Insuring Risks Large and Small: The Role of Preferences and Limited Consideration

Levon Barseghyan Isaac Cohen Yujie Feng*

February 16, 2026

Abstract

We develop and estimate a model of choice under risk with heterogeneous preferences and limited consideration, and apply it to U.S. auto-insurance choices across liability limits and collision/comprehensive deductibles. The framework combines a reference-dependent expected-utility specification, augmented with probability distortions, with a stochastic model of limited consideration in which alternatives enter the consideration set with alternative-specific probabilities. Two main findings emerge. First, a single preference framework can generate empirically plausible rankings across contexts with sharply different stakes and probabilities. Liability choices partially identify combinations of risk aversion and wealth exposure that imply extreme rankings over coverage limits and little sensitivity to in-sample price variation, while deductible choices are informative about probability distortions. Second, limited consideration, together with latent types, is central for matching observed choice variation and for avoiding biased preference inference. Latent types account for key features of the joint distribution across liability, collision, and comprehensive choices. Consideration depends on preferences: households are more likely to consider deductible pairs they prefer.

*We are grateful to many Cornell faculty, especially F. Molinari. T. O'Donoghue and M. Waldman, and to seminar participants at various institutions for helpful comments and suggestions. All remaining errors are our own. Barseghyan gratefully acknowledges financial support from NSF grants SES-1824448 and SES-2149374. *Author affiliations and contacts:* Barseghyan (corresponding author), Cornell University, lb247@cornell.edu; Cohen, Econic Partners, ic345@cornell.edu; Feng, Analysis Group, yf342@cornell.edu.

1 Introduction

Decision-making under risk is a foundational element of economics. The canonical framework—expected utility theory (EUT)—often underpins both positive and normative research. Predictions and welfare implications generated by models relying on this framework often vary dramatically with decision-makers' risk aversion, which in turn derives from the assumed degree of concavity of the Bernoulli (consumption-utility) function. The assumption of stable risk aversion—i.e., an invariant Bernoulli function across different choice situations—is implicit or explicit in most studies that discipline risk-aversion parameters using results from other studies. Yet both the canonical framework and preference stability have been challenged in experiments and in the field. To address the framework's shortcomings, a large behavioral and decision-theory literature has emerged. Loss aversion, disappointment aversion, probability weighting, reference dependence, and related departures from EUT are commonly employed, depending on the context, to generate a better fit to the data. While augmenting EUT with these features can be crucial for explaining high willingness to avoid risk in small-stakes contexts (Barseghyan et al. 2013), three key challenges remain. The first is Rabin's calibration argument: reconciling choices over low- and high-stakes risky prospects, since even modest curvature of the Bernoulli function can imply an implausibly high willingness to pay to avoid large-stakes lotteries (Rabin 2000). Two leading responses to this challenge are put forth in Kőszegi and Rabin (2007) and Chetty and Szeidl (2007). The former develops a reference-dependent utility framework that can accommodate seemingly inconsistent behavior across contexts through a dichotomy—choices over large stakes reflect modest curvature in the Bernoulli function, whereas choices over moderate or small stakes reflect reference dependence and loss aversion. The latter shows that, even with modest curvature of the Bernoulli function, consumption commitments can induce liquidity constraints that substantially increase a decision-maker's willingness to pay for insurance against small stakes without altering aversion to risk over large stakes. The second challenge is empirical evidence that, even in closely related field settings, preference-only accounts struggle to ex-

plain dominated choices (Abaluck and Gruber 2011; Barseghyan et al. 2016; Bhargava et al. 2017), instability in inferred preferences across contexts (Barseghyan et al. 2011; Einav et al. 2012), and the large effects of defaults and inertia (Honka 2014; Handel 2013). This challenge points to forces that shape which alternatives are considered and ultimately chosen, rather than only how they are evaluated. Understanding limited consideration—as a broad concept not confined to liquidity or other constraints—and how it interacts with risk preferences is essential for accurate inference, welfare analysis, and policy design. The third challenge is the substantial heterogeneity in choice under risk (Cohen and Einav 2007; Barseghyan et al. 2013). Identifying the sources of this heterogeneity may yield new insights into the stability of choice under risk across contexts, over time, and across the business cycle, and into the role of observable characteristics in shaping decision-makers' behavior.

Against this backdrop, we develop and estimate a model of choice under risk that allows for observed and unobserved heterogeneity in both preferences and limited consideration. We apply it to real-world data from contexts that differ sharply in stakes and probabilities. Our novel dataset comes from a U.S. auto-insurance provider and consists of repeated cross-sections spanning nearly two decades, with 322,448 household–policy observations covering three insurance lines—liability, collision, and comprehensive. Three features make the data well suited to our investigation. First, the structure of auto insurance contracts makes coverage choice largely a financial decision, unlike other domains in which non-financial factors—such as provider networks in health insurance—play an important role. Second, the dataset includes liability claims (and associated losses) as well as collision and comprehensive claims, which allows us to estimate individual-specific loss probabilities and the distribution of losses under liability coverage. While liability claims are relatively frequent, those exceeding the legally mandated minimum coverage limit of \$60,000—and hence relevant for household decisions—are extremely rare but potentially large. In contrast, collision and comprehensive claim probabilities are non-negligible, but the out-of-pocket cost in the event of an accident is the deductible. Third, we observe a measure of household wealth, a critical

variable because it affects both liability risk exposure and, potentially, risk preferences.

Our framework combines a preference model inspired by Kőszegi and Rabin (2007) with a flexible model of limited consideration. While Kőszegi and Rabin (2007) acknowledge the importance of, but abstract from, probability weighting (Kahneman and Tversky 1979), we incorporate it under a wider umbrella of probability distortions, as discussed in Barseghyan et al. (2013). To capture limited consideration, we adopt the model of Manski (1977) and Manzini and Mariotti (2014), in which each alternative is considered with an alternative-specific probability and the best among those considered is chosen. In addition to constraints, limited consideration may arise from behavioral processes such as elimination by aspects (Tversky 1972) or satisficing (Simon 1956; 1959),[1] or from reliance on subjective beliefs or third-party advice, leading decision-makers to exclude some options ex ante.

We present our findings in three steps. We begin with liability coverage. As a benchmark, we assume CARA utility. Provided households have non-trivial wealth at risk, even with modest risk aversion and no loss aversion or probability weighting, higher liability coverage yields higher expected utility. For instance, if households have at least $100,000 to lose and risk aversion exceeds 7×10^{-5}, then all households in our dataset should prefer more coverage to less.[2] Choices should then be insensitive to price variation. Conversely, if households have zero wealth exposure (or are risk-neutral and face actuarially unfair prices), they should opt for minimal coverage, and choices should again be uncorrelated with prices. Our estimates imply that (almost) all households have preferences consistent with a trivial ranking of coverage limits—from the highest to the lowest (the $\mathcal{R}$-types) or from the lowest to the highest (the $\mathcal{N}$-types)—and that liability choices are uncorrelated with prices.[3] Heterogeneity across decision-makers, observed and unobserved, is a crucial driver of liability choices. Roughly 70% of households are $\mathcal{R}$-types and 30% are $\mathcal{N}$-types. Conditional on type, the remaining

[1] Other behavioral phenomena such as extremeness aversion (Neumann et al. 2016) may also lead to limited consideration.

[2] This level of risk aversion is much lower than estimates derived from moderate-stakes insurance choices (Cohen and Einav 2007; Barseghyan et al. 2013).

[3] We label these types $\mathcal{R}$ and $\mathcal{N}$ based on our preferred interpretation ("Rabin-type" versus "nothing-to-lose"). We offer alternative interpretations in the paper.

variation reflects heterogeneity in consideration sets. Among observables, wealth stands out. Ceteris paribus, increasing households' wealth from its 5th to its 95th percentile raises the probability of choosing the two highest coverage levels, $500,000 and $1,000,000, from 46 to 65%, and lowers the probability of choosing the two lowest, $60,000 and $100,000, from 38 to 2%. We also find that higher unemployment is associated with less liability coverage, and that older households tend to purchase more of it.

We emphasize four points. First, accounting for limited consideration is essential for correctly inferring preferences. Under full consideration, our benchmark utility model (augmented with fixed effects and coverage-specific additive random noise, or with subjective beliefs) fits the data poorly and yields a misleading inference: the predicted sensitivity to price is strongly negative and the estimated preference distribution is biased. Second, while we establish the prevalence of households with a trivial ranking of coverages, their preferences are only partially identified. In the paper, we explore what types of preferences and consideration mechanisms are observationally equivalent to the ones we estimate. Third, wealth is a crucial determinant of household behavior, and we offer further evidence to support this finding. Fourth, in the liability domain, budget constraints or price-dependent consideration mechanisms are not the primary drivers of limited consideration.[4]

Next, we analyze collision and comprehensive deductible choices. We allow for both concave utility and non-linear probability distortions.[5] We highlight several peculiar features of the joint distribution of deductible choices, such as the fact that comprehensive deductibles are (almost) never larger than collision deductibles, and that some choices are never first best. While these patterns are difficult to reconcile under full consideration (Barseghyan et al. 2021a; Barseghyan and Molinari 2023), they are consistent with limited consideration, arising, for example, from interactions with insurance agents or from online sources.[6]

[4]Yet, as we detail in the paper, they are not inconsistent with the type of price-dependent consideration in Abaluck and Adams-Prassl (2021).

[5]In our setting, as shown in Barseghyan et al. (2013), probability distortions fully capture the combined effect of reference-dependent loss aversion, non-linear probability weighting as in Kahneman and Tversky (1979), and systematic deviations of subjective beliefs from actual loss probabilities.

[6]E.g., as Mortimer (2023) notes, in 2023 bankrate.com advised that "*having lower*" comprehensive than

We allow limited consideration to depend on preferences—an innovation made possible by observing choices across the two coverages—but otherwise preserve model parsimony. We find that, on average, households consider higher deductibles more frequently than lower ones, suggesting that liquidity constraints are unlikely to be the primary explanation for the overall level of willingness to pay for lower deductibles, but may still matter for cyclical variation in deductible choices, as discussed below. We find substantial probability overweighting, but only one quarter of households have a trivial ranking of deductibles, from the lowest to the highest. The resulting correlations between choices and base prices are -0.12 (collision) and -0.15 (comprehensive), closely matching their empirical counterparts. We find that households with higher willingness to pay for higher (or lower) coverages consider the respective coverages more often. This dependence of consideration on preferences alters choices for 7.5% of households and saves them, on average, \$31, more than a third of the total estimated average welfare cost of limited consideration of deductible pairs, \$78.

