\mathcal{T}(\emptyset, X_l)$ which is illustrated by the dotted line. We can assign the observed increase in shares $\alpha_h - \alpha_l$ to the types in between the two frontiers $\mathcal{T}(\emptyset, X_l)$ and $\mathcal{T}(\emptyset, X_h)$ (i.e., to $\mathcal{B}(X_h|\mathcal{M}_h) \setminus \mathcal{B}(X_l|\mathcal{M}_l)$). However, this type of plan variation sheds limited light on the heterogeneity in either dimension separately. The observed variation in plan choices could either be explained by risk variation or by preference variation only. The former is illustrated by the horizontal line in Figure 3, on which all types share the same preference $\bar{\sigma}$. In fact, the observed plan shares still do not allow to put any bounds on the preference distribution.

If the high-coverage plan $X_h$ is offered at a higher premium, it becomes less attractive than the low-coverage plan to some individuals, but remains more attractive to others if the premium increase is relatively small (i.e., $\mathcal{B}(X_{j'}|\mathcal{M}_{j'}) \nsubseteq \mathcal{B}(X_j|\mathcal{M}_j)$ for $j' \neq j$). Assumption 1 implies that among those types that are indifferent at $X$, those with high risks prefer to buy more coverage. This implies that that the type frontiers

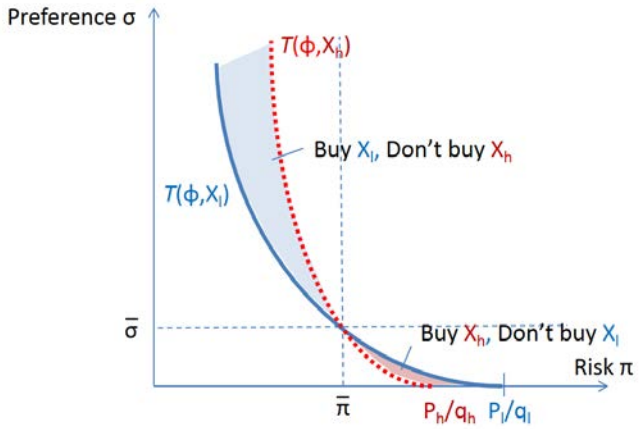


Figure 4: The solid and dotted lines show the type frontiers in $(\pi, \sigma)$-space for the binary menu $\mathcal{C} = \{\emptyset, X_l\}$ and $\mathcal{C} = \{\emptyset, X_h\}$ respectively. The type frontiers intersect at $(\bar{\pi}, \bar{\sigma})$. The cheaper low-coverage plan charges a higher price per unit of coverage and differentially attracts types with high risk aversion and low risk.

cross only once, as depicted in Figure 4. The high-coverage type frontier $\mathcal{T}(\emptyset, X_h)$, depicted by the dotted curve, is a clockwise "rotation" around $(\bar{\pi}, \bar{\sigma})$ relative to the low-coverage type frontier $\mathcal{T}(\emptyset, X_l)$, depicted by the solid curve. Low risk types between the two curves (with $\pi < \bar{\pi}$ and $\sigma > \bar{\sigma}$) buy the cheaper low-coverage plan but would remain uninsured when offered the more expensive plan, while high risk types between the two curves (with $\pi > \bar{\pi}$ and $\sigma < \bar{\sigma}$) remain uninsured when offered the cheaper low-coverage plan, but buy insurance when the plan provides the additional coverage so long as the premium increase is not too high. Note that the risk-neutral individual on the type frontier of plan $X_j$ has risk type $\pi = P_j/q_j$. Only if its price per unit of coverage remains lower than for the low-coverage contract ($P_h/q_h < P_l/q_l$), the high-coverage contract can differentially attract some types to buy insurance.

Clearly, we could now set identify an individual's type if we were to observe the individual's choice under the two menus. For example, an individual who switches out of the insurance plan when offered $X_h$ rather than $X_l$, must have a risk type higher than $\bar{\sigma}$ and a preference type lower than $\bar{\pi}$. Even if we observe only one choice per individual, but two random cross-sections of individuals choosing from the respective menus, we can put bounds on the mass of the distribution $H$ relative to the intersection point as stated in the following Lemma.

**Lemma 2** *Under Assumption 1, the type frontiers for the pairwise menus* $\{\emptyset, X_h\}$ *and* $\{\emptyset, X_l\}$ *with* $q_h > q_l$*, have a unique intersection* $(\bar{\pi}, \bar{\sigma})$ *if and only if* $P_h > P_l$*, but* $P_l/q_l \geq P_h/q_h$*. Moreover,*

$$\int_{\pi \geq \bar{\pi}} \int_{\sigma \leq \bar{\sigma}} dH \geq \alpha_h - \alpha_l \geq - \int_{\pi \leq \bar{\pi}} \int_{\sigma \geq \bar{\sigma}} dH. \tag{7}$$

***Proof.*** *See appendix.* ■

Very low risk types along the type frontier for contract $X_l$ have near zero marginal willingness to pay for insurance, so they will not buy the additional insurance offered by $X_h$. This ensures that the dotted curve in Figure 4 lies to the right of the solid curve at low risks. Moreover, by Assumption 1, the willingness to pay changes monotonically along the type frontier, so there can only be a unique type where the type frontiers cross: at that point all lower risks on the type frontier for contract $X_l$ would buy the additional insurance while all higher risks would not. If $P_l/q_l > P_h/q_h$, for risk-neutral preference ($\sigma = 0$) the dotted curve has to be to the left of the solid one, and so there will be a crossing, as shown in Figure 4. On the other hand, if the expensive insurance plan offers less coverage per dollar (i.e., $P_h/q_h > P_l/q_l$), the dotted curve would lie completely to the right of the solid curve.[11] In this case the type frontiers no longer intersect as the low-coverage contract dominates the high-coverage contract, and plan variation does not allow us to put bounds on preferences by a similar logic as that depicted in Figure 3.[12]

The Lemma clearly describes the plan variation required for the type frontiers to intersect: the high-coverage plan needs to be more expensive, but provide coverage at a lower price per unit. If more people buy the high-coverage plan, the difference in plan shares $\alpha_h - \alpha_l$ places a lower bound on the share of individuals with $\pi > \bar{\pi}$ and $\sigma < \bar{\sigma}$. The additional coverage is relatively more attractive to individuals with higher risk than to individuals with higher risk aversion. If more people by the low-coverage plan, the difference $\alpha_l - \alpha_h$ imposes a lower bound on the share of individuals with $\pi < \bar{\pi}$ and $\sigma > \bar{\sigma}$. The exact shape of the type frontiers could help putting tighter bounds on the joint distribution, but the more important observation is that the intersection of the frontiers allows to place bounds on the marginal distributions as well. For example, if the high-coverage plan is more popular, the share differential places a lower bound on the share of individuals with lower risk aversion, i.e., $H_\sigma(\bar{\sigma}) > \alpha_h - \alpha_l$. This is in contrast to the case discussed before where plan variation induced a shift in the type frontier (Figure 3) rather than a rotation (as in Figure 4). Intersections of the type frontiers are crucial for identification and more intersections help us to further tighten the bounds on the marginal distributions.

**Proposition 3** *Consider a binary costs $k \in \{0, L\}$ and observations $z^i = \left(C^i, \mathcal{M}^i\right)$ with $\mathcal{M}^i \in \{\mathcal{M}_1, .., \mathcal{M}_J\}$. With at least two appropriately chosen menus $\mathcal{M}_1$ and $\mathcal{M}_2$ offered to random cross-sections of individuals, there are permissable distributions for which we can identify bounds on preference and risk heterogeneity. With at least three appropriately chosen menus $\mathcal{M}_1, \mathcal{M}_2$ and $\mathcal{M}_3$, there are permissable distributions for which we can reject preference and risk homogeneity.*

[11] This is true since it lies to the right for both low and for high risks $\pi$ (and can only cross once).
[12] Only that here the labels between $X_h$ and $X_l$ are reversed.

***Proof.*** *See appendix.* ■

This proposition follows relatively straightforwardly from Lemma 2. Consider two menus $\{\emptyset, X_l\}$ and $\{\emptyset, X_h\}$ which generate type frontiers as depicted in Figure 4, with crossing point $(\bar{\pi}, \bar{\sigma})$. Assume an underlying distribution of types such that more agents choose the low-coverage contract than the high-coverage contract. That means that there are more types in the shaded area above $\bar{\sigma}$ (and below $\bar{\pi}$) than in the shaded area below. This puts a lower bound on the number of agents with preference types above $\bar{\sigma}$ (and below $\bar{\pi}$), but does not yet rule out that all agents have the same preference or risk type. Consider Figure 4 with another contract $X_h'$ providing even higher coverage than $X_h$ and the corresponding type frontier crossing the type frontier of the high-coverage contract to the south-east of the previous intersection ($\bar{\pi}' > \bar{\pi}, \bar{\sigma}' < \bar{\sigma}$). If more agents buy insurance when offered this new generous contract than when offered the original high-coverage contract, we know that there exists a set of types in the underlying distribution that have preference types below $\bar{\sigma}'$ (and above $\bar{\pi}'$). This places bounds on heterogeneity, since we can be sure that there are agents both with preferences above $\bar{\sigma}$ and below $\bar{\sigma}'$. The same holds for risks.

Proposition 3 suggests that more variation in insurance plans will further tighten bounds as the observation of each additional plan may provide an additional crossing relative to other plans. While it may not seem straightforward how this could lead to full identification, the textbook insurance model provides a simple setting that allows us to do exactly that.

### 3.4 Full Identification in the Textbook Model

The previous subsection demonstrated how plan variation can place non-parametric bounds on the distribution of preferences and risks. This subsection turns to the question whether variation in binary menus $\mathcal{M} = \{X_1, X_2\}$ across otherwise identical populations is enough to fully identify the joint distribution $H$ in principle. This question relates directly to the identification of risks and preferences in the standard textbook model of insurance choice, which has guided the thinking about insurance markets for decades.

In our binary setting, recall that a plan $X_n$ is fully characterized by the premium $P_n$ and the amount of insurance $q_n$. Defining the unit price of insurance as $p_n = P_n/q_n$, one can equivalently characterize the plan by $(p_n, q_n)$. The question is whether enough variation in these two components identifies the underlying heterogeneity. This analysis can be split into two parts. First, one can consider plans with identical unit price $p_n = \hat{p}$ and determine the fraction of agents that choose plan $(\hat{p}, \hat{q})$ over any other plan $(\hat{p}, q)$ through pairwise comparisons. Alternatively, one can ask individuals to directly choose their preferred plan amongst all plans $(\hat{p}, q)$ with unit price $\hat{p}$. This alternative formulation entails less information, so identification here also implies identification under

pairwise comparisons.[13] The alternative formulation is exactly the setup in textbook insurance models where individuals choose the optimal quantity of insurance at given unit price to cover a binary risk (see for example Kreps 1990, Varian 1992, Mas-Colell, Whinston and Greene 1995, Gravelle and Rees 2004). In our notation, this corresponds to the selection of an insurance plan from a menu $\mathcal{M}_{\hat{p}} = \{(P,q)\,|P/q = \hat{p}, q \in \mathbb{R}_+\}$, and from choice data we can observe the fraction of agents $D(q|\mathcal{M}_{\hat{p}})$ buying an unrestricted coverage level $q \in \mathbb{R}_+$ offered at unit price $\hat{p}$, as well as the cumulative $\tilde{D}(q|\mathcal{M}_{\hat{p}})$ of agents that choose a coverage level no larger than $q$. For notational convenience and to highlight the connection to standard results, we continue with this textbook model, instead of pairwise plan comparisons.[14]

This leads to the second step for identification: we also need variation in unit prices. Observing the fraction of individuals choosing between different coverage levels at constant unit price is not informative about risk or preference: following the logic of Lemma 2, the type sets $\mathcal{B}(q|\mathcal{M}_{\hat{p}})$ for any available coverage choice $q$ do not intersect as the price per unit of coverage remains constant, and we are in a choice environment akin to those depicted in Figure 3. However, consider randomly assigning groups to different unit prices. That is, for a first random cross-section we observe their insurance choices from $\mathcal{M}_{p_h}$ and for a second cross-section we observe choices from $\mathcal{M}_{p_l}$. Consumers with the same coverage choice for the price $p_h$ may choose different coverage levels at the reduced price $p_l < p_h$. The difference in willingness to buy additional coverage as prices change depends on the difference in their preferences and risks. In particular, due to the decreasing returns to coverage, the type with higher risk aversion (but lower risk) will increase her coverage less when the price decreases to $p_l$. This implies that the type sets $\mathcal{B}(q|\mathcal{M}_{p_h})$ will be flatter than the type sets $\mathcal{B}(q|\mathcal{M}_{p_l})$ at their respective intersections and allows us to use the difference in coverage shares to disentangle the heterogeneity in risk and preferences.