Turning to observables, the unemployment rate has a striking effect. An increase from the 5th to the 95th percentile in the contemporaneous unemployment rate lowers the probability of choosing higher deductible pairs (i.e., those comprised only of \$500 or \$1,000) by more than 15%. This finding is consistent with a liquidity channel: as the unemployment rate rises, for a non-trivial fraction of households, binding liquidity constraints shift consideration and choices from higher to lower deductibles. This channel, in turn, rationalizes a seeming puzzle: as unemployment rises, households choose less liability coverage but lower deductibles (i.e., more generous coverage). Higher wealth has the opposite effect, plausibly because for wealthier households liquidity constraints bind less often. Yet the strongest effect is due to age: an increase from the 5th to the 95th percentile in age lowers the probability of choosing

collision "*deductible is often logical*", presumably "*because collision claims tend to be more expensive than comprehensive claims, your collision deductible generally has a larger impact on your premium.*" Interpreted through an economist's lens, the advice yields an obvious fallacy: just because lowering the comprehensive deductible is cheaper than lowering the collision deductible, choosing a lower deductible in comprehensive than in collision is not necessarily optimal. The Wall Street Journal (WSJ June 17, 2024) recommends: "*only choose a deductible amount you can comfortably afford to pay from your savings if you have to file a claim. Otherwise, you risk being unable to pay the deductible after an accident and taking on debt to cover the bill.*"

higher deductible pairs by more than 25%. From a behavioral perspective, this pattern is consistent with Malmendier and Nagel (2011) on the role of past experiences in shaping behavior, as well as with status quo biases (Samuelson and Zeckhauser 1988; Kempf and Ruenzi 2021) and inertia in consideration sets (Klemperer 1995; Honka 2014; Handel 2013).[7]

Finally, we examine how choices across contexts relate to each other and identify a minimal modification to our model that accounts for this relation. Conditional on observables, loss probabilities, and prices, our benchmark framework predicts zero rank correlation between liability-limit and deductible choices. Yet we observe negative rank correlations, suggesting preference instability (Einav et al. 2012).[8] Such correlations can arise naturally if a household's latent type ($\mathcal{R}$ vs. $\mathcal{N}$) influences choices in both domains. We find that $\mathcal{R}$-types not only consider and prefer higher liability limits more often, but also consider and prefer the highest deductible more often and the lower deductibles less often; $\mathcal{N}$-types exhibit the opposite pattern. A natural interpretation is that $\mathcal{R}$-types have more wealth to protect—leading to higher liability coverage choices—while facing weaker liquidity constraints, which makes higher deductibles more attractive; $\mathcal{N}$-types behave as if their wealth at risk is small and liquidity is tight, pushing them toward lower liability limits and lower deductibles. This heterogeneity generates the negative rank correlation between choices across contexts with different stakes. Once accounted for, the correlations are negligible.

The rest of the paper is organized as follows. Section 2 describes our data. Sections 3 and 4 analyze liability limit choices and collision and comprehensive deductible choices, respectively. Section 5 discusses preference stability and what can be learned from choice patterns across all three coverages. Section 6 offers linkages to the literature and concluding remarks.

[7] We consider only the decision-maker's first choice with the company, so auto-renewals do not explain the patterns described, though the limited consideration mechanism could also account for them.

[8] Negative rank correlations between choices over coverages with different stakes were first documented in Barseghyan et al. (2018).

2 Data

The data are provided by a U.S. property-and-casualty insurance company. The main dataset documents households' choices over a bundle of auto insurance coverages: single-limit liability, collision, and comprehensive. We refer to these coverages as $\mathcal{S}$, $\mathcal{L}$, and $\mathcal{M}$, respectively, and use these symbols as superscripts to attribute variables to coverages. We focus on households' first-time purchases from the company,[9] yielding 322,448 observations over a nineteen-year period, from 2000 to 2018. The data include households' choices (Table 1), premiums paid, and the premiums associated with all available alternatives (Table 2). They also include a rich set of observables, $\mathbf{Z}$, including a proxy for households' total wealth—the value of their primary residence (Table 4),[10] which we condition on. We also condition on a fourth-degree polynomial in time to capture a secular shift toward higher liability limits and higher deductibles.

Claim probabilities are constructed using large auxiliary datasets. We model them assuming that, for each coverage, claims follow a Poisson process with arrival rate λ. We treat the λ's as latent random variables, the logarithms of which depend linearly on a set of observables, $\mathbf{X}^{\lambda}$, and an unobserved term. For each coverage, the unobserved term is distributed i.i.d. such that its exponent follows a gamma distribution with unit mean and unknown variance. We estimate these Poisson processes and, for each household in our main sample and for each coverage, construct the estimated λ based on observables and ex post claim realizations.[11] Table 3 reports summary statistics, and the results of the Poisson regressions are reported in Table A3 of the Appendix. Finally, we assume that losses, conditional on a claim, are drawn i.i.d. from a mixture of two exponential distributions fitted to match their empirical counterpart. The estimated function is reported in Equation A5 of the Appendix.

For each coverage, the premiums are a linear function of the household-specific base

[9] This avoids issues related to choice inertia (Handel 2013).

[10] This variable cannot be used by the company to set prices.

[11] This procedure is standard in the literature (see, e.g., Barseghyan et al. 2013). For applications in the actuarial literature, see Denuit et al. (2007), ch. 3.

price. For example, in liability: $p_j^{\mathcal{S}} = \delta_j^{\mathcal{S}} \bar{p}^{\mathcal{S}} + c^{\mathcal{S}}$, where $\bar{p}^{\mathcal{S}}$ is the base price. Base prices are determined by a set of variables, $\mathbf{X}^{\bar{p}}$, such as locality, time period, demographics, risk score, vehicle characteristics, and driving records. The pricing factors, $\delta_j^{\mathcal{S}}$, and the expense fee, $c^{\mathcal{S}}$, are invariant within each program offered by the company. We assume throughout the paper that, conditional on $\mathbf{Z}$, variation in base prices is exogenous to households' preferences. This assumption—standard in the literature (Cohen and Einav 2007; Barseghyan et al. 2013)—arises naturally because some determinants of insurance costs are orthogonal to risk preferences; that is, $\mathbf{X}^{\lambda} \backslash \mathbf{Z} \neq \varnothing$ and $\mathbf{X}^{\bar{p}} \backslash \mathbf{Z} \neq \varnothing$. These determinants include traffic patterns, meteorological conditions (e.g., the likelihood of severe weather events), wildlife (e.g., animals on roads), as well as local repair and transportation costs. The latter—along with many discounts—is also an example of a variable affecting prices but not loss probabilities, i.e., $\mathbf{X}^{\bar{p}} \backslash \mathbf{X}^{\lambda} \neq \varnothing$. As a result, although the correlations between base prices and claim probabilities are strong, they are far from perfect; see the last column of Table 3.[12]

The marginals of the choice distribution (Table 1) show that households rarely choose the highest or the lowest coverages: the two most common options—$300,000 and $500,000 liability limits, and $500 and $200 deductibles—account for at least 75% of all choices. The joint distribution shows that collision deductibles are almost always higher than comprehensive deductibles (Table 5) and that higher liability limits are associated with higher deductibles (i.e., lower coverage). We return to these patterns in Sections 4 and 5, respectively. As Table 2 shows, pricing across lines has a similar appearance and range, but this masks important differences in claim likelihood and severity. For liability, increasing the limit from $60,000 to $1 million raises the average premium by $104 (about 11 cents per additional $1,000 of coverage). For collision, lowering the deductible from $1,000 to $100 raises the average premium by $152 (or $169 per $1,000), while for comprehensive the corresponding increase is $57 (or $64 per $1,000).

[12] These numbers are similar if we do not use the ex post claim counts in estimating the λ's. As expected, the correlation is weakest for comprehensive coverage—the only coverage without experience rating.

Table 1: The choice distributions

Liability limits					
$60,000	$100,000	$200,000	$300,000	$500,000	$1,000,000
0.005	0.132	0.015	0.446	0.366	0.035
Collision deductibles					
$1000	$500	$250	$200	$100	
0.067	0.646	0.125	0.150	0.012	
Comprehensive deductibles					
$1000	$500	$250	$200	$100	$50
0.032	0.360	0.109	0.377	0.050	0.072

Notes: $50 deductible is not offered in collision. 0.09% (0.05%) of households have chosen collision (comprehensive) deductible of $750 or above $1000. These shares have been merged into the corresponding $1,000 bin.

Table 2: The average premia

Liability limits					
$60,000	$100,000	$200,000	$300,000	$500,000	$1,000,000
201	213	241	251	265	305
Collision deductibles					
$1000	$500	$250	$200	$100	
151	193	241	269	303	
Comprehensive deductibles					
$1000	$500	$250	$200	$100	$50
99	119	137	141	156	178

Notes: The averages are taken across the entire sample of 322,448.

Table 3: Base prices and estimated claim probabilities

	Base Prices				Loss Probabilities				Corr.
	Mean	SD	5th%	95th%	Mean	SD	5th%	95th%	$\rho(\lambda, \bar{p})$
Liability	251	126	112	477	0.0003	0.00012	0.0002	0.0005	0.50
Collision	269	164	102	562	0.058	0.020	0.033	0.096	0.38
Comprehensive	141	87	51	288	0.021	0.010	0.007	0.040	0.27

Notes: For liability, the claim probabilities correspond to the fraction of cases that result in a loss larger than the minimum mandated coverage of $60,000. Base prices are those of the $300,000 liability limit and the $200 deductible for collision and comprehensive. The last column is the correlation between the base price and the loss probability for each coverage.

Table 4: Selected demographic and other variables

	Mean	SD	5th%	95th%		Mean	SD	5th%	95th%
Demographics					**Wealth and Risk Score**				
Age	52.6	15.3	21	75	Home Value	238,000	133,000	95,000	480,000
Female	0.48				Risk Score	804	110	624	992
Married	0.68				**Local Economic Conditions**				
Single	0.21				GDP Growth	0.033	0.022	0.002	0.067
Young Driver	0.12				Unempl. Rate	0.057	0.016	0.038	0.089

Notes: The omitted groups for marital status are Widowed and Divorced. Young Driver indicates the presence of a driver below age 25 in the household. Home Value variables are rounded to the nearest $1000. The Risk Score takes into account financial variables including those in the insureds' credit history. The GDP growth and the Unemployment Rate, both in %, are at the county level.

Table 5: The joint distribution of deductibles, conditional on selected liability limits

	Conditional on liability choice $100,000				
			Collision		
Comprehensive	$100	$200	$250	$500	$1000
$50	0.02	0.05	0.02	0.02	0.00
$100	0.00	0.02	0.02	0.02	0.00
$200	0.00	0.15	0.06	0.21	0.01
$250	0.00	0.00	0.05	0.06	0.00
$500	0.00	0.00	0.00	0.27	0.01
$1000	0.00	0.00	0.00	0.00	0.01
	Conditional on liability choice $300,000				
			Collision		
Comprehensive	$100	$200	$250	$500	$1000
$50	0.01	0.03	0.02	0.02	0.00
$100	0.00	0.02	0.01	0.02	0.00
$200	0.00	0.12	0.06	0.22	0.01
$250	0.00	0.00	0.05	0.06	0.00
$500	0.00	0.00	0.00	0.32	0.02
$1000	0.00	0.00	0.00	0.00	0.02
	Conditional on liability choice $500,000				
			Collision		
Comprehensive	$100	$200	$250	$500	$1000
$50	0.01	0.02	0.01	0.02	0.00
$100	0.00	0.01	0.01	0.02	0.00
$200	0.00	0.08	0.05	0.20	0.01
$250	0.00	0.00	0.03	0.07	0.00
$500	0.00	0.00	0.00	0.38	0.03
$1000	0.00	0.00	0.00	0.00	0.05

Notes: The table reports the joint distribution of collision and comprehensive deductibles conditional on the three most common liability choices ($100,000, $300,000, and $500,000). Entries in each panel sum to 1. Overall, 0.19 (48.35) percent of households choose a lower(higher) collision deductible than comprehensive deductible. Unconditionally, the Spearman rank correlation between the liability choice and the collision deductible is -0.09; for the comprehensive deductible it is -0.12.

3 Liability Coverage Choices

The company offers s levels of liability coverage, which we refer to as the feasible set $\mathcal{S}$. Each coverage level $j \in \{1, 2, \ldots, s\}$ is characterized by a pair $(p_j^{\mathcal{S}}, \mathcal{S}_j)$, where $p_j^{\mathcal{S}}$ is the premium associated with coverage limit $\mathcal{S}_j$. Households are responsible for any claim amount exceeding their chosen coverage limit $\mathcal{S}_j$. The coverage levels are ordered such that $\mathcal{S}_j < \mathcal{S}_{j+1}$ and $p_j^{\mathcal{S}} < p_{j+1}^{\mathcal{S}}$. The premia are deterministic functions of the base price $\bar{p}^{\mathcal{S}}$ (when unambiguous, we

write $p_j^{\mathcal{S}}$ instead of $p_j^{\mathcal{S}}(\bar{p}^{\mathcal{S}})$). Conditional on observables $\mathbf{Z}$, both the base price and the claim probability $\lambda^{\mathcal{S}}$ vary across households. The probability of experiencing multiple claims is negligible and is assumed away. The associated loss l is distributed according to the function $\Phi(\cdot)$, which is invariant across households and $\mathbf{Z}$. Each household has rational expectations about $\lambda^{\mathcal{S}}$ and $\Phi(\cdot)$, and correctly believes that these are unaffected by its coverage choice.[13] Finally, each household evaluates and chooses from a subset of the available alternatives—the consideration set $\mathcal{K} \subseteq \mathcal{S}$, with $\mathcal{K} \neq \varnothing$. The consideration sets are unobserved and may vary across households.

3.1 Preferences: Expected utility theory with CARA

To start, we assume that the gain–loss utility of Kőszegi and Rabin (2007) does not affect rankings over lotteries with large stakes.[14] We also assume that decision makers (DMs) adhere to narrow bracketing (Read et al. 2000), i.e., their coverage-limit choice is made independently of other risks. We assume a constant absolute risk aversion (CARA) Bernoulli utility function. A DM's choices are governed by the risk aversion parameter ν and wealth-at-risk ω, defined as the maximum amount of wealth the DM believes they may lose when a claim results in excess liability. This may coincide with, or be lower than, the DM's total wealth. Both ν and ω are unobserved and may vary across DMs. We assume that they are drawn from an unknown distribution function $F(\cdot, \cdot \mid \mathbf{Z})$. To save space and streamline the exposition, we augment $\mathbf{Z}$ with $\lambda^{\mathcal{S}}$.[15] Conditional on observables, they are independent of

[13] That is, we allow for adverse selection but not for moral hazard—an approach common in the literature (e.g., Cohen and Einav 2007; Sydnor 2010; Barseghyan et al. 2013). This assumption is consistent with direct evidence in Jin (2023). While Jin and Vasserman (2021) present evidence that monitoring drivers affects their on-road behavior, they also assume that chosen coverage levels do not.

[14] Proposition 9 in Kőszegi and Rabin (2007) provides conditions under which this assumption is valid even when gain–loss utility affects rankings over lotteries with small stakes. Moreover, as we make clear below, our conclusions continue to hold even if gain–loss utility matters for large-stakes lotteries.

[15] From identification stand point, this is conservative, as we do not exploit variation in $\lambda^{\mathcal{S}}$. Including $\lambda^{\mathcal{S}}$ leaves our results unchanged, but it allows us to isolate the potential role of claim probabilities in shaping preferences and consideration. Recall that the $\lambda^{\mathcal{S}}$'s are constructed using ex post claim realizations and thus incorporate information about unobserved risk heterogeneity. In this sense, our approach is similar in spirit, though different in implementation, to Cohen and Einav (2007), who allow for dependence between unobserved risk heterogeneity and risk preferences.

the base price:

Assumption 1. $(\nu, \omega) \perp\!\!\!\perp \bar{p}^{\mathcal{S}} \mid \mathbf{Z}$.

The expected utility from choosing coverage level $\mathcal{S}_j$ is proportional to:

$$-E\mathcal{U}(p_j^{\mathcal{S}}, \mathcal{S}_j, \lambda^{\mathcal{S}}, \Phi \mid \nu, \omega) \propto \underbrace{\left[(1-\lambda^{\mathcal{S}}) + \lambda^{\mathcal{S}} \cdot \Phi(\mathcal{S}_j)\right] \cdot e^{-\nu(\omega - p_j^{\mathcal{S}})}}_{\text{term 1}} + \underbrace{\lambda^{\mathcal{S}} \int_{\mathcal{S}_j}^{\omega - p_j^{\mathcal{S}} + \mathcal{S}_j} e^{-\nu(\omega - p_j^{\mathcal{S}} - l + \mathcal{S}_j)} \, d\Phi(l)}_{\text{term 2}} + \underbrace{\lambda^{\mathcal{S}} \cdot \left(1 - \Phi(\omega - p_j^{\mathcal{S}} + \mathcal{S}_j)\right) \cdot e^{0}}_{\text{term 3}}. \tag{1}$$

The first term captures cases in which either no claim occurs or a claim occurs but its amount is below the coverage limit. In these cases, the DM's out-of-pocket expense is the premium paid. The second term integrates over states where the claim amount exceeds the coverage limit but the DM's total out-of-pocket expenses remain below ω. The final term captures cases in which the loss is so large that the DM loses their entire ω.[16] In what follows, for notational simplicity, we write

$$E\mathcal{U}_j(\bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}} \mid \nu, \omega) \equiv E\mathcal{U}(p_j^{\mathcal{S}}, \mathcal{S}_j, \lambda^{\mathcal{S}}, \Phi \mid \nu, \omega), \quad \forall j.$$

The preference structure above obeys a standard single crossing property, required later for semi–nonparametric identification (Barseghyan et al. 2021a):

Assumption 2 (Single Crossing)**.** *Given $\lambda^{\mathcal{S}}$ and ω, for any two coverage levels i and j and the base price $\bar{p}^{\mathcal{S}}$ there exists at most one $\nu \in (0, \infty)$ such that*

$$E\mathcal{U}_i(\bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}} \mid \nu, \omega) = E\mathcal{U}_j(\bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}} \mid \nu, \omega).$$

Given $\lambda^{\mathcal{S}}$ and ν, for any two coverage levels i and j and the base price $\bar{p}^{\mathcal{S}}$ there exists at most one $\omega \in (0, \infty)$ such that the equality above holds.

An immediate implication of this property is that, absent dominated coverage limits, for any positive ω, as risk aversion increases the DM prefers more coverage to less coverage.

[16] The DM's total wealth enters only through the coefficient of proportionality, which is invariant across j and therefore does not affect utility rankings.

That is, after some $\bar{\nu}$ their ranking of coverage limits becomes trivial: $\mathcal{S}_1 \prec \mathcal{S}_2 \prec \cdots \prec \mathcal{S}_s$. Conversely, as ω tends to zero, the DM's ranking of coverage limits takes the opposite direction: $\mathcal{S}_s \prec \mathcal{S}_{s-1} \prec \cdots \prec \mathcal{S}_1$.

3.2 Rabin's argument: Implications for empirical analysis

In settings like ours, the DM's wealth determines the stakes and, hence, their expected-utility rankings. As noted above, when $\omega = 0$, the DM has nothing to lose and therefore always prefers lower coverage limits to higher ones. With arbitrarily large ω, the opposite holds, echoing Rabin's critique (Rabin 2000): as ω increases and the stakes become larger, the certainty equivalent of any additional coverage increases.[17] The empirical relevance of Rabin's argument, however, is much broader. To illustrate it, we introduce two concepts.

Definition 1.

(ν, ω) is an $\mathcal{R}$-pair if for all observed $(\bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}})$ expected utility is strictly increasing in the coverage limit:

$$E\mathcal{U}_i(\bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}} \mid \nu, \omega) < E\mathcal{U}_j(\bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}} \mid \nu, \omega), \quad \forall i, j,\ i < j.$$

(ν, ω) is an $\mathcal{N}$-pair if for all observed $(\bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}})$ expected utility is strictly decreasing in the coverage limit:

$$E\mathcal{U}_i(\bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}} \mid \nu, \omega) > E\mathcal{U}_j(\bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}} \mid \nu, \omega), \quad \forall i, j,\ i < j.$$

For each pair (ν, ω), we classify it as an $\mathcal{R}$-pair (respectively, an $\mathcal{N}$-pair) by computing each household's expected-utility ranking and checking whether at least 99.75% of these rankings conform to the corresponding type.[18] Figure 1 maps the (ν, ω) space into regions that induce $\mathcal{R}$-type rankings, $\mathcal{N}$-type rankings, and non-trivial rankings in the observed data. As

[17] In the limit as $\omega \to \infty$, if, for example, the loss distribution $\Phi(\cdot)$ is exponential with parameter ρ, then the certainty equivalent of any (finite) coverage limit is $-\infty$ whenever $\rho > \nu$. Mirroring Rabin's calibration result, the DM would then be willing to pay an arbitrarily high price to avoid the remaining exposure to excess liability.

[18] The choice of labels reflects—for the $\mathcal{R}$-pairs—the logic of Rabin's calibration argument, and—for the $\mathcal{N}$-pairs—their "nothing-to-lose" interpretation, i.e., behaving as if wealth at risk were negligible.

Figure 1: R-types and N-types

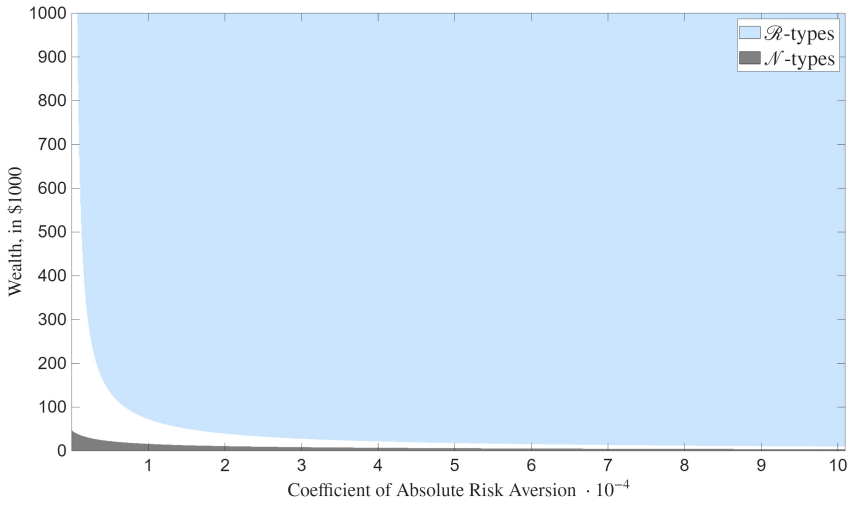


Figure 1 shows, the parameter space is dominated by $\mathcal{R}$-pairs, which populate the top-right area.[19] The $\mathcal{N}$-pairs populate the bottom-left area, and only (ν, ω) pairs in the white area induce a non-trivial ranking of coverage limits at some observed $(\bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}})$. The patterns in Figure 1 are a direct implication of Rabin's critique. Despite tiny claim probabilities, the loss distribution resulting from a claim is sufficiently fat-tailed that, with any sizable wealth at risk, even a small curvature of the Bernoulli function implies a strict preference for higher coverage limits. Put differently, the willingness to pay for higher coverage limits far exceeds not only the actuarially fair value but also the maximum observed price difference between higher and lower coverage limits. The remark below distills this observation into empirically testable implications:

Remark 1. *Suppose Assumption 1 holds. If all DMs are endowed with $\mathcal{R}$- or $\mathcal{N}$-pairs, then under full consideration:*

1. *All DMs prefer the highest (or the lowest) coverage limit offered; no other coverage limits are chosen.*

[19] The figure is constructed by computing, for each household, the expected utility of each coverage limit at the household's observed base price and claim probability. For numerical stability, we truncate ν below at 10^{-8}.

2. *The correlation between DMs' coverage limit choices and base prices, conditional on* $\mathbf{Z}$*, is zero.*

3.3 Limited consideration

For the purposes of this paper, limited consideration refers to all factors and features that do not affect how alternatives are evaluated but do affect which alternatives are evaluated and chosen from. As a benchmark, we impose independence—conditional on observables—between consideration and base prices, as well as between consideration and preferences. These conditions are required for semi–nonparametric identification in settings like ours (Barseghyan et al. 2021a):

Assumption 3.

3(a) Conditional on $\mathbf{Z}$*, the distribution of the consideration sets is independent of* $\bar{p}^{\mathcal{S}}$*.*

3(b) Conditional on $\mathbf{Z}$*, the distribution of the consideration sets is independent of preference parameters* (ν, ω)*.*

Under limited consideration, the first part of Remark 1 no longer holds, since limited consideration can force DMs to choose intermediate coverage limits even when they prefer the extremes. However, because limited consideration does not alter the expected-utility rankings, all DMs with $\mathcal{R}$-pairs would choose the highest coverage limit they consider, regardless of the base price, and all DMs with $\mathcal{N}$-pairs would choose the lowest coverage limit they consider. Since the consideration mechanism is independent of the base price, the second part of Remark 1 still holds.