The textbook model allows for a direct illustration of this intuition. An individual

[13] Intuitively, for a given agent, pairwise comparisons provide strictly more information, since it provides pairwise information even for choices that are not optimal to this particular agent. This intuition does not simply generalize for our comparison: in the textbook model one observes the optimal choice among many contracts for any given individual, while in the binary comparisions one does not see the preferred choice for one particular individual but only the relative attractiveness overall accross individuals. Nevertheless, note that in the textbook model, for a given agent the optimal choice $q^*$ is unique as his utility is strictly concave in $q$. Consider now an agent who has to choose between two options $q'$ and $q''$ who are either both larger or both smaller then his optimal $q^*$. Because of concavity he prefers the choice that is closest to his optimal choice. Now consider a binary choice set $\mathcal{M} = \{X_q, X_{q+\varepsilon}\}$ where both options have same unit price $p$ but $X_q$ has quantity $q$ while $X_{q+\varepsilon}$ has quantity $q+\varepsilon$. By the preceeding argument, all agents whose unconstraint choice $q^*$ is below $q$ prefer $X^q$, while those whose unconstrained choice is above $q+\varepsilon$ prefer $X_{q+\varepsilon}$. For $\varepsilon$ sufficiently small, the mass of agents that prefer the middle vanishes, and we have uncovered the fraction of agents that have optimal choices below $q$ as those that choose $X_q$. Formally, considering a sequence of populations we have $\lim_{\varepsilon\to 0} D\left(X_{q+\varepsilon}|\{X_q, X_{q+\varepsilon}\}\right) = \int^q D(x|\mathcal{M}_{\hat{p}})dx$. So pairwise comparisions entail at least the information from the textbook model.

[14] While, as mentioned before, pairwise plan comparisons provide in this setting at least the same information as what the textbook model provides, this is not generally true. We will discuss this further in Section 4.

chooses the level of coverage such that the marginal rate of substitution for her type equals the rate at which transfers can be made between the good and the bad state (as implied by the unit price),

$$\frac{\pi}{1-\pi}\frac{u'\left(m_b\left(q\right)|\sigma\right)}{u'\left(m_g\left(q\right)|\sigma\right)}=\frac{p}{1-p}. \tag{8}$$

An individual buys more coverage than another because she faces a higher risk or because she is more risk-averse. The variation in coverage choices across individuals at a constant price $p$ could therefore be entirely driven by heterogeneity in preferences or heterogeneity in risks alone. Now taking logs on both sides of equation (8) and approximating $\log\left[u'\left(m_b|\sigma\right)/u'\left(m_g|\sigma\right)\right]\cong-\frac{u''(m_g|\sigma)}{u'(m_g|\sigma)}\left[m_g-m_b\right]$, we find an individual's demand for coverage as a function of the unit price,

$$q\cong A+B\log\left(p/\left[1-p\right]\right) \tag{9}$$

with

$$A=L-\frac{\log\left(\frac{\pi}{1-\pi}\right)}{u''\left(m_g|\sigma\right)/u'\left(m_g|\sigma\right)}\text{ and }B=\frac{1}{u''\left(m_g|\sigma\right)/u'\left(m_g|\sigma\right)}. \tag{10}$$

While both higher risk and higher risk aversion increases coverage choices, the response to a change in the price only depends on risk aversion. Those with higher risk aversion tend to increase their coverage less and are thus less responsive to a change in the price.

The above approximation is exact for CARA preferences. For such preferences there is a one-to-one mapping between $(A,B)$ and $(\pi,\sigma)$, since $A=L+\log\left(\pi/(1-\pi)\right)/\sigma$ and $B=-1/\sigma$. Therefore, the distribution $H$ can be identified from the distribution of $A$ and $B$ in the population. We will show that sufficient price variation allows for such identification. The key step in this argument is to observe that prices determine the share of people with $(A,B)$ for whom $\alpha A+\beta B\leq t$ along any ray defined by $\alpha$ and $\beta$ and for any parameter $t$. In particular,

$$\Pr(\alpha A+\beta B\leq t)=\Pr\left(A+\frac{\beta}{\alpha}B\leq\frac{t}{\alpha}\right)=D\left(\frac{t}{\alpha}|M_{p(\alpha,\beta)}\right), \tag{11}$$

where $D\left(\frac{t}{\alpha}|M_{p(\alpha,\beta)}\right)$ is the observed share of people that buy no more insurance than $q=t/\alpha$ for for $p(\alpha,\beta)\equiv\exp\left(-\beta/\alpha\right)/[1+\exp\left(-\beta/\alpha\right)]$. With sufficient price variation this can be observed for any level of $\alpha,\beta$ and $t$. This amounts to observing the marginal distribution (11) of the weighted sum of $A$ and $B$, for all possible weights.

The remaining question is whether we can learn the joint distribution over $A$ and $B$ from observing all such marginal distributions over the sums of $A$ and $B$. Cai, Zhang and Peng (2005) provide an affirmative answer based on a proof in the space of characteristic functions which we replicate in the appendix to make our arguments self-contained. This yields the following insight:

**Proposition 4** *Consider a binary risk $k \in \{0, L\}$, a choice set $\mathcal{M}_p$ with constant unit price and CARA preferences. When observing the distribution of coverage choices in $\mathcal{M}_p$ for each price $p \in [0, 1]$, we can fully identify the distribution $H$ for any permissable set of observations.*

***Proof.*** *See appendix.* ■

Full identification of the non-parametric type distribution requires infinite variation in the unit price of insurance. However, with limited price variation, we can still reject homogeneity in preferences and risks and uncover key moments of the respective distributions. This can be seen along the following lines. The demand specification in (9) for CARA preferences implies

$$Var\left(q|p\right) = Var\left(A\right) + Var\left(\sigma^{-1}\right) \times \tilde{p}^2 - 2Cov\left(A, \sigma^{-1}\right)\tilde{p} \tag{12}$$

for $\tilde{p} = \log\left(p/\left[1-p\right]\right)$. Hence, with two exogenous prices, we obtain

$$\left[Var\left(q|p_1\right) - Var\left(q|p_2\right)\right] / \left[\tilde{p}_1 - \tilde{p}_2\right] = Var\left(\sigma^{-1}\right) \times \left[\tilde{p}_1 + \tilde{p}_2\right] - 2Cov\left(A, \sigma^{-1}\right). \tag{13}$$

We can use this to test for homogeneity in preferences and risks. This is easy to see for the preference types - the argument for the risk types is relegated to the appendix. Whenever the difference in variances in equation (13) is different from 0, we can reject that $\sigma$ (or $\sigma^{-1}$) is constant and thus that the preference type is homogeneous. For a constant $\sigma$, both $Var\left(\sigma^{-1}\right)$ and $Cov\left(A, \sigma^{-1}\right)$ would be equal to 0. Finally, when observing three (exogenous) prices, we can also identify the variance of the inverse of the coefficient of absolute risk aversion

$$Var\left(\sigma^{-1}\right) = \left[\frac{Var(q|p_1) - Var(q|p_2)}{\tilde{p}_1 - \tilde{p}_2} - \frac{Var(q|p_2) - Var(q|p_3)}{\tilde{p}_2 - \tilde{p}_3}\right] / \left[\tilde{p}_1 - \tilde{p}_3\right].$$

To summarize:

**Proposition 5** *Consider a binary risk $k \in \{0, L\}$, choice sets $\mathcal{M}_p$ with constant unit price and CARA preferences. Rejecting homogeneity in preferences (risks) is possible when observing the distribution of coverage choices in $\mathcal{M}_p$ for two prices in the unit interval. We can identify the variance in (inverse) preference types when observing the distribution of coverage choices in $\mathcal{M}_p$ for three prices.*

***Proof.*** *See appendix.* ■

In this textbook model the progress we can make with claims data instead of price variation is limited, for reasons akin to those mentioned in Section 3.1. Note that when we restrict attention to CARA preferences like in Proposition 4, claims data would be sufficient to reject homogeneity in preferences, but does not allow to identify any additional moments capturing the variation in preferences or risks. The issue is that we cannot establish or reject homogeneity in preference or risk types for the

individuals choosing the same coverage level $q$ at unit price $p$. The observed share of losses $D(L|q,p)$ pins down only the average risk type among these individuals and a preference type that rationalizes the coverage choice given this average risk type. Hence, there is no way to identify heterogeneity in preferences or risks beyond these average types that rationalize the respective coverage choices.[15]

Finally, note that in lieu of simple binary menus we introduced the text-book approach with several plans offered jointly. However, we restricted the menu to offer plans with constant unit price, providing limited opportunities for identification within a single menu. In the next section we discuss how multiple options offered together in a single menu can help uncover the underlying heterogeneity.

# 4 From Theory to Practice

In this section, we do two things to show how to implement our identification approach in practice. First, we move beyond binary risks and simple insurance plans. In practice, costs can take many values and insurance plans are often complex. The coverage that an insurance plan provides depends on the size and type of costs. For example, health plans typically specify deductibles, co-insurance rates and out-of-pocket maxima. Variation in these plan features helps identify type heterogeneity when (like in the binary case) these features differentially attract types along different dimensions. The increase in the dimensionality of the contract space - keeping fixed the dimensionality of the type space - provides additional opportunities for identification.

Second, we show how *within-menu* plan variation can be used for identification even if there is no *between-menu* plan variation (obtained via random variation in menus faced by similar individuals). Even choices from a single menu can be informative enough to place bounds on the distribution of types. This approach is particularly useful, as within-menu plan variation naturally arises in many settings, while between-menu variation typically requires experiments or quasi-experimental variation.

## 4.1 Plans and Expenses in Practice

When costs are continuous, a plan $X$ can in principle specify any out-of-pocket expense $x(k)$ for each possible cost $k \in \mathbb{R}_+$. We focus on three pre-dominant coverage features of insurance plans: a deductible $D$, below which all costs are paid out-of-pocket by the individual, an out-of-pocket maximum $M$ above which the out-of-pocket expenses cannot increase, and a co-insurance rate $\beta$ determining the individual's cost share in

[15] Moreover, when we observe a higher loss probability among the individuals who buy more coverage, in the spirit of Proposition 2, and we were to allow for any risk-averse preference type represented by some concave utility function $u(.)$, we can always construct a homogeneous preference type ex post that rationalizes the choice and claims data with only heterogeneity in risk types.

between. The out-of-pocket expense equals

$$x(k) = \begin{cases} k & \text{for} \quad k \leq D, \\ D + \beta(k - D) & \text{for} \quad k \in \left[D, \frac{1}{\beta}M - \frac{1-\beta}{\beta}D\right], \\ M & \text{for} \quad k > \frac{1}{\beta}M - \frac{1-\beta}{\beta}D. \end{cases}$$

**Simple Plans Covering High Expenses** We first show that the logic for identification remains essentially identical to the arguments from the previous sections if contracts cover high but not low expenses: consider insurance plans that set the deductible equal to the out-of-pocket maximum (i.e., $Z \equiv D = M$). This induces full cost sharing below $Z$ but no cost sharing above $Z$. Now, the setting resembles our stylized setting with binary risks studied before. The valuation of the insurance plan depends crucially on the probability $1 - F(Z|\pi)$ that the coverage is received.

Both high risk aversion and high expected costs increase the willingness to pay for such a plan. We can compute the marginal willingness to reduce the threshold $Z$ when the plan charges a premium $P$, analogous to (5). Alternatively we can invert this to get an expression analogous to the marginal rate of substitution (6) that guided our understanding in the binary risk case:

$$\begin{aligned} \frac{\frac{dP}{dZ}|_{U(X|\pi,\sigma)}}{1 - \frac{dP}{dZ}|_{U(X|\pi,\sigma)}} &= -\frac{[1 - F(Z|\pi)]\, u'(-P - Z|\sigma)}{\int_0^Z u'(-P - k|\sigma)\, f(k|\pi)dk} \\ &= -\frac{1 - F(Z|\pi)}{F(Z|\pi)} \frac{u'(-P - Z|\sigma)}{E[u'(-P - k|\sigma)\,|k \leq Z; \pi]}. \end{aligned} \tag{14}$$

The basic structure of this expression is very similar to (6) in the binary case. If risk types can be ranked in a first-order stochastic dominant way (i.e., $F(k|\pi_i) \leq F(k|\pi_j)$ for all $k$), we have that both individuals with higher risk and with higher risk aversion have a higher willingness-to-pay for additional coverage. However, the returns to coverage tend to decrease more rapidly for individuals with higher risk aversion. If among the marginal buyers of a plan, the marginal willingness to pay is indeed higher for those with higher risk but lower risk aversion, we can again invoke Lemma 2 and establish rotations of the type frontiers by changing the coverage and price paid.[16] Sufficient variation in prices and coverage allows us to uncover the underlying heterogeneity in the spirit of Proposition 4.

**Plans Covering High vs. Low Expenses** In practice, plans also differ in the type of expenses they cover: a plan could have lower deductible, but a higher out-of-pocket maximum, as well as different coinsurance rates. We will see that this can offer an additional channel for identification.

[16] Note that a risk-neutral type is indifferent about buying when $(1 - F(Z|\pi))\, E(k - Z|k > Z, \pi) = P$. By analogy to the binary case, to obtain a crossing of the type frontiers, we would need the expected coverage to increase by more than the price for this indifferent risk-neutral type.

The marginal expected utility from lowering the out-of-pocket expense $x(k)$ for a given cost $k$ equals

$$dU(X|\pi,\sigma) = f(k|\pi)\, u'(x(k)|\sigma)\, dx.$$

The willingness to purchase additional coverage depends on the probability of the underlying cost (which is determined by the risk type $\pi$) and the utility from reducing the out-of-pocket expense (which is determined by the risk preference $\sigma$).