Remark 2. *Suppose Assumptions 1–3(a) hold. If all DMs are endowed with* $\mathcal{R}$*- or* $\mathcal{N}$*-pairs, then under limited consideration:*

1. *Each DM prefers the highest (or the lowest) coverage limit they consider, and all coverage limits may be chosen.*

2. *The correlation between DMs' coverage limit choices and base prices, conditional on* $\mathbf{Z}$*, is zero.*

3.4 Completing the model

To complete the model, we take two additional steps. First, we specify a particular consideration mechanism. As a benchmark, we adopt the Alternative-specific Random Consideration (ARC) model, which is commonly used in the literature (Manski 1977; Manzini and Mariotti 2014; Abaluck and Adams-Prassl 2021; Barseghyan et al. 2021a). Under the ARC model, each alternative j, $1 \leqslant j \leqslant s$, is considered with probability $\psi_j(\mathbf{Z})$, which implies the following probability of forming consideration set $\mathcal{K}$:

$$\Pr(\mathcal{K} \mid \mathbf{Z}) = \prod_{j \in \mathcal{K}} \psi_j(\mathbf{Z}) \prod_{k \notin \mathcal{K}} \big(1 - \psi_k(\mathbf{Z})\big).$$

While the ARC model offers clear merits—parsimony, flexibility, and ease of implementation—it is only one of many possible specifications.[20] Second, we must specify wealth-at-risk, ω. As noted above, ω is potentially different from households' total wealth. (The latter may influence the stakes through ω, and it may also directly affect risk preferences and the consideration mechanism.) While we conduct an extensive sensitivity analysis and explore a variety of alternative specifications in Section 3.10, our benchmark specification sets $\omega =$ \$100,000 for all households. With this structure in place, model identification reduces to the semi–nonparametric identification of the distribution function of the risk aversion parameter, $F(\nu \mid \mathbf{Z})$, and the consideration probabilities $\{\psi_j(\mathbf{Z})\}_{j=1}^{s}$. With an additional minimal restriction, identification follows from Barseghyan et al. (2021a), provided that $\bar{p}^{\mathcal{S}}$

[20] A technical issue in the ARC model is that $\mathcal{K} = \varnothing$ can occur with probability $\prod_{k \in \mathcal{S}} \big(1 - \psi_k(\mathbf{Z})\big)$. If an outside option is available, it is chosen in this case. If no outside option is available, the model must be completed with an additional assumption—a completion rule. For example, one can rule out empty consideration sets by assuming that at least one alternative is always considered. Alternatively, one can assume that if $\mathcal{K} = \varnothing$, the DM chooses uniformly at random from $\mathcal{S}$. A third approach is to assume that if the realized consideration set is empty, the DM redraws a new one using the same procedure; this is equivalent to renormalizing $\Pr(\mathcal{K} \mid \mathbf{Z})$ over $\mathcal{K} \neq \varnothing$ by the factor $\Theta(\mathbf{Z}) = \big[1 - \prod_{k \in \mathcal{S}} \big(1 - \psi_k(\mathbf{Z})\big)\big]^{-1}$. To avoid an ad hoc choice of an always-considered alternative, we adopt the latter approach. In our applications, the second and third approaches yield very similar results.

exhibits sufficient variation. We elaborate on this point in Section 3.8.

3.5 Likelihood

In our model, conditional on preferences (ν, ω) and given $(\bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}})$, coverage limit j is chosen if and only if j is considered and no coverage limit that yields higher expected utility is considered.[21] Let $\mathcal{B}_j(\bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}} \mid \nu, \omega)$ be the set of coverage limits that strictly dominate j:

$$\mathcal{B}_j(\bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}} \mid \nu, \omega) \equiv \left\{k \in \mathcal{S} : \; E\mathcal{U}_k(\bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}} \mid \nu, \omega) > E\mathcal{U}_j(\bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}} \mid \nu, \omega)\right\}.$$

Then, the conditional choice probability of j is

$$\Pr\big(j \mid \bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}}; \mathbf{Z}\big) = \Theta(\mathbf{Z}) \cdot \psi_j(\mathbf{Z}) \int \prod_{k \in \mathcal{B}_j(\bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}} \mid \nu, \omega)} \big(1 - \psi_k(\mathbf{Z})\big)\, dF(\nu \mid \mathbf{Z}),$$

where $\Theta(\mathbf{Z})$ is defined as in Footnote 20. Since our model can rationalize any choice, it can be estimated as is, and in particular does not require logit-type errors (see more on this point in Sections 3.11 and 3.12). The absence of the latter allows for a major computational advantage: at any (ν, ω) and for any household, the expected-utility rankings are computed only once to determine the corresponding $\mathcal{B}_j(\bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}} \mid \nu, \omega)$.

3.6 Estimation

We model the distribution of the risk aversion parameter as follows. We restrict its support to an interval where the data are informative. Specifically, we set support $F(\nu \mid \mathbf{Z}) \subseteq [0, \bar{\nu}(\omega)]$, where $\bar{\nu}(\omega)$ is the smallest ν such that (ν, ω) is an $\mathcal{R}$-pair. Since all $\nu \geqslant \bar{\nu}(\omega)$ are observationally equivalent to $\bar{\nu}(\omega)$, the only quantity that estimation can recover is the share of DMs whose risk aversion weakly exceeds $\bar{\nu}(\omega)$.

While $F(\cdot)$ is identified nonparametrically, for estimation purposes we parameterize $F(\cdot)$ as a mixture between a Beta distribution $\mathbf{B}(\cdot;\cdot)$ and a point mass at $\bar{\nu}(\omega)$:

$$F(\nu \mid \mathbf{Z}) = \alpha_\nu(\mathbf{Z}) \cdot \mathbf{B}(\nu; \beta_1, \beta_2) + \big(1 - \alpha_\nu(\mathbf{Z})\big) \cdot \mathbf{1}\big(\nu = \bar{\nu}(\omega)\big),$$

[21] For a detailed discussion, see Section III of Barseghyan et al. (2021a).

where the distribution parameters (β_1, β_2) and the share $\alpha_\nu(\mathbf{Z})$ are unknown. The latter is parameterized as a logistic transformation of a linear function:

$$\alpha_\nu(\mathbf{Z}) = \frac{1}{1 + \gamma_{\nu,0} \exp(\gamma_\nu \mathbf{Z})}. \tag{2}$$

The limited consideration probabilities, $\{\psi_j(\mathbf{Z})\}_{j=1}^s$, are parameterized as:

$$\psi_j(\mathbf{Z}) = \alpha_\psi(\mathbf{Z}) \cdot \underline{\psi}_j + \big(1 - \alpha_\psi(\mathbf{Z})\big) \cdot \bar{\psi}_j,$$

where

$$\alpha_\psi(\mathbf{Z}) = \frac{1}{1 + \exp(\gamma_\psi \mathbf{Z})}. \tag{3}$$

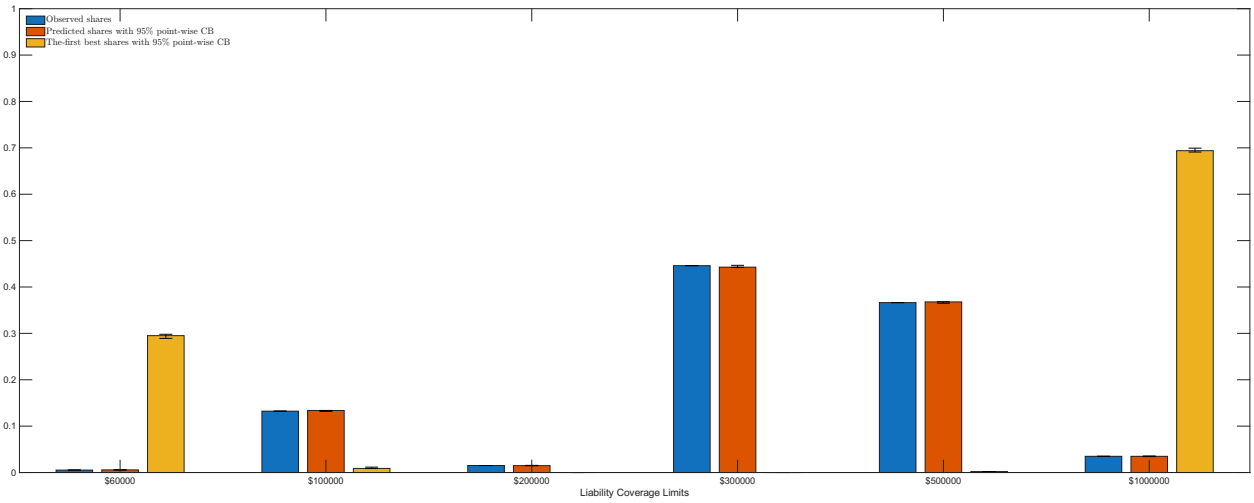
We impose no restrictions on $\{\underline{\psi}_j, \bar{\psi}_j\}_{j=1}^s$ beyond requiring that they lie in $[0, 1]$. Each $\psi_j(\mathbf{Z})$ is a linear function of $\alpha_\psi(\mathbf{Z})$, with an alternative-specific slope and intercept. The substantive restriction, imposed to ensure parsimony and computational feasibility, is that $\alpha_\psi(\mathbf{Z})$—more specifically, γ_ψ—is invariant across j.[22] To see the flexibility afforded by this parameterization, suppose that $\underline{\psi}_j = \bar{\psi}_j$, $\underline{\psi}_k > \bar{\psi}_k$, and $\underline{\psi}_l < \bar{\psi}_l$. Then variation in $\mathbf{Z}$ would have no effect on the consideration probability of coverage limit j, while it would have opposite effects on the consideration probabilities of coverage limits k and l, and hence induce a negative correlation between them.

With this structure in place, estimation reduces to finding the parameter vector

$$\{\beta_1, \beta_2, \gamma_\nu; \{\underline{\psi}_j, \bar{\psi}_j\}_{j=1}^s, \gamma_\psi\}$$

that maximizes the quasi-log-likelihood. After presenting our estimation results, we provide a detailed discussion of which features of the data yield identification, how these features pin down the consideration probabilities, and what alternative consideration mechanisms would be consistent with our findings.

Figure 2: Choice distributions



3.7 Benchmark results: Choices

Figure 2 shows the empirical distribution of coverage limit choices, the estimated distribution, and the first-best distribution—that is, the choice distribution that would arise under full consideration. The estimated and empirical distributions are nearly identical, with total variation of 0.0031 ($CI_{95\%} = [0.0004, 0.0036]$).[23] Overall, essentially all households have preference rankings consistent with either $\mathcal{R}$-pairs (about 70%) or $\mathcal{N}$-pairs (30%). These patterns align with Remarks 1 and 2. Under full consideration, one would observe a choice distribution almost entirely concentrated at the lowest and highest coverage limits. When at least one of these extremes is not considered, choices may shift to intermediate limits. The estimated (conditional on $\mathbf{Z}$ and $\lambda^{\mathcal{S}}$) correlation between coverage limit choices and the base price is -0.003, while in the data it is 0.003.[24] In contrast, the correlation between

[22] As specified, the model has 47 parameters. Allowing γ_ψ to vary with j would add 64 parameters (for a total of 111), making estimation infeasible.

[23] The confidence intervals are reversed-quantile (basic bootstrap) intervals (Hansen 2022), computed using a 66% subsampling procedure (Politis et al. 1999) with 750 replications. All estimates are bias-corrected, except those reported alongside their empirical counterparts, which we leave uncorrected to facilitate exact comparisons. Bias-correcting these estimates yields results that are nearly indistinguishable from those reported here; the bias-corrected versions are available from the authors upon request.