Arbitrary non-linear insurance plans could vary the out-of-pocket expenses for each cost realization $k$. Such plan variation allows separating heterogeneity in risk and preferences. Yet even standard insurance contracts provide valuable identification. Out-of-pocket maxima, for example, affect the coverage for high expenses, while deductibles affect coverage for low expenses. For given risks, individuals with high risk aversion care more about reducing high out-of-pocket expenses than reducing low out-of-pocket expenses. In particular, a type with extreme risk aversion chooses based on the out-of-pocket maximum and premium only, trying to reduce spending in the worst case, in which both are paid. As a result, decreasing the wedge between out-of-pocket maximum and deductible attracts the more risk-averse and discourages the less risk-averse types from buying insurance. This tends to rotate the decreasing type frontier counter-clockwise.

How much individuals with different risk care about reducing the out-of-pocket maximum rather than the deductible depends on the likelihood ratio of the different expenses. Starting from a contract for which deductible and out-of-pocket maximum coincide at $Z$, the marginal willingness to reduce the deductible relative to the out-of-pocket maximum simplifies to the product of co-insurance and hazard rate:

$$\frac{dM}{dD}|_{U(X|\pi,\sigma)} = (1-\beta)\frac{f(Z|\pi)}{1-F(Z|\pi)}. \tag{15}$$

If the hazard rate were to decrease for higher risk types, they care more about reducing the out-of-pocket maximum.[17] Decreasing the wedge between the out-of-pocket maximum and deductible then tends to rotate type frontiers clockwise. Whether we can establish a clean identification result like Lemma 2 will depend on these potentially opposing forces.

### 4.2 Within-Menu Plan Variation

In practice, we often observe individuals picking a plan out of a menu providing the choice between several, different plans. We demonstrate how *within-menu* variation in plans can still be exploited for identification and link this to the *between-menu* variation

[17] Note that when risk types are ranked by first-order stochastic dominance, this does not imply that the hazard rate and thus the marginal rate of substitution between $D$ and $M$ is monotone. A monotone likelihood ratio property for the risk types (i.e., $f(k+\varepsilon|\pi)/f(k|\pi)$ increasing in $\pi$ for $\varepsilon > 0$), however, would imply both a first-order stochastic dominance ranking and a monotone hazard rates.

in plans analyzed before.

The first practical insight is that if identification is not possible for plans offered in different menus (i.e. from between-menu variation, as in our previous setting), identification is not possible either when these plans are offered together (i.e. from within-menu variation). This is the case when type frontiers do not intersect, as in Figure 3. Consider again our original binary risk setting, but now with contracts $X_h$ and $X_l$ offered together in a three-plan menu $\mathcal{M} = \{\emptyset, X_l, X_h\}$. If contract $X_h$ provides more insurance at higher unit price (such that $\mathcal{T}(\emptyset, X_h)$ lies above $\mathcal{T}(\emptyset, X_l)$), identification is not possible using choices from this menu, as any different choice can be explained either by higher risk aversion or higher risk. Starting from a type that buys no insurance, an agent switches first to the low-coverage plan $X_l$, when increasing either her risk or preference type, and eventually to the high-coverage plan $X_h$.[18]

The counterpart of this result is that plan variation that leads to identification across menus can also provide identification when plan are offered together in one menu. Consider any two plans $X_j$ and $X_{j'}$ for which the type frontiers $\mathcal{T}(\emptyset, X_j)$ and $\mathcal{T}(\emptyset, X_{j'})$ intersect, as illustrated before in Figure 4. The type $(\bar{\pi}, \bar{\sigma})$ at the intersection of the two frontiers is indifferent between all three options (including the outside option $\emptyset$). This type $(\bar{\pi}, \bar{\sigma})$ is a natural candidate to provide a bound on the support of one of the two plans.

We illustrate this in our original binary risk setting. Consider again contracts $X_h$ and $X_l$, but with $X_h$ providing more coverage at lower price per unit. Figure 5 plots the different type sets corresponding to the choice of each of the plans when the plans are offered within the same menu $\mathcal{M} = \{\emptyset, X_l, X_h\}$. The low-coverage plan provides an intermediate option, but as it charges a higher price per unit of coverage, this is only attractive to individuals with relatively high risk aversion (and relatively low risk type). Such individuals strongly value the basic coverage provided by the low-coverage plan, but place less value on the additional coverage provided by the high-coverage plan. Hence, when increasing the risk type of an individual with risk aversion higher than $\bar{\sigma}$, she will first switch from no insurance to the low-coverage plan before eventually switching to the high-coverage plan. In contrast, individuals with risk aversion lower than $\bar{\sigma}$ will never buy the low-coverage plan. Their marginal valuation of coverage is more constant. As a consequence, these individuals remain uninsured when their risk type is low, but switch immediately to the high-coverage plan (charging a low price per unit) when their risk type is high.

In Figure 5 this gives rise to an area above $\bar{\sigma}$ where agents buy $X_l$, but not below. As a consequence, the share of individuals buying the low-coverage plan $X_l$ places a

[18] Note that the agent will switch to the low-coverage plan when crossing the type frontier $\mathcal{T}(\emptyset, X_l)$ and to the high-coverage plan when crossing type frontier $\mathcal{T}(X_l, X_h)$. The latter lies above $\mathcal{T}(\emptyset, X_h)$ in this case and does intersect with $\mathcal{T}(\emptyset, X_l)$ either. The share of individuals buying the low-coverage plan $X_l$ allows us to allocate mass to the parameter region in between the loci $\mathcal{T}(\emptyset, X_l)$ and $\mathcal{T}(X_l, X_h)$, but this variation is not sufficient to place any bounds on the heterogeneity in preferences.

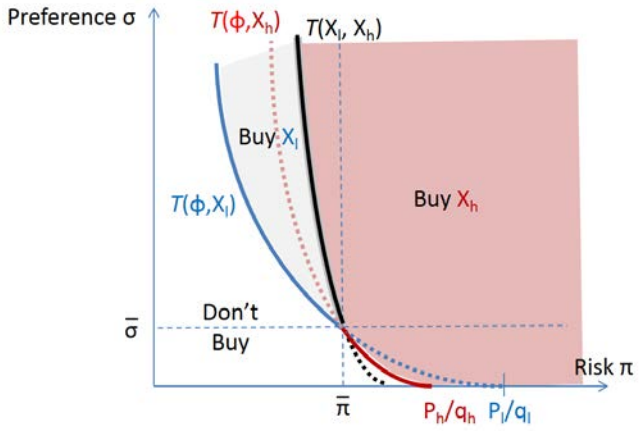


Figure 5: The figure shows the choices for types in $(\pi, \sigma)$-space from the menu $\mathcal{C} = \{\emptyset, X_l, X_h\}$. The lines show the type frontiers for any binary choice. All type frontiers intersect at $(\bar{\pi}, \bar{\sigma})$. Like in Figure 4, the low-coverage plan charges a higher price per unit of coverage and therefore differentially attracts types with high risk aversion (and low risk).

lower bound on $1 - H_\sigma(\bar{\sigma})$. The following Lemma summarizes identification using within-menu variation, in line with the potential of between-menu variation described in Lemma 2:

**Lemma 3** *Under Assumption 1, the three type sets rationalizing the respective plan choices from the menu $\mathcal{M} = \{\emptyset, X_l, X_h\}$ with $q_h > q_l$ meet at a unique pair $(\bar{\pi}, \bar{\sigma})$ if and only if $P_h > P_L$ and $P_l/q_l > P_h/q_h$. Moreover,*

$$\int_{\pi \leq \bar{\pi}} \int_{\sigma \geq \bar{\sigma}} dH \geq D(X_l|\mathcal{M}).$$

***Proof.*** *See appendix.* ■

Comparing Lemmas 2 and 3, we note two important differences from observing plan shares when plans are offered jointly rather than pairwise. First, for a given set of plans, the plan shares when all plans are offered jointly allow for tighter bounds. For pairwise comparisons the bounds need to be constructed using share differentials. The rotations of the type frontiers imply that plans differentially attract and discourage individuals from buying the plan, but we only observe the net differences. Second, depending on the underlying distribution the net difference may be positive or negative, which results in either a lower or upper bound on the support. For a given set of plans, with all plans offered jointly, the bounds only go in one direction, regardless of the underlying distribution.[19] However, this would not be an issue in more complex

[19] This is due to our restricted contract space in the binary risk framework. The dimensionality of the contract space and decreasing returns to coverage make that $\mathcal{T}(\emptyset, X_h)$ is a clockwise rotation of $\mathcal{T}(\emptyset, X_l)$ if and only if $\frac{P_h}{q_h} < \frac{P_l}{q_l}$, which implies that the type frontier $\mathcal{T}(X_l, X_h)$ is a clockwise rotation

contractual environments for which the dimensionality exceeds the dimensionality of the type space as we demonstrate in our empirical application. Such environments allow for bounds from both sides, which are necessary to reject homogeneity.

# 5 Application to Massachusetts' Health Insurance Exchange

In this section, we use health insurance plan choices by consumers on the Massachusetts Health Insurance Exchange (HIX) to illustrate our identification method. We use within-menu plan variation and derive informative bounds on the CDFs of risk preferences and expected costs of these consumers.

## 5.1 Exchange Context and Choice Menu

Established by the 2006 Massachusetts Health Reform, the Massachusetts HIX is the forerunner of the HIXs established across the U.S. by the 2010 Affordable Care Act (ACA). Data from the Massachusetts HIX allows us to examine consumer choice from a menu with a variety of plans, offered at posted prices on a guaranteed issue, non-health rated basis. We examine the choices of first-time choosers on the exchange, so that choices are "active choices" uncontaminated by inertia (Ericson 2014, Handel 2013).[20]

On the HIX, plans are available in multiple tiers based on actuarial value (the fraction of health care costs that would be insured for a representative sample of the population). In the time period we consider (January and February 2010), cost-sharing characteristics were standardized across insurers within each tier of quality: bronze low, bronze medium, bronze high, silver low, silver medium, gold. As a result, a bronze low plan would have the same deductible, coinsurance, and maximum out of pocket spending regardless of which insurer it was offered by.[21]

The menu of plans was designed by the HIX regulator, while prices were set by individual insurers. Within each tier, multiple insurers compete, offering plans that differ in premiums, brand and insurer network. Premiums thus vary by plan tier and by insurers. However, due to modified community rating regulation, the premium for a given insurer-plan combo is only allowed to vary by geography and age. In particular, premiums are only allowed to differ for each 5-year age group. This provides arguably exogenous price variation for comparable populations around age cut-offs, but one would need a larger sample to achieve sufficient statistical power to use this *between-menu* plan variation (as analyzed in Subsection 3.3). However, all age groups can

of $\mathcal{T}(\emptyset, X_h)$. In a more complex contractual environment, $\mathcal{T}(X_1, X_2)$ can still be a counter-clockwise rotation of $\mathcal{T}(\emptyset, X_1)$ when $\mathcal{T}(\emptyset, X_2)$ is a clockwise rotation of $\mathcal{T}(\emptyset, X_1)$. This allows for bounds in the opposite direction.

[20] The HIX is described in detail in Ericson and Starc (2012a, 2012b).

[21] Ericson and Starc (2013) describes the standardization process in more detail. The Massachusetts HIX tiers in this time period are slightly different from the ACA tiers—for instance, gold on the Massachusetts HIX is similar to Platinum on the ACA exchanges.

choose from a menu with several plans differing in coverage and price. We use this *within-menu* plan variation for identification (as analyzed in Subsection 4.2).

In order to model choice from this menu, we translate each plan design into a simplified plan design characterized solely by a deductible $D$, a coinsurance rate $\beta$, and maximum out-of-pocket spending $M$. In a contract characterized solely by these parameters, an individual's out of pocket spending is simply a plan-specific function of their total spending. However, contracts are in fact quite complex, with per-visit copayments that vary based on service used and per admission charges to the hospital. Modeling choice from such a complex contract would require modeling a very detailed level of health care utilization: for instance, how often consumers expect to use each type of specialist, each tier of prescription drug, and differentiating between expenditures for lab tests, durable medical equipment, allergy treatment, and inpatient spending. Our simplification procedure is also reasonable since it is unlikely that consumers observed, understood, and had well-formed expectations of the probability that they would use each of these varied services.

To translate the actual plan design into a simplified plan design $X$, we entered the original characteristics of each plan into the Center for Consumer Information & Insurance Oversight's (CCIIO) actuarial value calculator—including details such as per visit copayments, which produced an estimated actuarial value (AV) for that plan. Then, we solve for the coinsurance rate (given that plan's actual deductible $D$ and maximum OOP $M$) that would produce the same AV for the simplified version of each plan characterized by $(D, \beta, M)$.[22] We explore results using a variety of other alternative plan translations in the Empirical Appendix (see Appendix Figure A.2). Table 1 presents the results of this exercise, while Table A.1 describes the detailed design of the plans as sold on the Massachusetts HIX.[23] Premiums are different for each 5-year age group; we present the premiums for the lowest and highest priced age group, and focus our analysis on these groups.[24]

Plans in the table are ordered by their actuarial value, from least to most generous. While the actuarial values of the Bronze plans are quite similar, the plans vary in where they apply coverage: Bronze High has a very low deductible but correspondingly

[22] However, because the actual plans did indeed provide some coverage for spending below deductible (e.g. a \$100 doctor's visit resulted in a \$30 copay even if the deductible was not met), our method underestimated the degree of coinsurance. While the results were reasonably representative of the plans' characteristics, this method produced a 0% coinsurance rate for the Bronze Medium plan, even though this plan in fact did include cost-sharing after the deductible. We used a corrected coinsurance of 5% for Bronze Medium, based on dividing the \$500 hospital copay (as in the original plan characteristics) by the mean 2010 hospital stay cost of \$9700 (as reported in Pfuntner, Wier and Steiner 2013).

[23] In some months, a Silver Medium plan is also offered; when it is, we drop it from our plan menu, along with the small number of people who choose it from our calculation of market shares. Because the remainder of the individuals revealed they preferred one of the other plans to Silver Medium, our bounds are still describing the preferences and beliefs of our sample population. (The bounds we present are slightly looser than if we had used information about Silver Medium.)

[24] Premiums are averaged over the two months (there is small variation between January and February) and across zipcodes for all people offered the Neighborhood Health Plan (most people live in the Boston region).

higher coinsurance than Bronze Medium; all Bronze plans have the same maximum OOP. (Note that despite having a slightly higher actuarial value than Bronze Medium, Bronze High is priced slightly lower.) Silver Low is quite different as it has a lower maximum OOP, but a higher deductible relative to Bronze High. Silver High and Gold are quite similar again: both have zero deductible and a maximum OOP of $2000. While Gold is more generous based on actuarial value and has a lower coinsurance rate, it has higher premiums.

While multiple insurers offer plans, we focus our analysis on the price menu of the most popular insurer (Neighborhood Health Plan), which has approximately 50% market share. (The price for each plan design varies across insurers; we have explored using the prices for other insurers, which give similar results.) In all cases, our results apply to the population of individuals who chose this insurer. We do not explicitly model individual's choice of insurers. Tighter bounds could be obtained by modeling individuals' pattern of substitution between insurers, but we have limited data to identify these patterns.[25]

The final columns of Table 1 present market shares for the plan designs, broken down by broad age groups. Though prices vary by 5 year age groups, we group those above and below age 45 to get more accurate estimates of market shares (doing so reduced sampling error). See Appendix Table A.2 for detailed market shares within each 5 year age bin category.

### 5.2 Individual Model of Choice

We model individuals as having CARA utility over consumption: $u\left(-P-x\left(k\right)\right) = -\exp\left(\sigma\left(P+x\left(k\right)\right)\right)/\sigma$, where OOP expenses $x\left(k\right)$ are a function of the individual's healthcare spending $k$ and the insurance plan they choose. Individuals vary on two dimensions. First, they vary in their CARA coefficient $\sigma$. Second, they vary in their beliefs about the distribution of their own healthcare spending. While there are many dimensions on which individuals might vary in their distributional beliefs, we summarize variation in expected claims in a single risk-type index, $\pi$. For each risk-type $\pi$, the expected claims distribution is assumed to follow a log normal distribution with $mean = \pi$ and $variance = \frac{\pi}{4053}\left[\frac{1}{2}\times 10451\right]^2$. Note that variance of expenditures scales with the mean expected risk. We take the $4053 mean spending number from the 2010 Medical Expenditure Panel Survey, persons with private insurance. The standard deviation

[25] Our model is consistent with a variety of different ways in which individuals trade off their preferred plan design versus price and preferred insurer. For instance, individuals could make a hierarchical decision, choosing their preferred insurer first (based on insurer network versus insurer's average price), then choosing their preferred plan design. Then, our results simply describe the population of people whose preferred insurer was Neighborhood Health Plan. Alternatively, an individual may have a more complex pattern of substitution—for instance, a Blue Cross Bronze High plan may be the closest substitute to a Neighborhood Health Plan Silver Low plan. In this case, our bounds on preferences and beliefs still describe the population of individuals whose preferred plan was offered by Neighborhood Health Plan, since the plan they chose was indeed revealed preferred to all other plans offered by this insurer.

of expenses is \$10451. Someone with $\pi = 4053$ has the population average as his or her mean claim, but because individuals have information about their own risk type (age, gender, particular diseases, and expected patterns of care), we assume the individual's expected standard deviation is half the population standard deviation. Little is known about risk types and their structure. Under our assumptions, the variance of claims is lower for an individual with lower mean expected claims.[26] We have explored alternative variance assumptions, including a model of constant variance of claims across all risk types.[27] Note as well that we have assumed no moral hazard: expected healthcare spending is the same, regardless of which contract individuals choose.

To determine what can be learned from consumers choosing from the menu of options in Table 1, we construct a grid of $(\pi, \sigma)$ pairs, with $\sigma$ ranging from $10^{-15}$ to $0.5 \times 10^{-2}$ and $\pi$ ranging from 1/100 the population expected claims (about \$40 in expected claims) to 5 times the population expected claims (about $\$20,000$ in expected claims). Each $(\pi, \sigma)$ pair represents a combination of expected healthcare costs and risk aversion. We then calculate the plan that maximizes expected utility for each pair.

The first column of Figure 6 displays the optimal plan choice for the youngest group (Panel A, upper panel) and oldest group (Panel B, lower panel). Recall that prices vary between age groups, and the older group faces a higher marginal cost of more generous coverage. For both groups, only individuals with relatively low expected costs choose the Bronze Low (dark black) plan: it is chosen for only the lowest value of $\pi$ in Panel A, and the lowest two values of $\pi$ in Panel B. It is attractive for all individuals with such low expected costs regardless of risk aversion. Bronze Medium is similar to Bronze Low but with a lower coinsurance rate and priced slightly higher. It is only chosen by the older consumers at this set of relative prices (it does not appear in Panel A), and attracts relatively risk averse, but low-risk individuals. Bronze High is the most popular plan with a market share of 40.2% and 29.0% for the young and old respectively. The plan is attractive to relatively risk-neutral individuals with a wide range of expected claims, and to low expected-cost individuals with a wide range of risk aversion. The plan has a low deductible (\$250 vs. \$2000 for the other Bronze plans) and is cheaper than Bronze Medium, but has a higher co-insurance rate above the deductible.

Turning to Silver plans, we find that individuals with the highest expected costs choose Silver Low rather than Silver High; individuals with intermediate expected costs choose Silver High.[28] While the two silver plans have the same maximum OOP, the Silver High plan has a lower deductible but higher coinsurance; from the perspective of risk averse individuals, paying for first dollar coverage is less valuable than paying for

[26] For instance, both the mean and variance of medical spending rises with age.

[27] Appendix Figure A.1 shows how choices would shift if alternative variance structures were assumed. Intuitively, higher variance at a given amount of expected costs tends to increase demand for insurance.

[28] Silver High is slightly more generous than Silver Low (both in the sense of actuarial value and of higher premiums).

lower coinsurance. Despite the fact that Silver Low is preferred for many $(\pi, \sigma)$ pairs, the market share of Silver Low is relatively small: only about 3%. This indicates that there is not a large subset of the population with both very high risk aversion and very high expected claims.

Finally, note that no one in this menu chooses a Gold plan: its only advantage over Silver High is lower coinsurance, but it has substantially higher premiums. Thus, even though the Gold plan has the highest actuarial value, it exposes individuals to a worse worst-case scenario than the Silver plans. Someone who hits the maximum OOP of \$2000 in both Silver High and Gold will spend more in the Gold plan due to the higher premiums (an additional \$1392 at the premiums faced by older individuals). This explains why Gold is actually less attractive than Silver for someone who is very risk averse and expects to hit the OOP maximum.[29]

### 5.3 Bounds from Plan Choices

From these choices, we can use the plan shares in Table 1 to construct bounds on the CDFs of $\pi$ and $\sigma$. Column 2 of Figure 6 presents CDFs of each independently. The upper panel shows that choice provides virtually no restriction on the distribution of risk preferences in the population facing the young prices. Any single choice of the risk aversion parameter $\sigma$ (except the most risk neutral one) could rationalize all the choices. The only restriction on the distribution is that individuals choosing Silver Low cannot have the most risk neutral value of $\sigma$. This bound, however, is coming from our restriction on the domain of risk types, having assumed that an individual's expected claims cannot exceed $\$20,000$.

The bottom panel of Figure 6 shows that there must be at least some relatively risk-averse individuals to rationalize choice for older individuals given the prices they face. The bound is coming from the difference in plan features between Bronze and Silver plans which differentially attract types along the risk and preference dimension. Bronze Medium offers relatively generous coverage for intermediate costs and only attracts types with risk aversion $\sigma \geq \bar{\sigma}_I = 9.32 \times 10^{-4}$. Types with lower risk aversion should either buy Bronze High, providing more generous coverage for low costs, or Silver Low, providing more generous coverage for high costs. Similarly, we find that Silver High only attracts types with risk aversion $\sigma \geq \bar{\sigma}_{II} = 0.0011$.[30]

In line with Lemma 3, the share of older individuals with risk aversion greater than $\bar{\sigma}_{II} = 0.0011$, $1 - H_\sigma(\bar{\sigma}_{II})$, is at least as high the market share of Silver High and thus provides an upper bound on the CDF. The share of individuals with risk aversion above $\bar{\sigma}_I = 9.32 \times 10^{-4}$ is at least the sum of the market shares of Silver High and

[29] The market share of Gold is relatively small (only 8% for the old), but non-zero. In exploratory analysis, we do find that the plan becomes rationalizable under certain menus and variance assumptions.

[30] Types with lower risk aversion and relatively low risk should buy Bronze, providing lower coverage but at substantially lower premium. Types with lower risk aversion but high risk should again buy Silver Low.

Bronze Medium, providing a tighter upperbound on the CDF for this lower level of risk aversion. Despite our informative upper bound on the CDF, we cannot reject homogeneity in risk preferences because the available plan variation does not place any lower bounds on the CDF. As a consequence, we can fit a degenerate CDF that jumps from zero to one for risk-aversion levels above $\bar{\sigma}_{II} = 0.0011$. So while we can reject that all individuals would have relatively low risk aversion, we cannot reject that all individuals have some relatively high yet homogeneous risk aversion.

Turning to the distribution of risk types ($\pi$), we note that for each bound on risk aversion coming from the plan variation corresponds to a bound on risk as well. For example, Bronze Medium attracts types who not only have relatively high risk aversion ($\sigma \geq \bar{\sigma}_I$), but also expect low costs ($\pi \leq \bar{\pi}_I = \$1170$). Types with higher expected costs prefer the higher actuarial value of Bronze High or Silver depending on their risk preferences. The same is true for Silver High, which only attracts types with expected expenses $\pi \leq \bar{\pi}_{II} = \$1067$. In addition, the choice of Bronze Low, which provides the lowest coverage, can only be rationalized for types with very low expected costs ($\pi \leq \bar{\pi}_{III} = \$383$). The cumulative market shares of Bronze Low, Silver High and Bronze Medium provide a lower bound on the CDF of expected costs at respectively $\bar{\pi}_{III}, \bar{\pi}_{II}$ and $\bar{\pi}_I$. This is illustrated in the bottom figure of Column 2 of Figure 6

For the distribution of risk types, the market shares can also be used to provide upper bounds on the CDF. When risk preferences cannot exceed the extremely risk averse[31] $\sigma = 0.005$, as illustrated in Column 1 of Figure 6, we find strictly positive lower bounds on the support of expected expenses for each of the plan choices other than Bronze Low. The market shares for these plans allow us to construct upper bounds on the CDF of expected costs. Note that when we relax the constraint on the preference domain, we still find informative lower bounds on the support for some plans. For example, for the older individuals, Silver High (Bronze Medium) will only attract types with expected expenses above \$1069 (\$383), regardless of the risk preferences.[32]

The derived upper and lower bounds on the CDF imply that we can reject homogeneity in expected expenses. (We cannot fit a degenerate CDF jumping from 0 to 1 for some $\pi$.) Hence, while we can rationalize the different plan choices with only heterogeneity in expected expenses, we can not do it with only heterogeneity in risk preferences. Note that we have considered a wide candidate range for $(\sigma, \pi)$. To the extent you are willing to put further restrictions on the range of reasonable parameters, tighter bounds can be obtained.

[31] For $\sigma = 0.005$, an individual is indifferent between getting \$139 for certain and a 50-50 gamble for \$10,000 or \$0.

[32] Note that also in the high-variance specification in Panel B of Appendix Figure A.1, we can only rationalize Bronze Medium for a limited range of risk aversion parameters. The market share of Bronze Medium thus provides both a lower and upperbound on the CDF.