[24] Throughout, we compute such correlations using the rank—i.e., the menu position—of the chosen coverage, rather than its dollar amount. We caution that this near-zero correlation between coverage choices and the within-firm base price does not imply that consumers ignore price. Consumers may still be price sensitive when selecting insurers, which—together with substantial rate regulation—disciplines markups. More generally, the correlation between choices and the base price has no structural interpretation and, in

coverage-limit choices and $\lambda^{\mathcal{S}}$ (conditional on $\mathbf{Z}$ and $\bar{p}^{\mathcal{S}}$) is 0.056 in the data and 0.060 in the model ($CI_{95\%} = [0.059, 0.067]$). Higher $\lambda^{\mathcal{S}}$ is associated with higher coverage choices through a higher share of $\mathcal{R}$-types, echoing the findings of Cohen and Einav (2007). Absent this dependence, choices would not respond to variation in loss probabilities under extreme preference rankings. This observation, in turn, motivates a more nuanced question about the overall role of prices and how they affect choices, which we address in Section 3.13.

In sum, the estimation effectively partitions households into $\mathcal{R}$-types and $\mathcal{N}$-types, while using limited consideration to rationalize intermediate coverage choices. Finally, the prevalence of extreme types implies that preferences are only partially identified. The behavior of $\mathcal{R}$-types is consistent with any pair (ν, ω) such that $\nu \geqslant \bar{\nu}(\omega)$ (i.e., any point in the upper-right region of Figure A10). Similarly, for any ν, the behavior of $\mathcal{N}$-types can be replicated with sufficiently low ω (i.e., by selecting a pair from the lower-left region of Figure A10).

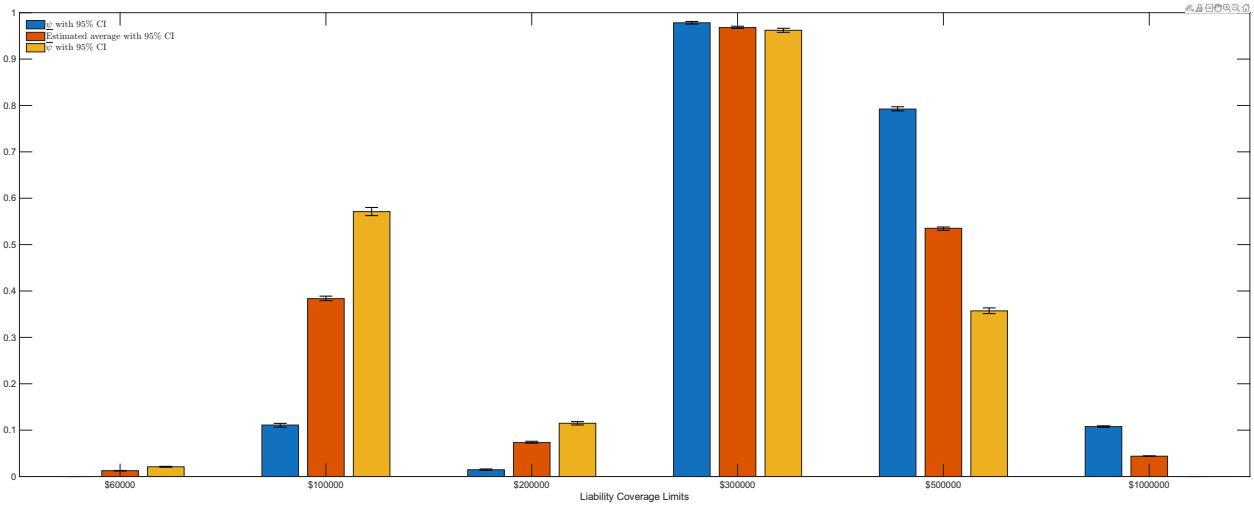
3.8 Benchmark results: Types and Consideration

Figure 3 reports estimated consideration probabilities: $\underline{\psi}$, $\overline{\psi}$, and $\psi(\mathbf{Z})$ evaluated at the mean values of the observables. The \$300,000 limit is almost always considered, regardless of observables, since $\underline{\psi}$ is close to $\overline{\psi}$. In contrast, all other limits are considered far less often, with probabilities that depend on observables; the lowest and highest coverage levels have the smallest probability of entering the consideration set.

To interpret these results and their limits, we emphasize two points. First, the prevalence of $\mathcal{R}$-types and $\mathcal{N}$-types stems from the negligible correlation between coverage-limit choices and the base price. If preferences generated a non-trivial mass of intermediate choices, they would also imply a strong negative correlation between the base price and coverage-limit choices, even under limited consideration. Second, while the prevalence of extreme preference types is robust, the exact interpretation of $\mathcal{R}$-types and $\mathcal{N}$-types is influenced

particular, is not a measure of price elasticity.

Figure 3: Consideration probabilities



by our choice of consideration mechanism. To see why, consider the identification argument underlying our estimation (for a detailed discussion and proofs, see Barseghyan et al. 2021a). For a given ranking of alternatives, the probability that coverage limit j is chosen equals the probability that j is considered and that no alternative that yields higher expected utility is considered. With $\mathcal{N}$-type preferences, whose share is $\alpha_\nu(\mathbf{Z})$, coverage limit j is chosen if it is considered and lower coverage limits are not. With $\mathcal{R}$-type preferences, coverage limit j is chosen if it is considered and higher coverage limits are not. Consequently, we arrive at the key identifying equation:[25]

$$\Pr(j \mid \mathbf{Z}) = \Theta(\mathbf{Z})\Big[\alpha_\nu(\mathbf{Z}) \cdot \psi_j(\mathbf{Z}) \prod_{k<j}\big(1-\psi_k(\mathbf{Z})\big) + \big(1-\alpha_\nu(\mathbf{Z})\big) \cdot \psi_j(\mathbf{Z}) \prod_{k>j}\big(1-\psi_k(\mathbf{Z})\big)\Big]. \quad (4)$$

Consider instead a model in which all households have $\mathcal{R}$-type preferences, but there are two ARC consideration types. The first type is exactly as estimated above and has share $\big(1-\alpha_\nu(\mathbf{Z})\big)$. The other type has consideration probabilities $\phi_j(\mathbf{Z})$, constructed recursively

[25]The choice probabilities $\Pr(j \mid \mathbf{Z})$ are observed in the data. For a fixed $\alpha_\nu(\mathbf{Z})$, and modulo the normalization issues discussed in Footnote 20, this delivers full nonparametric identification of $\{\psi_j(\mathbf{Z})\}_{j=1}^s$, as in Barseghyan et al. (2021a). Joint identification of $\alpha_\nu(\mathbf{Z})$ and $\{\psi_j(\mathbf{Z})\}_{j=1}^s$ requires one additional restriction (e.g., a scale normalization or a functional-form restriction) imposed by parameterization; this restriction replaces the large-support assumption on the special regressor used in Barseghyan et al. (2021a).

as follows:

$$\phi_s(\mathbf{Z}) = \frac{1}{\alpha_\nu(\mathbf{Z})}\Big[\frac{\Pr(s \mid \mathbf{Z})}{\Theta(\mathbf{Z})} - \big(1-\alpha_\nu(\mathbf{Z})\big)\psi_s(\mathbf{Z})\Big],$$

$$\phi_j(\mathbf{Z}) = \frac{1}{\alpha_\nu(\mathbf{Z})}\left[\frac{\Pr(j \mid \mathbf{Z})}{\Theta(\mathbf{Z})} - \big(1-\alpha_\nu(\mathbf{Z})\big)\psi_j(\mathbf{Z})\prod_{k>j}\big(1-\psi_k(\mathbf{Z})\big)\right] \cdot \frac{1}{\prod_{k>j}\big(1-\phi_k(\mathbf{Z})\big)}, \quad \forall\, j < s.$$

By construction, this model with two consideration types generates exactly the same fit as the model above with two preference types. Hence, the interpretation of $\mathcal{R}$- and $\mathcal{N}$-types can vary with the assumed flexibility of the consideration mechanism.

Finally, when households have extreme preferences, the ARC consideration mechanism admits alternative interpretations. E.g., with $\mathcal{R}$-type preferences, it is observationally equivalent to a mechanism in which consideration sets are truncated from above: the consideration set equals $\{\mathcal{S}_1, \ldots, \mathcal{S}_j\}$ with probability $\Theta(\mathbf{Z})\psi_j(\mathbf{Z})\prod_{k>j}\big(1-\psi_k(\mathbf{Z})\big)$. Indeed, with $\mathcal{R}$-type preferences, the choice is the highest coverage limit considered, so $\mathcal{S}_j$ is chosen whenever higher coverages are not considered; whether lower coverages are considered probabilistically or with certainty is irrelevant. Thus, the estimated model is also equivalent to one in which all households are $\mathcal{R}$-types and there are two consideration types that differ only in their propensity to truncate the consideration set from above, e.g., because of their subjective beliefs, as we discuss in Section 3.11, or because of additional supplemental liability coverage.[26]

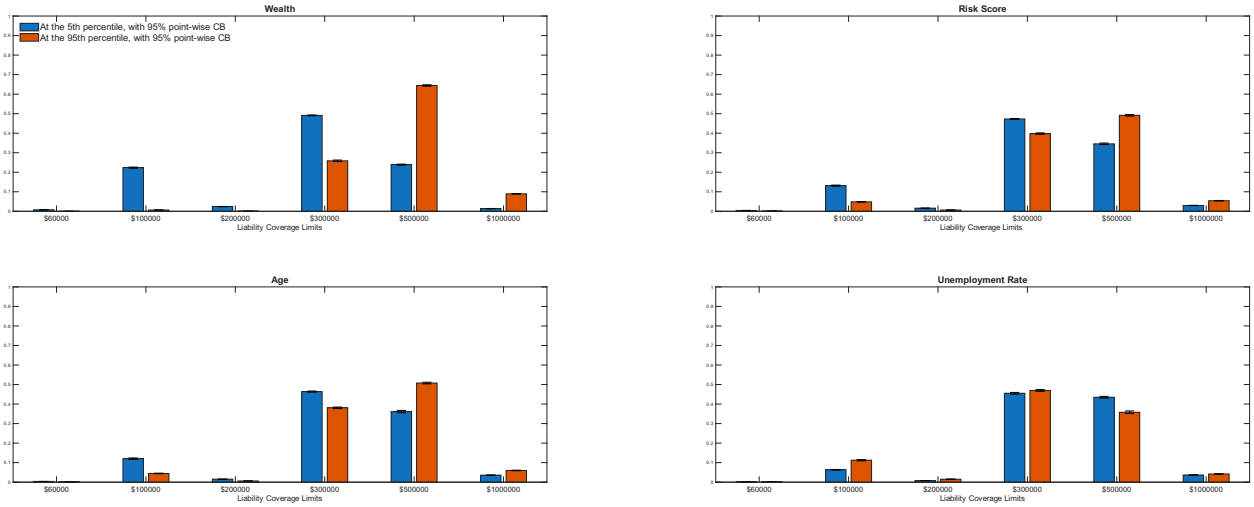
In sum, liability choices are explained by limited consideration and extreme preference types. Whether there are two extreme types or just one depends on the flexibility of the consideration mechanism. Given the type(s), the consideration mechanism is identified via Equation 4.

3.9 Benchmark results: The role of observables

The first panel of Figure 4 illustrates how household wealth shapes coverage choices. (Estimates of γ_ν and γ_ψ appear in Table A1 of the Appendix.) Holding all other observables

[26] For a detailed discussion of potential issues with supplemental liability coverage see Appendix A2.

Figure 4: The role of observables



at their averages, households at the 95th percentile of wealth almost never select limits below $300,000, and more than three-fifths of their choices exceed this threshold. In contrast, households at the 5th percentile rarely choose limits above $300,000, with over two-fifths of their choices falling below it. This pattern reflects two effects: wealthier households are more likely to be $\mathcal{R}$-type and more likely to consider higher limits while excluding lower ones. Higher risk scores similarly shift choices toward higher limits (see the second panel of the figure). As shown in the third panel, older households prefer more extensive coverage: shifting the age of the primary driver from the 5th to the 95th percentile generates about a 15-percentage-point reallocation from lower to higher coverage limits (see the fourth panel of the figure). Finally, local economic conditions influence coverage only marginally: moving from the 5th to the 95th percentile of the unemployment rate reallocates roughly 8 percentage points from the two highest coverage limits to the rest. The effects of the remaining variables are shown in Figure A1 of the Appendix.

3.10 The role of wealth and other expected utility formulations

The importance of total wealth that we uncover invites further scrutiny. First, we establish robustness with respect to the modeling of wealth-at-risk, ω. In the benchmark case, it is set to \$100,000. Resetting it to \$50,000 or to \$150,000 (with corresponding changes in the support of the risk aversion parameter) leaves the estimation results virtually unchanged. Second, we investigate whether dependence between ω and total wealth could alter our results. To do so, we model ω as a power function of the wealth proxy, home value, v_h:

$$\omega(v_h) = a_0\Big[(1-a_1) + a_1 \cdot \Big(\frac{v_h}{\bar{v}_h}\Big)^{a_2}\Big],$$

where $\bar{v}_h$ is the average home value. We entertain various combinations of $\{a_1, a_2\}$, with the restriction that $\omega(\bar{v}_h) = \$100{,}000$ and an appropriate adjustment of the support of the risk aversion parameter. The results are again virtually unchanged.[27] The reason, ex post, is transparent. Changing ω alters expected utility. Because DMs must have extreme preferences to generate a negligible correlation between choices and the base price, allowing ω to depend on v_h is redundant: v_h already affects the mix of preference types via α_ν (Equation 2) and affects consideration via α_ψ (Equation 3).

A more general point concerns the expected-utility construct. As long as the family of Bernoulli functions is flexible enough to allow for both the $\mathcal{R}$-types and the $\mathcal{N}$-types, our results stand. For example, gain–loss utility of Kőszegi and Rabin (2007) may be non-trivial, or ω and ν could be correlated. Moreover, the $\mathcal{R}$-types could overweight claim probabilities, becoming even more risk averse, and the $\mathcal{N}$-types could underweight them—without changing any of the model's predictions.[28]

[27] In total, twelve parameterizations are considered: $a_0 = \$100{,}000$; $a_1 \in \{0, \frac{1}{3}, \frac{1}{2}, \frac{2}{3}\}$; and $a_2 \in \{\frac{1}{2}, 1, 2\}$.

[28] In fact, with expected utility dependent on a number of unobserved parameters, one can construct the $\mathcal{R}$-types and the $\mathcal{N}$-types in a variety of ways, all of which would be consistent with our findings.

3.11 Robustness: Beliefs about losses

Our benchmark analysis assumes that DMs know the claim probabilities and the loss distribution, $\{\lambda^{\mathcal{S}}, \Phi(\cdot)\}$. Here, we address two questions: (i) what structure of subjective beliefs could be consistent with our findings yet appear empirically as limited consideration, and (ii) to what extent common subjective beliefs over $\{\lambda^{\mathcal{S}}, \Phi(\cdot)\}$ can explain DMs' choices? For example, an $\mathcal{R}$-type choosing a \$500,000 coverage limit could either be failing to consider the \$1,000,000 limit or believing that $\Phi(\cdot)$ is truncated (and reweighted) so that no loss above \$500,000 is possible.[29] In the extreme, a belief that a claim is certain and will equal exactly \$500,000 makes that limit optimal for any weakly risk-averse agent with $\omega > p^{\mathcal{S}}_{500,000} - p^{\mathcal{S}}_{60,000}$. By this logic, the choice of any single household in our data could be rationalized by some set of subjective beliefs. Though with a different source of unobserved heterogeneity, such behavior would appear identical to limited consideration and hence be indistinguishable from the mechanism discussed in the final paragraph of Section 3.8. Therefore, absent restrictions on beliefs (e.g., commonality across DMs conditional on observables, parametric structure, or limited heterogeneity), beliefs are not separately identified from limited consideration.

The question, then, is whether—absent (or with restricted) unobserved heterogeneity—subjective beliefs can allow a standard full-consideration framework to match the data. That is, suppose DMs have potentially non-extreme preferences and consider all coverage limits. In addition, suppose that, conditional on observables, they share common but otherwise arbitrary subjective beliefs about the loss distribution. Can a model with this structure explain the choices in our sample? The answer is no. Under full consideration, such beliefs can make intermediate coverage limits optimal only by altering expected-utility rankings, which would imply a strong negative correlation between choices and base prices—a pattern inconsistent with the empirical evidence. Intuitively, under full consideration, making an intermediate limit optimal requires beliefs that compress valuation differences across lim-

[29] One would also need a belief system consistent with why a \$1,000,000 limit is offered at all, but that is beyond the scope of this paper.

its; combined with observed premium variation, this pushes choices to move systematically against the base price, generating a counterfactually strong negative correlation.

As a quantitative illustration, consider a model in which all DMs are characterized by a pair (ν, ω) and subjective beliefs $\{\mu_j(\mathbf{Z})\}_{j=1}^s$, where $\mu_j(\mathbf{Z})$ is the probability that the loss amount exceeds $\omega + \mathcal{S}_j$ (with $\mu_1(\mathbf{Z}) = 1$). In this exercise, we allow flexible subjective beliefs about the loss distribution conditional on a claim (captured by $\mu_j(\mathbf{Z})$), while maintaining the claim probability $\lambda^{\mathcal{S}}$ as in the benchmark. By construction, $\mu_j(\mathbf{Z})$ is weakly decreasing in j; thus, $\mu_j(\mathbf{Z})$ summarizes the perceived probability of an "excess-loss" event that triggers the maximum out-of-pocket expense ω under limit j. The expected utility from choosing limit j is

$$-E\mathcal{U}(p_j^{\mathcal{S}}, \mathcal{S}_j, \lambda^{\mathcal{S}}, \{\mu_j(\mathbf{Z})\}_{j=1}^s \mid \nu, \omega) \propto \left(1 - \lambda^{\mathcal{S}} + \lambda^{\mathcal{S}}\left(1 - \mu_j(\mathbf{Z})\right)\right) e^{\nu p_j^{\mathcal{S}}} + \lambda^{\mathcal{S}} \mu_j(\mathbf{Z})\, e^{\nu(p_j^{\mathcal{S}} + \omega)}.$$

To make this model as flexible as the benchmark model, we parameterize $\{\mu_j(\mathbf{Z})\}_{j=1,\ldots,s}$ as

$$\mu_j(\mathbf{Z}) = \alpha_\mu(\mathbf{Z})\, \underline{\mu}_j + (1 - \alpha_\mu(\mathbf{Z}))\, \bar{\mu}_j,$$

where $\alpha_\mu(\mathbf{Z}) = \frac{1}{1+\exp(\gamma_\mu \mathbf{Z})}$. Risk aversion is parameterized as before, and ω is set to the benchmark value of \$100,000. Transforming $E\mathcal{U}$ into certainty equivalents gives

$$-CE(p_j^{\mathcal{S}}, \mathcal{S}_j, \lambda^{\mathcal{S}}, \{\mu_j(\mathbf{Z})\}_{j=1}^s \mid \nu, \omega) = p_j^{\mathcal{S}} + \frac{1}{\nu} \log\left[1 - \lambda^{\mathcal{S}} + \lambda^{\mathcal{S}}\left(1 - \mu_j(\mathbf{Z})\right) + \lambda^{\mathcal{S}} \mu_j(\mathbf{Z})\, e^{\nu\omega}\right],$$

and, to make estimation feasible, we add i.i.d. extreme-value type 1 (EVT1) errors with an unknown dispersion parameter. Modulo normalizations, this model has the same number of parameters as our benchmark model, plus one dispersion parameter. The fundamental difference between this model and the benchmark is the source of unobserved heterogeneity: in the benchmark model it is preferences and the consideration set; here it is preferences and the realizations of the McFadden error, which are i.i.d.

This model reproduces the empirical distribution of coverage limits reasonably well (total variation = 0.038, see Figure A3 of the Appendix), though not as well as our benchmark model (total variation = 0.0031). However, it predicts a strong negative correlation between

coverage choices and base prices (-0.16). The model also overshoots the correlation between the loss probability $\lambda^{\mathcal{S}}$ and choices: 0.12 versus the empirical value of 0.056. As anticipated, this discrepancy arises because, under full consideration, matching intermediate choices requires non-extreme valuations, which necessarily induce substantial sensitivity to prices and loss probabilities in this setting.

3.12 Robustness: Random utility under full consideration

To further contrast our approach with a standard full-consideration model, we consider an alternative setting. Suppose there is full consideration, and the utility from choosing coverage limit j is given by the expected utility in Equation 1, augmented with two additional terms:

$$\mathcal{RU}(p_j^{\mathcal{S}}, \mathcal{S}_j, \lambda^{\mathcal{S}}, \Phi(\cdot) \mid \nu, \omega; \mathbf{Z}) = \Lambda_j(\mathbf{Z}) + E\mathcal{U}_j(\bar{p}^{\mathcal{S}}, \lambda^{\mathcal{S}} \mid \nu, \omega) + \varepsilon_j,$$

where $\Lambda_1(\mathbf{Z}) = 0$ and $\Lambda_j(\mathbf{Z})$, $j > 1$, are functions of $\mathbf{Z}$ known up to a finite-dimensional parameter vector, and the ε_j are i.i.d. extreme-value type 1 random variables with unknown dispersion parameter. When $\Lambda_j(\mathbf{Z}) = 0$ for all j, the model reduces to the standard EUT specification used to estimate risk preferences.[30] By adding coverage-limit–specific fixed effects, we make the standard model robust to concerns such as: "What if decision makers simply prefer certain coverage limits over others? What if they receive advice that directly alters their willingness to pay for specific coverage limits?" We parameterize $\Lambda_j(\mathbf{Z})$ as

$$\Lambda_j(\mathbf{Z}) = \alpha_\Lambda(\mathbf{Z}) \cdot \underline{\Lambda}_j + \big(1 - \alpha_\Lambda(\mathbf{Z})\big) \cdot \bar{\Lambda}_j,$$

where $\alpha_\Lambda(\mathbf{Z}) = \frac{1}{1+\exp(\gamma_\Lambda \mathbf{Z})}$. Risk aversion is parameterized as before, and ω is set to the benchmark value of \$100,000. Hence, in terms of parameters, this model—like the one above with subjective beliefs—is as flexible as our benchmark model. The key difference is $\Lambda_j(\mathbf{Z})$ versus $\psi_j(\mathbf{Z})$ in the benchmark model; both objects are parameterized in the same way, including their dependence on observables. This model also matches the observed distribution

[30] Although widely used, this model has notable shortcomings; see Apesteguia and Ballester (2018).

of coverage-limit choices reasonably well, though it fits worse than our benchmark model (total variation = 0.053, see Figure A4 of the Appendix). However, as in the subjective-beliefs exercise above, this fit comes at the cost of a substantially overstated correlation between choices and the base price (-0.13). Similarly, the implied correlation between the loss probability and choices is 0.18, roughly three times its empirical value. The cause of this overshooting is the difference in the estimated distributions of ν. While estimating the benchmark model generates only extreme types, estimating the model here produces 7% share of extreme types, while the remaining 93% come from a distribution that lies well within $[0, \bar{\nu}(\omega)]$. Specifically, the estimated beta parameters are 35 and 12, respectively, implying a bell-shaped distribution heavily concentrated around the mean value of 5.2×10^{-5} and essentially no mass at the boundaries. In sum, the data reject this model, and its risk-preference estimates fall outside the partially identified region implied by our benchmark.