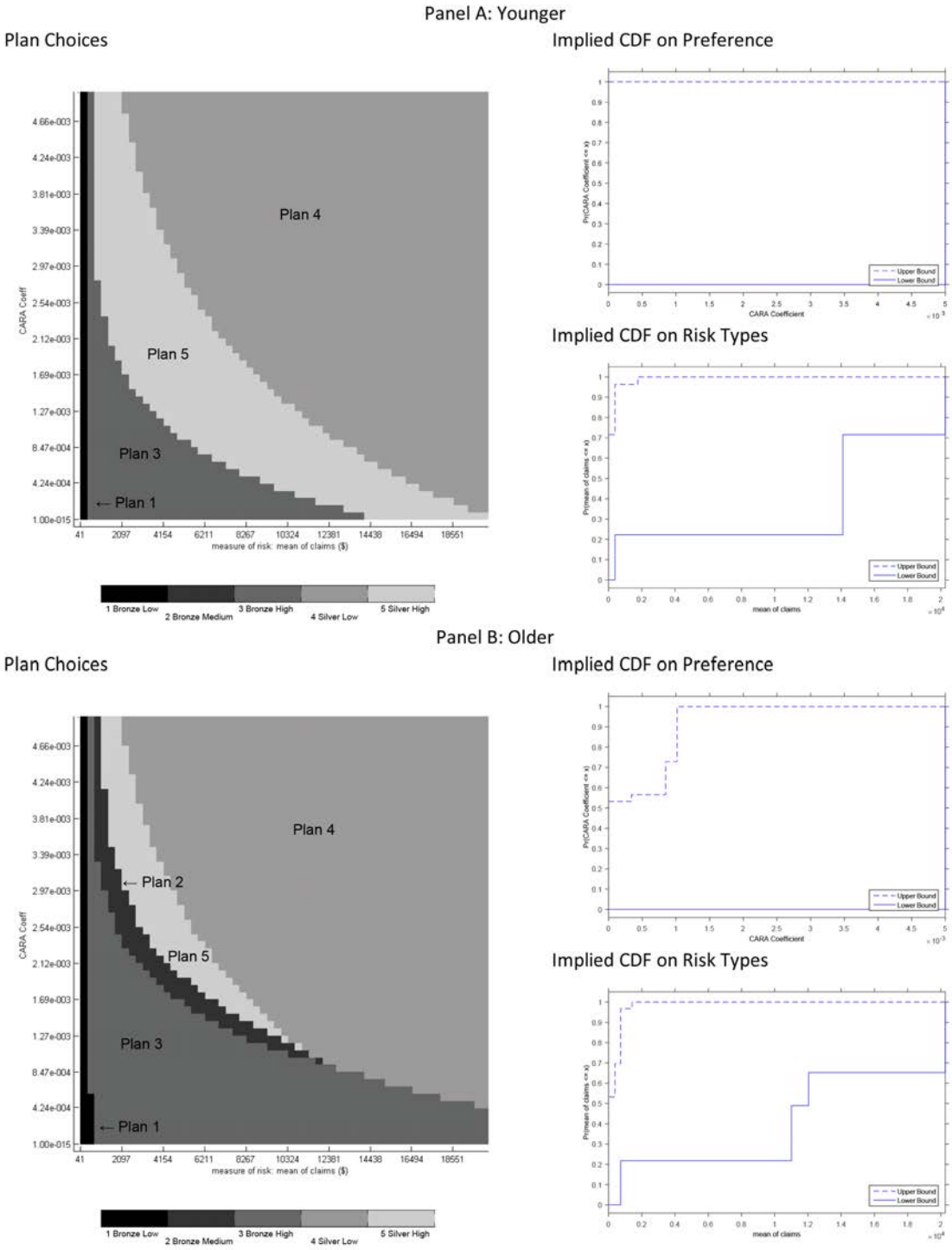


Figure 6: Choices and implied bounds on risk preferences and risk perceptions.

### 5.4 Discussion

A large empirical literature has argued that heterogeneity in risk preferences is a key feature of insurance markets and explains why adverse selection is a minor issue in several markets. The implementation of our non-parametric approach does not allow us to validate this claim in our empirical context. We cannot reject that all individuals have the same preferences, while they must differ in their (perceived) risks.

The non-parametric bounds on risk preferences, using only plan variation, do not allow us to distinguish between quite extreme forms of preference heterogeneity. A more structural approach could help to tighten bounds on preferences and prove complementary to our approach, but the tighter bounds would rely on the validity of the imposed structure. For comparison, Figure 7 plots our bounds on CARA preferences for the old group with some well-known examples in the insurance literature of parametric estimates of CARA distributions using standard random utility models. These estimates are obtained from different contexts and potentially very different populations. Our bounds do not reject the vast dispersion in risk aversion estimated by Cohen and Einav (2007), but are also consistent with the more homogeneous distribution estimated in Handel and Kolstad (2015). Interestingly, this is no longer true for the estimates in Handel and Kolstad (2015) obtained by augmenting the standard random utility model with survey data on information frictions. This could indicate that it is not sufficient to account for people's risk perceptions, and that our expected utility model should be augmented with other informational or behavioral frictions to provide consistent and tighter bounds on preference heterogeneity. Finally, more plan variation would allow us to further tighten bounds as well. The regulation of plan features or prices could provide promising variation for identification.[33]

## 6 Conclusion

This paper has shown how to identify both consumer risk preferences and their risk perceptions, using only insurance choice data. Our method uses variation in insurance plans that differentially attracts individuals along the preference and risk type dimensions, exploiting the fact that marginal willingness to buy insurance is more rapidly decreasing in coverage for individuals with high risk aversion (but low risk) than for individuals with low risk aversion (but high risk).

Our approach allows us to relax strong assumptions about (rational) expectations and parametric type distributions, as well as to identify preferences and risk perceptions when claims data is unavailable. We applied our method to the Massachusetts

[33] The discussed price variation across age groups would be useful for identification in combination with within-menu plan variation. Comparing the type sets at the young prices and the old prices reveals that changes in prices change the parameter values that bound the support of particular plans. When the price variation is exogenous, plan share differentials may be attributable to particular parameter ranges and thus provide further bounds.

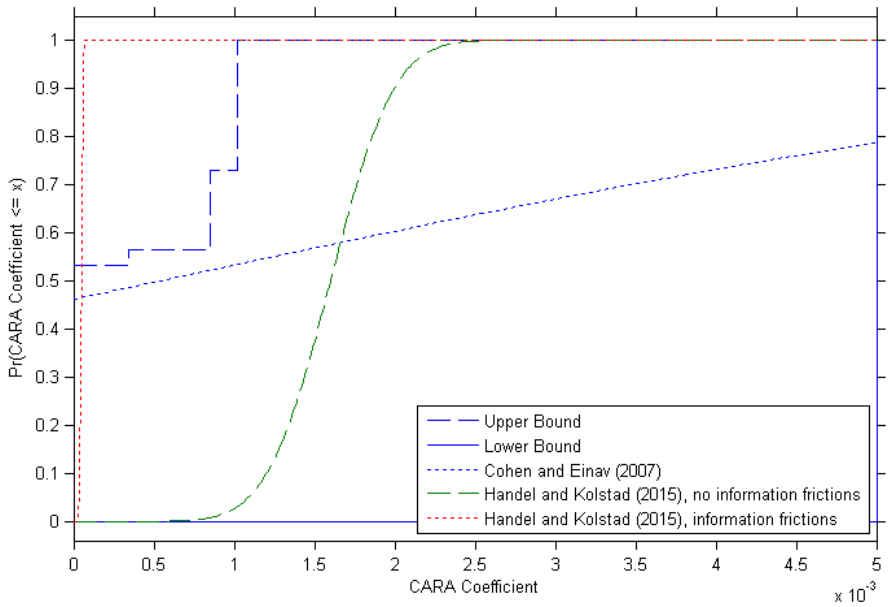


Figure 7: Comparing bounds on risk preferences for older individuals on the Massachusetts HIX to estimates from the literature

HIX. For these individuals, we can reject homogeneity in risks, but not homogeneity in preferences. We estimate bounds on the distribution of preferences that are consistent with other papers. We also highlight the type of variation that is necessary to obtain tighter bounds on the distribution of preferences, which may be useful for experimentalists eliciting preferences. Future empirical work could pair our approach with claims data to directly test the assumption of rational expectations about individuals' distribution of insurance claims. Moreover, future theoretical work could change the micro-foundations of the choice model (e.g., by adding loss aversion or ambiguity aversion) and then analyze which type of plan variation would allow to identify the primitives of that model.

# 7 References

Abaluck, J., and J. Gruber, 2011. Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program. American Economic Review, 101(4), 1180–1210.

Adams, A., Cherchye, L., De Rock, B., and E. Verriest. 2014. Consume Now or Later? Time Inconsistency, Collective Choice, and Revealed Preference. American Economic Review, 104(12): 4147-83.

Barseghyan, L., Molinari, F., O'Donoghue, T., and J. Teitelbaum, 2013. The Nature of Risk Preferences: Evidence from Insurance Choices. American Economic Review, 103(3), 2499-2529.

Barseghyan, L., Molinari, F., O'Donoghue, T., and J. Teitelbaum, 2015a. Estimating Risk Preferences in the Field. Working Paper.

Barseghyan, L., Molinari, F., and J. Teitelbaum, 2015b. Inference under Stability of Risk Preferences. Quantitative Economics, Forthcoming.

Bundorf, K., Levin, J. and N. Mahoney, 2012. Pricing and Welfare in Health Plan Choice. American Economic Review, 102(7), 3214-3248.

Cai, Q., C. Zhang and C. Peng, 2005. Learning Probability Density Functions from Marginal Distributions with Applications to Gaussian Mixtures. Proceedings of International Joint Conference on Neural Networks, Montreal, Canada, July 31- August 4, 1148-1153.

Chetty, R., and A. Finkelstein 2013. Social Insurance: Connecting Theory to Data. In the Handbook of Public Economics, Volume 5.

Chiappori, P., and B. Salanié, 2014. Asymmetric Information in Insurance Markets: Predictions and Tests. in Handbook of Insurance, 2nd edition (G. Dionne, ed.).

Chiappori, P., Gandhi, A., Salanié, B., and F. Salanié, 2009. Identifying Preferences under Risk from Discrete Choices. American Economic Review P&P 99(2), 356-362.

Chiappori, P., Gandhi, A., Salanié, B., Salanié, F., 2012. From Aggregate Betting Data to Individual Risk Preferences. working paper.

Choi, S., Fisman, R. ,Gale, D. and S. Kariv, 2007. Consistency and Heterogeneity of Individual Behavior under Uncertainty. American Economic Review 97(5), 1921-1938.

Cohen, A., Einav, L., 2007. Estimating Risk Preferences from Deductible Choice. American Economic Review 97(3), 745-788.

Cohen, A., Siegelman, P., 2010. Testing for Adverse Selection in Insurance Markets. Journal of Risk and Insurance 77(1), 39-84.

Crawford I., 2010. Habits Revealed. Review of Economic Studies, 77(4), 1382-1402.

Crawford I., De Rock B. 2014. Empirical Revealed Preference. Annual Review of Economics 6, 503-524.

Crawford I., Pendakur K. 2013. How many types are there? Economic Journal 123: 77-95.

Cutler, D., Finkelstein, A., and K. McGarry, 2008. Preference Heterogeneity and Insurance Markets: Explaining a Puzzle of Insurance. American Economic Review, 98(2), 157-162.

Dean M., Caplin A. 2015. Revealed Preference, Rational Inattention and Costly Information Acquisition, American Economic Review, 105 (7): 2183-2203.

Dean M., Martin D. 2015. Measuring Rationality with the Minimum Cost of Revealed Preference Violations. Review of Economics and Statistics, Forthcoming

Einav, L., Finkelstein, A., Cullen, M., 2010. Estimating Welfare in Insurance Markets using Variation in Prices. Quarterly Journal of Economics 125(3), 877-921.

Einav, L., Finkelstein, A., Levin, J., 2010. Beyond Testing: Empirical Models of Insurance Markets. Annual Review of Economics 2, 311-336.

Einav, L., Finkelstein, A., Schrimpf, P., 2010. Optimal Mandates and The Welfare Cost of Asymmetric Information: Evidence from the U.K. Annuity Market. Econometrica, 78(3), 1031-1092.

Ericson, Keith M. 2014. Consumer Inertia and Firm Pricing in the Medicare Part D Prescription Drug Insurance Exchange. American Economic Journal: Economic Policy 6 (1), 38–64.

Ericson, Keith M., and Amanda Starc, 2012a. Designing and Regulating Health Insurance Exchanges: Lessons from Massachusetts. Inquiry: The Journal of Health Care Organization, Provision, and Financing 49 (4), 327–38.

Ericson, Keith M., and Amanda Starc, 2012b. Heuristics and Heterogeneity in Health Insurance Exchanges: Evidence from the Massachusetts Connector. American Economic Review P&P 102 (3), 493–97.

Ericson, Keith M., and Amanda Starc, 2013. How Product Standardization Affects Choice: Evidence from the Massachusetts Health Insurance Exchange. NBER Working Paper 19527.

Ericson, Keith M. and Amanda Starc, 2015 Pricing Regulation and Imperfect Competition on the Massachusetts Health Insurance Exchange. Review of Economics and Statistics 97(3)-667-682.

Gandhi, A., and R. Serrano-Padial, 2014. Does Belief Heterogeneity Explain Asset Prices: The Case of the Longshot Bias. Review of Economic Studies, Forthcoming.

Grubb, M., 2015. Behavioral Consumers in Industrial Organization: An Overview. Review of Industrial Organization 47(3), 247-258

Handel, B., 2013. Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts. American Economic Review, 103 (7), 2643-2682.

Handel, B., and J. Kolstad, 2015. Health Insurance for Humans: Information Frictions, Plan Choice, and Consumer Welfare. American Economic Review, 105(8), 2449-2500.

Handel, B., Kolstad, J., and J. Spinnewijn, 2015. Information Frictions and Adverse Selection: Policy Interventions in Health Insurance Markets. NBER working paper 21759.

Johnson, E., Hershey, J., Meszaros, J., and H. Kunreuther, 1993. Framing, Probability Distortions, and Insurance Decisions. Journal of Risk and Uncertainty, 7(1): 35-51.

Lieber, E., 2015. Does it Pay to Know the Prices in Health Care?. University of Notre Dame. Working Paper.