3.13 Robustness: Price-Dependent Consideration

As mentioned earlier, constraints—in particular, budget constraints—can be interpreted as a form of limited consideration. If households "budget" a fixed amount for total auto-insurance spending, then, ceteris paribus, households facing higher overall prices across all three coverages should consider high liability limits less often. To assess whether this mechanism is present—and, more generally, whether prices operate directly through consideration—we perform two additional checks. First, we add the collision and comprehensive base prices $(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})$ as additional controls in our benchmark model; second, we add the full vector of base prices $(\bar{p}^{\mathcal{S}}, \bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})$ as additional controls. This allows both consideration and, for symmetry with the benchmark case, the preference mix to depend on prices. We note that when $\bar{p}^{\mathcal{S}}$ is included as a control, semi–nonparametric identification is generally difficult to obtain, since Assumption 1 is then violated. However, when only extreme types are present (i.e., $\mathcal{R}$ and $\mathcal{N}$), identification follows from standard arguments. In our parametric specification, convergence to the same estimates from multiple starting values suggests that the model

is identified under the same assumptions as our benchmark model, including allowing for potentially non-extreme preferences.[31] In either scenario, adding base prices as observables has no discernible effect on our previously reported results. As before, the choice distribution is matched nearly perfectly (see Figures A5 and A6 of the Appendix), and the shares of extreme types are the same as in the benchmark case. In the first scenario, increases in the collision and comprehensive base prices shift liability choices only marginally, and in *different* directions. Since these base prices are strongly positively correlated, their combined impact on choices is negligible. Similarly, in the second scenario, increases in the liability and comprehensive base prices have marginal and countervailing effects, while changes in the collision base price are even more muted.[32] In sum, we find no evidence that an ex ante insurance budget drives liability-limit choices.

3.14 Welfare Cost of Limited Consideration

Under our benchmark assumptions, the estimated welfare cost of limited consideration (the difference between the certainty equivalents of the simulated first-best and chosen coverages) is \$151 ($CI_{95\%} = [150, 153]$), which exceeds the average price difference between the highest and lowest coverage limits, \$104. We caution, however, that because preferences are only partially identified, the welfare cost of limited consideration can, in principle, be much larger. For $\mathcal{R}$-types, it increases monotonically in both risk aversion and wealth-at-risk. For $\mathcal{N}$-types, the welfare cost decreases monotonically in wealth-at-risk; in the limit, it converges to the price difference between the lowest and chosen coverages.[33]

[31] For identification of models with price-dependent consideration in markets with many firms, each offering products with their own price variation, see Abaluck and Adams-Prassl (2021). In our setting such variation is absent: the company sets a single base price, which determines the prices of all coverage limits.

[32] Figures A7 and A8 report, for the first and second scenarios respectively, the effect of moving the relevant base prices from their 5th to 95th percentiles. For reference, they also include the corresponding effect of changes in the loss probability, which is nearly identical to the benchmark case. Finally, with $\bar{p}^{\mathcal{L}}$ and $\bar{p}^{\mathcal{M}}$ included as controls, the correlation between $\bar{p}^{\mathcal{S}}$ and choices is -0.02 both in the data and in the model.

[33] With a concave Bernoulli function, buying more insurance than the first-best has a welfare cost that is bounded by the certainty-equivalent cost of paying the higher premium for sure. Under CARA, this bound coincides with the premium difference when the premium enters as a uniform shift across states; in our

4 Collision and Comprehensive Coverage Choices

The next step in our analysis examines collision and comprehensive coverages, which protect against damage to the insured vehicle. Collision coverage applies when the vehicle is in motion, whereas comprehensive coverage applies when the vehicle is parked.[34] The company offers l levels of collision coverage, which we refer to as the feasible set $\mathcal{L}$. Each coverage level $j \in \{1, 2, \ldots, l\}$ is characterized by a pair $(p_j^{\mathcal{L}}, \mathcal{L}_j)$, where $p_j^{\mathcal{L}}$ is the premium associated with deductible $\mathcal{L}_j$. In the event of a claim, the household pays $\mathcal{L}_j$, and the company covers the remainder.[35] Coverage levels are ordered such that $\mathcal{L}_j > \mathcal{L}_{j+1}$ and $p_j^{\mathcal{L}} < p_{j+1}^{\mathcal{L}}$. Premiums are deterministic functions of a base price $\bar{p}^{\mathcal{L}}$. Both the base price and the claim probability $\lambda^{\mathcal{L}}$ vary across households. We assume each household correctly believes that $\lambda^{\mathcal{L}}$ is independent of the coverage choice.

Comprehensive coverage has an identical structure, but may have a different deductible menu, base prices, and claim probabilities. We denote the number of coverage levels by m, the deductibles by $\mathcal{M}_k$, the associated prices by $p_k^{\mathcal{M}}$, the claim probability by $\lambda^{\mathcal{M}}$, and the feasible set by $\mathcal{M}$. In total, there are $l \times m$ possible deductible pairs. We assume decision makers only consider a subset of these: $\mathcal{K} \subseteq \mathcal{L} \times \mathcal{M}$, with $\mathcal{K} \neq \varnothing$. The specific consideration set is unobserved and may vary across households. As in our data, we assume that the first l elements of the deductible menus coincide: $\mathcal{L}_j = \mathcal{M}_j$ for $j = 1, \ldots, l$. That is, the set $\mathcal{M}$ may contain lower deductibles not available in $\mathcal{L}$.

setting it is only approximate given the full utility specification in Equation 1. In contrast, buying less insurance than the first-best need not satisfy such a bound; indeed, with CARA preferences the certainty equivalent of a lottery converges to the value of its worst outcome as risk aversion grows (Apesteguia and Ballester 2018).

[34] There are exceptions. For example, colliding with an animal on the road is covered under comprehensive coverage.

[35] We assume that potential losses exceed the highest deductible offered.

4.1 Kőszegi–Rabin (2007) model with distorted probabilities

We continue to assume that decision makers follow narrow bracketing (Read et al. 2000) and that the Bernoulli function exhibits CARA. Under these assumptions, the (negative) expected utility of a binary prospect with outcomes $\{(p_j^{\mathcal{L}},\, 1-\lambda^{\mathcal{L}}),\ (p_j^{\mathcal{L}}+\mathcal{L}_j,\, \lambda^{\mathcal{L}})\}$ can be written as

$$-E\mathcal{U}(p_j^{\mathcal{L}}, \mathcal{L}_j, \lambda^{\mathcal{L}} \mid \nu, \Omega(\cdot)) \;\propto\; \big(1-\Omega(\lambda^{\mathcal{L}})\big)\cdot e^{\nu p_j^{\mathcal{L}}} + \Omega(\lambda^{\mathcal{L}})\cdot e^{\nu(p_j^{\mathcal{L}}+\mathcal{L}_j)},$$

where $\Omega(\cdot)$ maps objective probabilities into decision weights and, as before, ν is the coefficient of absolute risk aversion. The distortion function $\Omega(\cdot)$ encompasses Kahneman and Tversky (1979) probability weighting and Kőszegi and Rabin (2007) loss aversion under "choice-acclimating personal equilibrium," as well as other behavioral phenomena such as the disappointment aversion of Gul (1991). It can also capture subjective beliefs.[36] The certainty equivalent associated with $\mathcal{L}_j$ is

$$CE(p_j^{\mathcal{L}}, \mathcal{L}_j, \lambda^{\mathcal{L}} \mid \nu, \Omega(\cdot)) = -\,p_j^{\mathcal{L}} - \frac{1}{\nu}\log\big[\big(1-\Omega(\lambda^{\mathcal{L}})\big) + \Omega(\lambda^{\mathcal{L}})\cdot e^{\nu\mathcal{L}_j}\big].$$

The DM then evaluates

$$CE(p_j^{\mathcal{L}}, \mathcal{L}_j, \lambda^{\mathcal{L}} \mid \nu, \Omega(\cdot)) + CE(p_k^{\mathcal{M}}, \mathcal{M}_k, \lambda^{\mathcal{M}} \mid \nu, \Omega(\cdot))$$

for each considered bundle $\mathcal{I}_{j,k} \equiv \{\mathcal{L}_j, \mathcal{M}_k\}$ and selects the bundle with the highest value.

To complete the model, we specify the probability distortion function and assumptions about preference heterogeneity. As a benchmark, we set $\nu = 1.1 \times 10^{-3}$, which is the estimated mean value of the CARA coefficient reported in Barseghyan et al. (2013) (the median value is 1×10^{-3}; see their Appendix Table A.17).[37] For Ω we choose the family of constant relative sensitivity (Abdellaoui et al. 2010) weighting functions:

$$\Omega(\lambda) = \begin{cases} \delta^{\eta}\lambda^{1-\eta}, & \text{if } \lambda < \delta, \\ 1-(1-\delta)^{\eta}(1-\lambda)^{1-\eta}, & \text{otherwise,} \end{cases}$$

[36] For a detailed discussion of $\Omega(\cdot)$, see Barseghyan et al. (2013).

[37] This is fully consistent with having two extreme preference types, $\mathcal{R}$ and $\mathcal{N}$, with respect to liability coverage limits. $\nu = 1.1 \cdot 10^{-3}$ combined with $w \geqslant 11{,}000$ implies an $\mathcal{R}$-type, while the same ν combined with sufficiently small w implies an $\mathcal{N}$-type.

which is close to the formulation in Prelec (1998) but more flexible. We set the inflection point δ to 0.5. Since in our data $\lambda < 0.5$ throughout, the case $\eta = 1$ is extreme: it implies $\Omega(\lambda) = 0.5$ for all relevant λ, i.e., maximal overweighting up to 0.5. We then parameterize the distribution of the random coefficient η as a mixture between a Beta distribution, $\mathbf{B}(\eta; \beta_1, \beta_2)$, and a point mass at $\eta = 1$:

$$F(\eta \mid \mathbf{Z}) = \alpha_\eta(\mathbf{Z}) \cdot \mathbf{B}(\eta; \beta_1, \beta_2) + \big(1 - \alpha_\eta(\mathbf{Z})\big) \cdot \mathbf{1}(\eta = 1),$$

where the distribution parameters (β_1, β_2) and the share $\alpha_\eta(\mathbf{Z})$ are unknown. The latter is parameterized as

$$\alpha_\eta(\mathbf{Z}) = \frac{1}{1 + \exp(\gamma_\eta \mathbf{Z})}.$$

4.2 The limited consideration mechanism

We continue to rely on the ARC framework. Each deductible bundle $\{\mathcal{L}_j, \mathcal{M}_k\}$ is considered with probability $\Psi_{j,k}(\eta; \mathbf{Z})$, so that consideration depends on both observables and the preference parameter η.