Manski, C., 2004. Measuring Expectations, Econometrica, 72(5): 1329-1376

Mas-Colell, A., Whinston, M., Green, J., 1995. Microeconomic Theory. Oxford University Press. New York, NY.

Pfuntner, A., Wier, K. and C. Steiner, 2013. Costs for Hospital Stays in the United States, 2011." HCUP Statistical Brief #146

http://www.hcup-us.ahrq.gov/reports/statbriefs/sb146.pdf

Spinnewijn, J., 2012. Heterogeneity, Demand for Insurance and Adverse Selection. LSE mimeo

Sydnor, J., 2010. (Over)insuring Modest Risks. American Economic Journal: Applied Economics, 2(4), 177–99.

Table 1: HIX Plan Menu

| | | | | | Monthly Premium | | Market Share | |
|---|---|---|---|---|---|---|---|---|
| | Deductible | Coinsurance | Max OOP | AV | Youngest | Oldest | Under 45 | Over 45 |
| Bronze Low | 2000 | 11.20% | 5000 | 73.1 | $193 | $388 | 17.9% | 19.9% |
| Bronze Medium | 2000 | 5.00% | 5000 | 79.8 | $210 | $420 | 7.0% | 14.9% |
| Bronze High | 250 | 15.40% | 5000 | 85.2 | $202 | $405 | 40.2% | 29.0% |
| Silver Low | 1000 | 2.50% | 2000 | 85.6 | $273 | $540 | 3.4% | 2.9% |
| Silver High | 0 | 12.20% | 2000 | 92.2 | $275 | $543 | 19.6% | 25.4% |
| Gold | 0 | 10.30% | 2000 | 93 | $336 | $659 | 12.0% | 8.0% |

Note: Deductible and maximum OOP are taken directly from the original plan design. Coinsurance rate calculated as defined in the text. Actuarial values are calculated from original plan design using the CCIIO calculator. Premiums and market shares are for Neighborhood Health Plan, Jan. and Feb. 2010. Premiums are averaged across the two sample months and across ZIP codes.

# A.1 Appendices

## A.1.1 Proofs

**Proof of Proposition 1**

Recall that we consider insurance contracts ($q \in (0, L]$) that are in principle rationalizable ($P \in (0, q)$), and consumers with domain of risk types $\pi \in [0, 1]$ and risk-averse preference types $\sigma \in \mathbb{R}_+$. An individual of type $(\pi, \sigma)$ buys plan $X$ if and only if condition (4) is satisfied.

First, we show that preference homogeneity cannot be rejected. Assume that all individuals have identical preference type $\sigma > 0$. By (4), we can find a sufficiently low risk type $\pi_l$ for which it is optimal to remain uninsured (choose $\emptyset$) and a sufficiently high risk type $\pi_h$ for which it is optimal to buy insurance (choose $X$). Assigning the appropriate fractions of agents to each of these risks then rationalizes the observed choices.

Second, we cannot put any bounds on preferences. Assume a bound of the form $H_\sigma(\bar{\sigma}) \geq \alpha$ for $\alpha > 0$. Yet we just showed that any homogeneous $\sigma > 0$ can rationalize any dataset under appropriate risk assumptions, and therefore in particular $\sigma > \bar{\sigma}$. Such a homogeneous $\sigma$ then means $H_\sigma(\bar{\sigma}) = 0$, yielding a contradiction. Similarly, for bound $1 - H_\sigma(\bar{\sigma}) \geq \alpha$ for $\alpha > 0$, we get a contradiction by rationalizing the observations for homogeneous $\sigma \leq \bar{\sigma}$.

Regarding risk types, the restriction to risk-averse preferences implies that risk types that would buy no insurance are bounded from above by risk type $\pi = P/q$, corresponding to the risk for which a risk-neutral type is indifferent between $X$ and $\emptyset$. Hence, we can use the share of individuals not buying insurance to bound $H_\pi$, but we can still not reject homogeneity in risks. For any risk type $\pi \in [\varepsilon, P/q]$ for $\varepsilon > 0$, by Assumption 1 we can find a (sufficiently low) preference type $\sigma_l$ for which it is optimal to remain insured and a (sufficiently high) preference type $\sigma_h$ for which it is optimal to buy insurance. The former is ensured because the risk-neutral agent is assumed to be in the domain, and for the latter it is assumed that risks span up to infinite risk aversion. With homogeneous risk type $\pi$ but appropriate fractions of agents with preference $\sigma_l$ and $\sigma_l$ any fraction of observed choices can be rationalized.□

**Proof of Proposition 3**

This proof provides rigor to the outline in the main text. Using Lemma 2, we can find two menus $\{\emptyset, X_h\}$ and $\{\emptyset, X_l\}$ with $q_h > q_l$ that intersect at an interior intersection $(\bar{\pi}, \bar{\sigma})$. If $\alpha_h = D(X_h | \{\emptyset, X_h\})$ is higher than $\alpha_l = D(X_l | \{\emptyset, X_l\})$, we know that $H_\sigma(\bar{\sigma}) \geq \alpha_h - \alpha_l$ and thus $H_\sigma(\sigma) \geq \alpha_h - \alpha_l$ for any $\sigma \geq \bar{\sigma}$ since the CDF is (weakly) increasing. At the same time, $1 - H_\pi(\bar{\pi}) \geq \alpha_h - \alpha_l$ and thus $H_\pi(\pi) \leq H_\pi(\bar{\pi}) \leq 1 - [\alpha_h - \alpha_l]$ for any $\pi \leq \bar{\pi}$. Hence, the plan share difference $\alpha_h - \alpha_l$ provides a lower bound on the CDF of preferences (for $\sigma \geq \bar{\sigma}$) and its complement an upper bound on the CDF of risks (for $\pi \leq \bar{\pi}$). Similarly, if $\alpha_h < \alpha_l$, the plan share difference $\alpha_l - \alpha_h$

places an upper bound on the CDF of preferences (for $\sigma \leq \bar{\sigma}$) and its complement an upper bound on the CDF of risks (for $\pi \geq \bar{\pi}$). Hence, any permissable distribution with $\alpha_h \neq \alpha_l$ places a bound on the marginal CDFs.

Consider now a third menu $\{\emptyset, X_h'\}$, where the plan $X_h'$ provides more coverage than the previous high-coverage plan $X_h$ (i.e., $q_h' > q_h > q_l$). If the price of the new plan were set at $P_h'$ such that the price per unit of coverage remains unchanged relative to the old high-coverage plan ($P_h'/q_h' = P_h/q_h$), the type $(\bar{\pi}', \bar{\sigma}')$ that is indifferent between these two plans is the risk-neutral type $(P_h/q_h, 0)$, while otherwise Assumption 1 implies that the type frontier $\mathcal{T}\{\emptyset, X_h'\}$ would be strictly steeper and therefore strictly above the type frontier of the previous plan $\mathcal{T}\{\emptyset, X_h\}$. Instead of this price, assume the price $P_h^{'}$ is set slightly lower so that $P_h^{'}/q_h' < P_h/q_h$ but still $P_h^{'}/q_h' \approx P_h/q_h$. The risk-neutral type $(P_h/q_h, 0)$ now strictly prefers the new plan over the old high-coverage plan, but by continuity the intersection $(\bar{\pi}', \bar{\sigma}')$ between $\mathcal{T}\{\emptyset, X_h'\}$ and $\mathcal{T}\{\emptyset, X_h\}$ remains close to $(P_h/q_h, 0)$. Since the intersection $(\bar{\pi}, \bar{\sigma})$ between the original plans $\mathcal{T}\{\emptyset, X_h\}$ and $\mathcal{T}\{\emptyset, X_l\}$ was placed in the interior of the type space, it had strictly higher risk-aversion and strictly lower risk than this risk-neutral type, and we have $\bar{\sigma} > \bar{\sigma}'$ and $\bar{\pi} < \bar{\pi}'$.

If now for a permissable distribution more agents choose the low contract $X_l$ over no insurance than choose the high contract $X_h$ over no insurance ($\alpha_l > \alpha_h$), but also more agents choose the new contract $X_h'$ over no insurance than those that choose the old high contract over no insurance ($\alpha_h < \alpha_h' \equiv D(X_h' | \{\emptyset, X_h'\})$), we will have that $H_\sigma(\bar{\sigma}) \leq 1 - [\alpha_l - \alpha_h] < 1$ while $H_\sigma(\bar{\sigma}') \geq \alpha_h' - \alpha_h > 0$ by the logic of the first paragraph of this proof. Since a CDF is weakly increasing and $\bar{\sigma}' < \bar{\sigma}$, we cannot fit a degenerate CDF between this lower and upper bound. That is, the lower bound becomes binding at $\bar{\sigma}'$, before the upper bound stops binding at $\bar{\sigma}$. We can thus reject homogeneity in preferences. The same is true for risks.

The final step in the proof is to show that such a distribution exists. To do this, define for any risk $\pi$ the preference $\sigma_l(\pi)$ that makes the person indifferent between no insurance and the low contract, i.e., $(\pi, \sigma_l(\pi)) \in \mathcal{T}\{\emptyset, X_l\}$, when it exists. Otherwise, $\sigma_l(\pi) = 0$. Define $\sigma_h(\pi)$ ($\sigma_h'(\pi)$) analogously via indifference between no insurance and the high insurance (new higher insurance) contract. The non-empty set of types $\Delta_{l,h} = \{(\pi, \sigma) | \sigma_h(\pi) > \sigma > \sigma_l(\pi)\}$ then prefer the low contract to no insurance which they prefer to the original high contract. Similarly, the non-empty set of types $\Delta_{h',h} = \{(\pi, \sigma) | \sigma_h(\pi) > \sigma > \sigma_h'(\pi)\}$ prefer the new contract to no insurance which they prefer to the old high coverage contract. Now we can construct a type distribution $H$ by placing strictly positive mass on types both in $\Delta_{l,h}$ and in $\Delta_{h',h}$, but nowhere else. This implies that $\alpha_l > 0, \alpha_h' > 0$ but $\alpha_h = 0$, which fulfills the premise of the previous paragraph (as do an uncountable number of other distributions with less stark properties).□

**Proof of Proposition 4**

Equation (11) in the main text showed that $F_{(\alpha,\beta)}(t) = \Pr(\alpha A + \beta B \leq t)$ is observed for all $\alpha, \beta$ and $t$. So we observe the marginal distribution $F_{(\alpha,\beta)}$ of $\alpha A + \beta B$, for all $\alpha, \beta$.

Therefore we know its characteristic function $\hat{F}_{(\alpha,\beta)}(\tau)$ for all $\alpha$ and $\beta$. We are interested in the joint cumulative distribution function $F(A,B)$ over $A$ and $B$, or equivalently in its characteristic function $\hat{F}(a,b)$.

The following just recalls the definition of the characteristic function for a random vector in $\mathbb{R}^k$ with cumulative distribution function $G(x)$ with $x \in \mathbb{R}^k$. Its characteristic function $\hat{G}(\omega)$ with $\omega \in \mathbb{R}^k$ is defined as

$$\hat{G}(\omega) = \int e^{i\omega^T x} dG(x)$$

where $\omega^T$ is the transpose of $\omega$ and $i$ is the imaginary unit.

The remaining identification follows the proof in Cai, Zhang and Peng (2005). At any value of $\alpha$ and $\beta$ we can apply the definition of the characteristic function twice (once for the two-dimensional random vector and once for the one-dimensional marginal random vector) to obtain

$$\begin{aligned} \hat{F}(\alpha\tau, \beta\tau) &= \int e^{i(\alpha\tau A + \beta\tau B)} dF \\ &= \int e^{i\tau(\alpha A + \beta B)} dF = \hat{F}_{(\alpha,\beta)}(\tau). \end{aligned}$$

Therefore, $\hat{F}_{(\alpha,\beta)}(1)$ varied over all $\alpha$ and $\beta$ identifies $\hat{F}(\alpha,\beta)$ and therefore identifies $F(A,B)$. Finally, by the one-to-one mapping between $(A,B)$ and $(\pi,\sigma)$ in case of CARA preferences, this identifies the distribution of risk and preference types as well.□

**Proof of Proposition 5.**

It is left to show that we can test for homogeneity in risks with two prices. For this we can use equations (12) - (13) for the variance and we can exploit similar expressions for the average:

$$\begin{aligned} E(q|p) &= E(A) - E\left(\sigma^{-1}\right) \times \tilde{p}, \text{ and} \\ E(q|p_1) - E(q|p_2) &= E\left(\sigma^{-1}\right) \times [\tilde{p}_2 - \tilde{p}_1] \end{aligned} \tag{16}$$

Using the fact that $A = L + \log\left(\frac{\pi}{1-\pi}\right)\sigma^{-1}$ under CARA, we know that if the the risk type $\pi$ were to be homogenous, we could infer the homogeneous risk type from

$$\begin{aligned} E(q|p) &= E(A) - E\left(\sigma^{-1}\right) \times \tilde{p} \\ &= L + \log\left(\frac{\pi}{1-\pi}\right) E\left(\sigma^{-1}\right) + E\left(\sigma^{-1}\right) \times \tilde{p}, \end{aligned}$$

where we know $E\left(\sigma^{-1}\right)$ from the difference in coverage choices in (16). For a homoge-

neous risk type, we also know that

$$\begin{aligned} Var\left(A\right) &= \log\left(\frac{\pi}{1-\pi}\right)^2 Var\left(\sigma^{-1}\right) \\ Cov\left(A,\sigma^{-1}\right) &= \log\left(\frac{\pi}{1-\pi}\right) Var\left(\sigma^{-1}\right). \end{aligned}$$

and thus

$$\begin{aligned} Var\left(q|p_k\right) &= Var\left(A\right) + Var\left(\sigma^{-1}\right) \times \tilde{p}_k^2 - 2Cov\left(A,B\right)\tilde{p}_k \\ &= \left[\log\left(\frac{\pi}{1-\pi}\right)^2 + \tilde{p}_k^2 - 2\log\left(\frac{\pi}{1-\pi}\right)\tilde{p}_k\right] Var\left(\sigma^{-1}\right). \end{aligned}$$

Hence, we can reject homogeneity in risk types if

$$\frac{Var\left(q|p_1\right)}{Var\left(q|p_2\right)} \neq \frac{\log\left(\frac{\pi}{1-\pi}\right)^2 + \tilde{p}_1^2 - 2\log\left(\frac{\pi}{1-\pi}\right)\tilde{p}_1}{\log\left(\frac{\pi}{1-\pi}\right)^2 + \tilde{p}_2^2 - 2\log\left(\frac{\pi}{1-\pi}\right)\tilde{p}_2}.$$

□

**Proof of Lemma 1.**

For CARA preferences $u(k|\sigma) = -e^{-\sigma k}/\sigma$, the marginal rate of substitution (6) can be written as

$$MRS \equiv -\frac{dm_g}{dm_b}|_{U(X|\pi,\sigma)} = \frac{\pi}{1-\pi}\frac{e^{\sigma(P+L-q)}}{e^{\sigma P}} \tag{17}$$

The type frontier $\mathcal{T}(\emptyset, X)$ is the set of types $(\pi, \sigma)$ for which (4) holds with equality, which for CARA preferences reads as:

$$\frac{\pi}{1-\pi}\frac{-e^{\sigma(P+L-q)} + e^{\sigma L}}{-1 + e^{\sigma P}} = 1 \tag{18}$$

Note that smaller $\pi$ are associated with larger $\sigma$, and $\pi \to 0$ is associated with $\sigma \to \infty$. Since we evaluate (17) only along (18), we can substitute the latter into the former to obtain a marginal willingness to pay along the type frontier of

$$\begin{aligned} MRS|_{(\pi,\sigma)\in\mathcal{T}(\emptyset,X)} &= \frac{-1 + e^{\sigma P}}{-e^{\sigma(P+L-q)} + e^{\sigma L}}\frac{e^{\sigma(P+L-q)}}{e^{\sigma P}} \\ &= \frac{1 - e^{-\sigma P}}{-1 + e^{\sigma(q-P)}}. \end{aligned}$$

Since $P < q$, it is immediate that $\lim_{\sigma\to\infty} MRS|_{(\pi,\sigma)\in\mathcal{T}(\emptyset,X)} = 1/\infty = 0$, which establishes that $MRS$ goes to zero as $\pi$ goes to zero. Moreover, $MRS$ is monotonically

decreasing in $\sigma$ along the type frontier (and thus monotonically increasing in $\pi$) if

$$\frac{d\,MRS|_{(\pi,\sigma)\in T(\emptyset,X)}}{d\sigma} = \frac{Pe^{-\sigma P}\left(-1+e^{\sigma(q-P)}\right)-(q-P)e^{\sigma(q-P)}\left(1-e^{-\sigma P}\right)}{\left(-1+e^{\sigma(q-P)}\right)^2}$$

is strictly negative. This arises if the denominator is strictly negative, i.e., if

$$\begin{aligned} Pe^{-\sigma P}\left(-1+e^{\sigma(q-P)}\right)-(q-P)e^{\sigma(q-P)}\left(1-e^{-\sigma P}\right) &< 0 \\ \Leftrightarrow -q(1-e^{-\sigma P})+P(1-e^{-\sigma q}) &< 0 \\ \Leftrightarrow P\left(1-e^{-\sigma P}\right)^{-1}-q(1-e^{-\sigma q})^{-1} &< 0. \end{aligned}$$

which holds since $P<q$ and $x/(1-e^{-\sigma x})$ is increasing in $x$.□

**Proof of Lemma 2.**

This proof follows the outline in the main text. We consider the type frontiers for two menus $\mathcal{M}_h=\{\emptyset,X_h\}$ and $\mathcal{M}_l=\{\emptyset,X_l\}$ with $q_h>q_l$. We first establish that if the two type frontiers intersect, they only intersect once and the high-coverage type frontier $\mathcal{T}(\emptyset,X_h)$ is a clockwise rotation of the low-coverage type frontier $\mathcal{T}(\emptyset,X_l)$. Denote the type at which the two frontiers intersect by $(\bar\pi,\bar\sigma)$. Consider the case where $q_h=q_l+\varepsilon$ for some small $\varepsilon$. By Assumption 1, any type with higher risk $\pi$ (lower preference $\sigma$) on $\mathcal{T}(\emptyset,X_l)$ than the type at the intersection, who is indifferent between the high-coverage and low-coverage plan, has higher marginal willingness to pay for the additional coverage. Therefore, they strictly prefer $X_h$ to both $X_l$ and $\emptyset$, which they are indifferent about. Hence, the type frontier $\mathcal{T}(\emptyset,X_h)$ lies to the left of $\mathcal{T}(\emptyset,X_l)$ for $\pi>\bar\pi$ and $\sigma<\bar\sigma$. Any type with lower risk $\pi$ (higher preference $\sigma$) has lower willingness to pay for the additional coverage and thus strictly prefers $X_l$ and $\emptyset$ to $X_h$. Hence, the type frontier $\mathcal{T}(\emptyset,X_h)$ lies to the right of $\mathcal{T}(\emptyset,X_l)$ for $\pi<\bar\pi$ and $\sigma>\bar\sigma$. This proves that $\mathcal{T}(\emptyset,X_h)$ intersects $\mathcal{T}(\emptyset,X_l)$ once and clockwise, if the two intersect. Now for a larger difference in coverage, we can find a sequence of contracts $X_k$ with coverage $q_k$ and price $P_k$, starting from $X_l$ and converging to $X_h$, such that type $(\bar\pi,\bar\sigma)$ is indifferent among any two contracts. The reasoning above now applies for any two consecutive contracts. Our sequence thus corresponds to a sequence of type frontiers that intersect only once and imply clockwise rotations around $(\bar\pi,\bar\sigma)$. Hence, this is also true for $\mathcal{T}(\emptyset,X_h)$ relative to $\mathcal{T}(\emptyset,X_l)$.

We now establish when the two type frontiers intersect. Consider first the case $P_h/q_h>P_l/q_l$ (i.e., the average price per unit is higher for the high-coverage contract $X_h$). This implies that the risk-neutral type with $\pi=P_l/q_l$ strictly prefers $X_l$ and $\emptyset$ (which he is indifferent about) to buying $X_h$. Hence, the type frontier $\mathcal{T}(\emptyset,X_h)$ lies to the right of the type frontier $\mathcal{T}(\emptyset,X_l)$ for $\sigma=0$. This implies that the two frontiers cannot intersect, since $\mathcal{T}(\emptyset,X_h)$ would be a clockwise rotation of $\mathcal{T}(\emptyset,X_l)$ and thus to the left of it for $\sigma=0$ in case the type frontiers were to intersect.

Consider now the case that $P_h/q_h \leq P_l/q_l$. In this case, the risk neutral type with $\pi = P_l/q_l$ prefers $X_h$ above $X_l$ and $\emptyset$. Moreover, since the marginal willingness to pay for the additional coverage converges to zero when moving up along the frontier $\mathcal{T}(\emptyset, X_l)$, there is a type with sufficient low risk (and high preference) that prefers $X_l$ (and thus $\emptyset$) above $X_h$ as long as $P_h > P_l$. Hence, the two type frontiers intersect. However, if $P_h \leq P_l$, all types on $\mathcal{T}(\emptyset, X_l)$ strictly prefer $X_h$ above $X_l$ and thus $\emptyset$. The two type frontiers again do not intersect. This proves the first part of the Proposition.

Since $\mathcal{T}(\emptyset, X_h)$ is a clockwise rotation of $\mathcal{T}(\emptyset, X_l)$ around $(\bar{\pi}, \bar{\sigma})$, the high-coverage contract $X_h$ differentially attracts types with high risk, but low preference. Types that prefer $X_h$ above $\emptyset$, but $\emptyset$ above $X_l$ (i.e., $\mathcal{B}(X_h|\{\emptyset, X_h\}) \setminus \mathcal{B}(X_l|\{\emptyset, X_l\})$), need to have preference $\sigma \leq \bar{\sigma}$ and risk $\pi \geq \bar{\pi}$. Only individuals with such types could rationalize that plan $X_h$ attracts a larger share of the population than plan $X_l$. Similarly, types that prefer $X_l$ above $\emptyset$, but $\emptyset$ above $X_h$ (i.e., $\mathcal{B}(X_l|\{\emptyset, X_l\}) \setminus \mathcal{B}(X_h|\{\emptyset, X_h\})$), need to have preference $\sigma \geq \bar{\sigma}$ and risk $\pi \leq \bar{\pi}$. Only these types could rationalize that plan $X_l$ attracts a larger share of the population than plan $X_h$. Hence, we have

$$\begin{aligned}
\int_{\pi \geq \bar{\pi}} \int_{\sigma \leq \bar{\sigma}} dH &\geq \int_{\mathcal{B}(X_h|\{\emptyset,X_h\}) \setminus \mathcal{B}(X_l|\{\emptyset,X_l\})} dH \\
&\geq \int_{\mathcal{B}(X_h|\{\emptyset,X_h\}) \setminus \mathcal{B}(X_l|\{\emptyset,X_l\})} dH - \int_{\mathcal{B}(X_l|\{\emptyset,X_l\}) \setminus \mathcal{B}(X_H|\{\emptyset,X_H\})} dH \\
&= \alpha_h - \alpha_l \\
&\geq -\int_{\mathcal{B}(X_l|\{\emptyset,X_l\}) \setminus \mathcal{B}(X_H|\{\emptyset,X_H\})} dH \\
&\geq -\int_{\pi \leq \bar{\pi}} \int_{\sigma \geq \bar{\sigma}} dH,
\end{aligned}$$

which proves the second part of the proposition. Note that if the type frontiers do not intersect, the support of the set of types that prefer the one plan, but not the other, covers the entire range of the preference domain. The differential plan share no longer places a bound on the distribution of preferences.□

**Proof of Lemma 3.**

By Lemma 2, we know that type frontiers $\mathcal{T}(\emptyset, X_h)$ and $\mathcal{T}(\emptyset, X_l)$ intersect if and only if $P_h/q_h \leq P_l/q_l$ and $P_h > P_l$. We denote this intersection by $(\bar{\pi}, \bar{\sigma})$. In this case, the type frontier $\mathcal{T}(X_h, X_l)$ intersects both frontiers again at $(\bar{\pi}, \bar{\sigma})$, since this intersection type is indifferent among both plans and the option not to buy insurance. Moreover, the type frontier $\mathcal{T}(X_h, X_l)$ is a clockwise rotation of $\mathcal{T}(\emptyset, X_h)$, which is a clockwise rotation of $\mathcal{T}(\emptyset, X_l)$. Note first that the willingness to choose the high-coverage plan over the low-coverage plan is increasing in both risk and preference. The type frontier is monotonically decreasing in $(\pi, \sigma)$-space, just like the original two frontiers. Now consider a type on the frontier $\mathcal{T}(\emptyset, X_h)$ above the intersection (with low risk, but high preference). This type strictly prefers $X_l$ to $\emptyset$ and thus $X_h$, since

$\mathcal{T}(\emptyset, X_h)$ is to the right of $\mathcal{T}(\emptyset, X_l)$. Hence, the type frontier $\mathcal{T}(X_h, X_l)$ is to the right of $\mathcal{T}(\emptyset, X_h)$. The set of types choosing $X_l$ above both $X_h$ and $\emptyset$, i.e., $\mathcal{B}(X_l|\{\emptyset, X_l, X_h\})$ corresponds to this region between the two frontiers $\mathcal{T}(\emptyset, X_l)$ and $\mathcal{T}(X_h, X_l)$ above $(\bar{\pi}, \bar{\sigma})$. Indeed, consider a type on the frontier $\mathcal{T}(\emptyset, X_h)$ below the intersection (with high risk, but low preference). This type strictly prefers $\emptyset$ and thus $X_h$ to $X_l$. Hence, the type frontier $\mathcal{T}(X_h, X_l)$ is to the left of $\mathcal{T}(\emptyset, X_h)$ (and thus to the left of $\mathcal{T}(\emptyset, X_l)$). This implies that no type with $\sigma < \bar{\sigma}$ or $\pi > \bar{\pi}$ will choose the low-coverage plan. It immediately follows that the share of individuals buying the low-coverage plan (out of this 3-options menu) puts the following lower bound,

$$\int_{\pi \leq \bar{\pi}} \int_{\sigma \geq \bar{\sigma}} dH \geq \int_{\mathcal{B}(X_l|\{\emptyset, X_l, X_h\})} dH \geq D(X_l|\mathcal{C}).$$

For completeness, the set of types choosing $X_h$ above $X_l$ and $\emptyset$, i.e., $\mathcal{B}(X_h|\{\emptyset, X_l, X_h\})$ corresponds to the region to the right of $\mathcal{T}(\emptyset, X_h)$ below $(\bar{\pi}, \bar{\sigma})$ and to the right of $\mathcal{T}(X_h, X_l)$ above $(\bar{\pi}, \bar{\sigma})$, as illustrated in Figure 5.

Note that if $P_h \leq P_l$, no type will ever buy the low-coverage plan. Hence, the only relevant type frontier is $\mathcal{T}(\emptyset, X_h)$. If $P_h > P_l$ and $P_h/q_h > P_l/q_l$, none of the type frontiers intersect. The type frontier $\mathcal{T}(X_h, X_l)$ now lies to the right of the type frontier $\mathcal{T}(\emptyset, X_h)$, which lies to the right of type frontier $\mathcal{T}(\emptyset, X_l)$. Types to the right of $\mathcal{T}(X_h, X_l)$ will buy the high-coverage plan. Types to the left of $\mathcal{T}(\emptyset, X_l)$ will buy no insurance. Types in between will buy the low-coverage plan. Since the support of any of the choices corresponds to the full preference domain, we can place no bounds on the distribution of preferences.□

## A.2 Empirical Appendix

In this Appendix, we present the optimal plan choice results under different assumptions. In Figure A.1, we model alternative relationships between the mean and variance of claims. In our main analyses, for risk type $\pi$, the expected claims distribution is assumed to follow a log normal distribution with $mean = \pi$ and $variance = \frac{\pi}{4053}\left[\frac{1}{2} \times 10451\right]^2$

For comparison purposes, we show, once again, the optimal choice for each $\pi$, $\sigma$ pair for older individuals in Panel C of Figure A.1. We then model choice with more or less variability in claims. In Panel A of Figure A.1, we show the choice assuming that the variance of claims is half of that in our main specifications: $variance = \frac{1}{2}\frac{\pi}{4053}\left[\frac{1}{2} \times 10451\right]^2$. This would correspond to a case in which individuals have additional information predicting about their expected costs, reducing variability. Then, in Panel B, we run a high variance specification where variance is twice that in our main specifications: $variance = 2 \times \frac{\pi}{4053}\left[\frac{1}{2} \times 10451\right]^2$. The results are intuitive: more variability increases the demand for more generous insurance.

We then turn to alternative menu designs in Figure A.2, again showing optimal choices for older individuals. Panel A examines a modified menu, in which the Bronze Medium plan has a deductible of \$1462 instead of \$2000; we make this modification so that the actuarial value of the Bronze Medium plan as modeled matches the actuarial value of the more complex Bronze Medium plan on the exchange. This menu leads to some modest changes in choice as compared to our main specification. In Panel B, we consider the case in which the Bronze Medium plan has zero coinsurance as produced by our original method described in the text. Unsurprisingly, this leads to Bronze Medium being a very favored plan. However, this is unlikely to be a faithful representation of the Bronze Medium characteristics. Finally, Panel C of Figure A.2 examines a very different menu design. For Panel C, we construct coinsurance values (for plans that have copayments instead of coinsurance) by taking the hospital copayment value and dividing by the mean cost of a hospital admission of \$9700. This method, however, does not do a good job modeling the relative quality of Silver Low, as Silver Low requires paying the deductible and then has zero hospital copayment. We then drop Silver Low from this menu. The menu of coinsurance values used in Panel C is given below:

| | Coinsurance for Panel C of Figure A.2 |
|---|---|
| Bronze Low | 0.2 |
| Bronze Medium | 0.05 |
| Bronze High | 0.35 |
| Silver High | 0.05 |
| Gold | 0.02 |

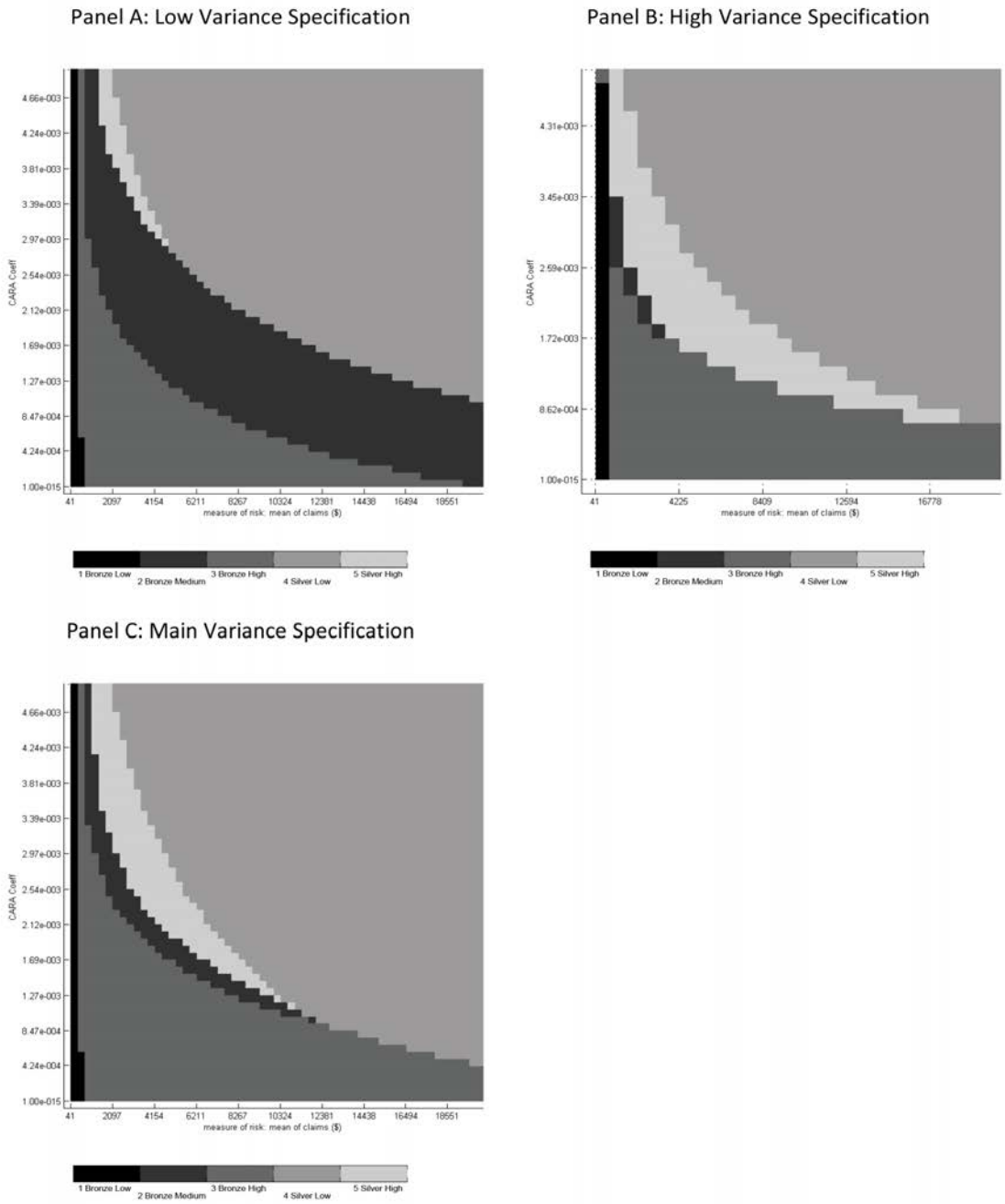


Figure A.1: Optimal plan choices for older individuals under alternative variance assumptions.

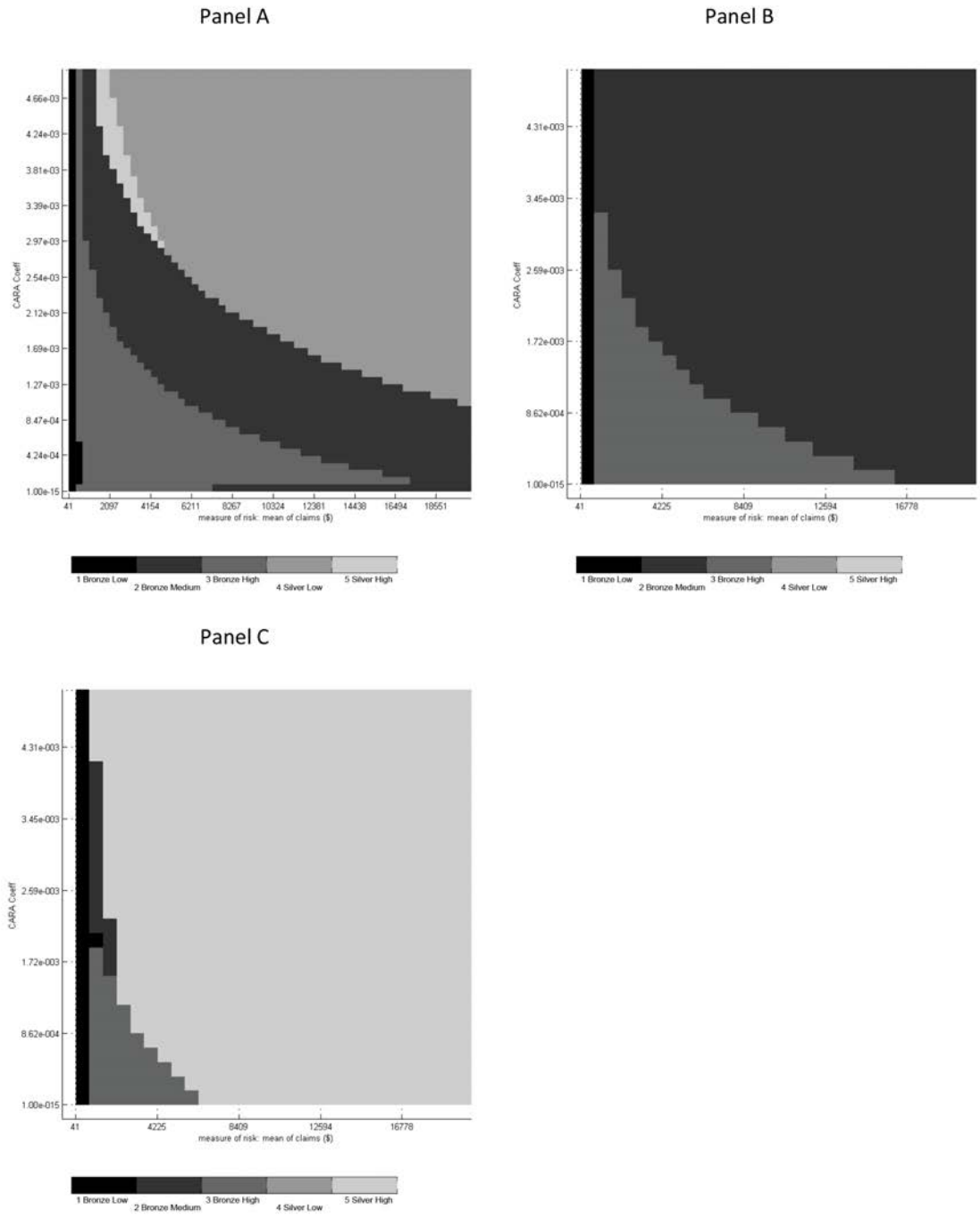


Figure A.2: Optimal plan choices for older individuals under alternative menu designs.

Table A.1: Summary of detailed plan parameters, taken from the HIX's website

| Plan Design | Deductible | Max OOP | Doctor Visit | Generic Rx | Emergency Room | Hospital Stay |
|---|---|---|---|---|---|---|
| Bronze Low | $2000 | $5000 | deduct., then $25 copay | deduct., then $15 copay | deduct., then $100 copay | deduct., then 20% co-insurance |
| Bronze Medium | $2000 | $5000 | $30 copay | $10 copay | deduct., then $150 copay | deduct., then $500 copay |
| Bronze High | $250 | $5000 | $25 copay | $15 copay | $150 copay | deduct., then 35% co-insurance |
| Silver Low | $1000 | $2000 | $20 copay | $15 copay | deduct., then $100 copay | deduct., then no copay |
| Silver High | $0 | $2000 | $25 copay | $15 copay | $100 copay | $500 copay |
| Gold | $0 | None | $20 copay | $15 copay | $75 copay | $150 copay |

Table A.2: Detailed Plan Shares, among individuals who chose Neighborhood Health Plan.

| | Age Group | | | | | | |
|---|---|---|---|---|---|---|---|
| | 27-29 | 30-34 | 35-39 | 40-44 | 45-49 | 50-54 | 55+ |
| Bronze Low | 13.3% | 20.0% | 20.0% | 18.0% | 20.5% | 21.6% | 18.2% |
| Bronze Medium | 7.1% | 7.3% | 6.0% | 8.0% | 10.3% | 18.2% | 15.5% |
| Bronze High | 49.0% | 38.2% | 35.0% | 38.0% | 37.2% | 15.9% | 33.6% |
| Silver Low | 1.0% | 4.5% | 5.0% | 2.0% | 1.3% | 3.4% | 3.6% |
| Silver High | 19.4% | 19.1% | 19.0% | 22.0% | 21.8% | 28.4% | 25.5% |
| Gold | 10.2% | 10.9% | 15.0% | 12.0% | 9.0% | 12.5% | 3.6% |