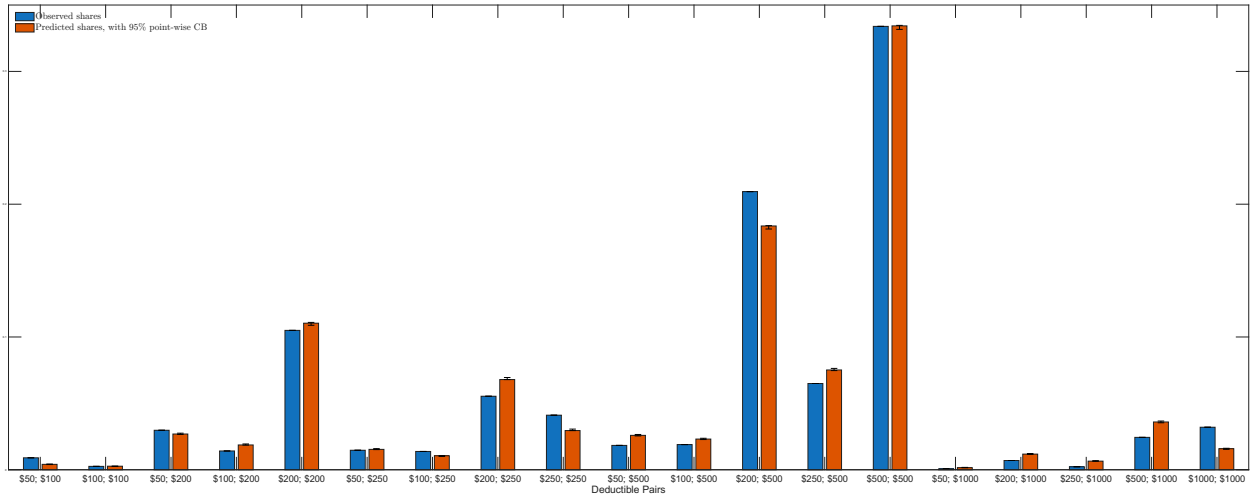
Definition 2 ($ARC^{\triangle}$ model)**.** *An ARC model is an $ARC^{\triangle}$ model if the consideration probability of each bundle $\mathcal{I}_{j,k}$, $\Psi_{j,k}(\eta; \mathbf{Z})$, satisfies:*

- $\Psi_{j,k}(\eta; \mathbf{Z}) = 0$ *if* $j > k$*; and*
- $\Psi_{j,k}(\eta; \mathbf{Z}) = \psi_j^{\mathcal{L}}(\eta; \mathbf{Z}) \cdot \psi_k^{\mathcal{M}}(\eta; \mathbf{Z})$ *otherwise.*

Semi–nonparametric identification of such a rich model generally requires multiple sources of variation in the data (for a detailed analysis of the case with linear utilities, see Barseghyan and Molinari 2025). In our setting, we rely on variation in two base prices and on the fact that the joint distribution of choices is upper triangular. The identification argument is provided in Appendix A3.[38]

[38] In special cases, such as proportional consideration of Barseghyan et al. (2021a), semi–nonparametric identification can be attained with a single excluded regressor. For partial identification results that allow for consideration formation to depend on preferences see Barseghyan et al. (2021b).

Figure 5: Deductible Choices



For parsimony, we assume that consideration probabilities for collision deductibles are the same as for comprehensive deductibles: $\psi_j^{\mathcal{L}}(\eta; \mathbf{Z}) = \psi_j^{\mathcal{M}}(\eta; \mathbf{Z})$, for $j = 1, \ldots, l$. We set

$$\psi_k^{\mathcal{M}}(\eta; \mathbf{Z}) = \alpha_\psi(\eta; \mathbf{Z}) \cdot \underline{\psi}_k^{\mathcal{M}} + \big(1 - \alpha_\psi(\eta; \mathbf{Z})\big) \cdot \bar{\psi}_k^{\mathcal{M}},$$

where

$$\alpha_\psi(\eta; \mathbf{Z}) = \eta^{q_\eta} \cdot \left[\frac{1}{1 + \exp(\gamma_\psi \mathbf{Z})}\right]^{q_\psi}, \quad q_\eta, q_\psi \geqslant 0.$$

Because both η and $\frac{1}{1+\exp(\gamma_\psi \mathbf{Z})}$ lie in the unit interval, α_ψ is also in $[0, 1]$. If $q_\eta = 0$ (respectively $q_\psi = 0$), then consideration is independent of preferences (respectively observables). In sum, estimation reduces to finding the parameter vector

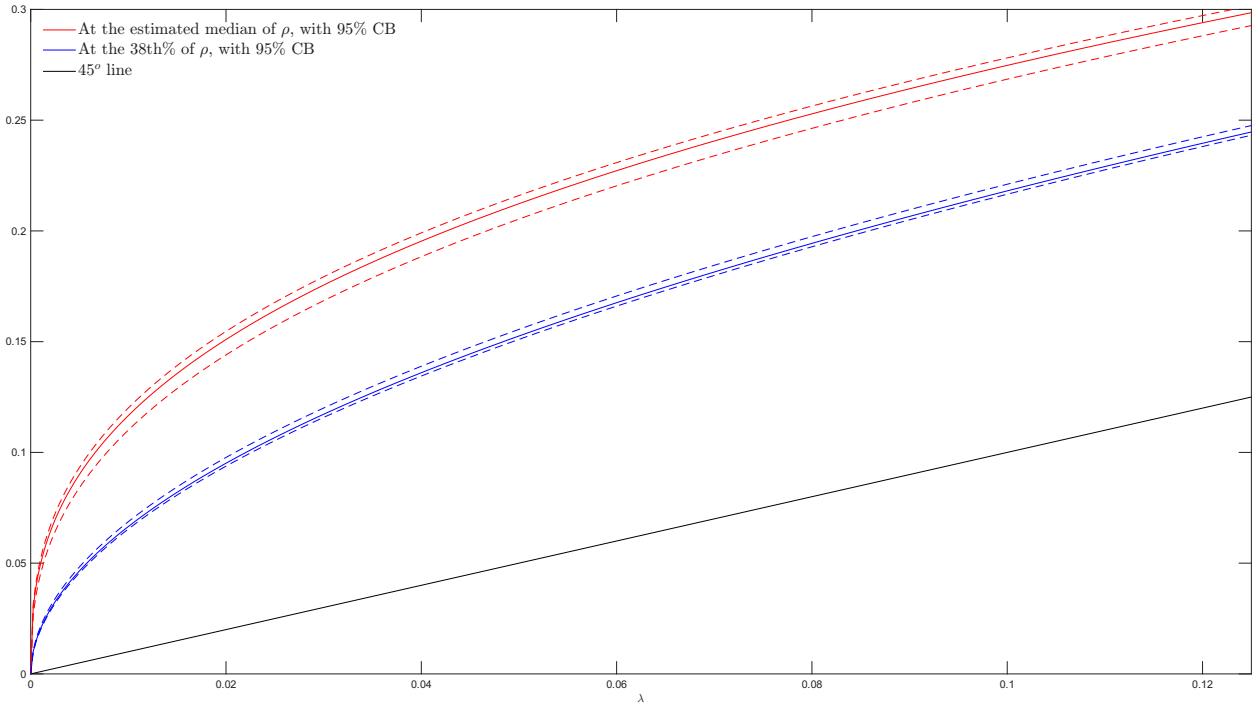
$$\{\beta_1, \beta_2, \gamma_\eta;\ \{\underline{\psi}_k^{\mathcal{M}}, \bar{\psi}_k^{\mathcal{M}}\}_{k=1}^m,\ \gamma_\psi,\ q_\eta,\ q_\psi\}$$

that maximizes the quasi–log-likelihood.

4.3 Benchmark results

Despite its parsimonious parameterization, the model predicts a distribution of deductible bundles that closely matches its empirical counterpart (Figure 5). The joint distribution contains nineteen nonzero shares, while, conditional on observables, only six menu-specific

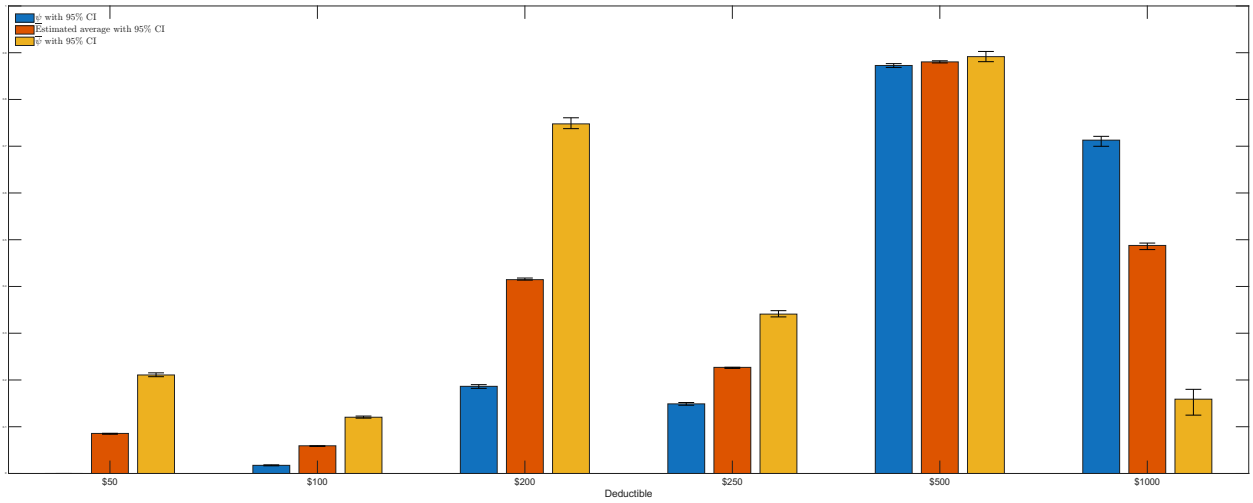
Figure 6: The Probability Distortion Function



consideration parameters (one per deductible level) are estimated. Consistent with prior work (Cohen and Einav 2007; Sydnor 2010; Barseghyan et al. 2013), we find substantial willingness to pay for lower deductibles, with about 25% of households trivially ranking deductibles from lowest to highest. The remainder exhibit substantial over-weighting of small probabilities. Figure 6 shows the probability distortion function at the median and at the 38$^{\text{th}}$ percentile of the estimated η distribution.[39] Turning to consideration, Figure 7 reports estimated consideration parameters: $\underline{\psi}$, $\overline{\psi}$, and $\psi(\mathbf{Z})$ evaluated at the average values of the observables. As the figure shows, limited consideration is substantial, with the resulting average welfare cost of \$78 (with $CI_{95\%} = [\$73, \$80]$), or about a third of the average price difference between the cheapest and the most expensive deductible pairs. Both $\underline{\psi}^{\mathcal{M}}$ and $\bar{\psi}^{\mathcal{M}}$ are highest and nearly equal to each other for the \$500 deductible (at 0.87 and 0.89, respectively), implying that, irrespective of observables, (\$500, \$500) is the most often

[39] The 38$^{\text{th}}$ percentile corresponds to the median of the distribution, conditional on the type not being extreme.

Figure 7: Consideration parameters



considered pair. For the lower deductibles, the $\underline{\psi}^{\mathcal{M}}$ parameters are significantly lower than the corresponding $\bar{\psi}^{\mathcal{M}}$ parameters, while for the \$1000 deductible the opposite is true. Hence, as $\alpha_{\psi}(\eta; \mathbf{Z})$ increases, consideration of lower deductibles increases, while consideration of the \$1000 deductible decreases. In turn, $\alpha_{\psi}(\eta; \mathbf{Z})$ depends both on the preference parameter and on observables.

The dependence on the preference parameter is governed by q_{η}, estimated to be 0.36, with a 95% confidence interval of $[0.34, 0.42]$. As a result, consideration varies with preferences: households with a stronger willingness to pay for higher (or lower) coverage levels tend to consider those options more frequently. To gauge the importance of this dependence, we simulate choices twice. The simulations are identical except that the first uses the estimated model, while the second uses the same model but sets $q_{\eta} = 0$, thereby removing the link between consideration and preferences. Choices differ in 7.5% of cases, and among these households, more than nine out of ten achieve higher expected utility under the benchmark specification.[40] The average difference in certainty equivalents among these 7.5% of households is \$31. One interpretation of this finding is that limited consideration may arise for

[40]For the remaining households, variation in prices and loss probabilities implies that higher deductibles become more attractive when probability distortions are stronger; thus, failing to consider these deductibles can be costly.

a number of reasons not directly related to households' preferences. Despite these reasons, households appear to search purposefully. While our framework is that of simultaneous search, we note that the correlation between preferences and consideration may also arise with sequential search. In the seminal sequential-search framework of Weitzman (1979), Pandora's rule specifies that (1) the next alternative evaluated is the one with the highest reservation utility, and (2) search stops when the highest (expected) utility among considered alternatives exceeds the reservation utility of those remaining. When reservation utilities broadly align with preferences, considered alternatives will, on average, have higher expected utility than those excluded. Moreover, even with uniform priors over alternatives' utilities, sequential search—much like the satisficing process of Simon (1956)—can produce a positive relationship between an alternative's utility ranking and the likelihood it is considered.[41]

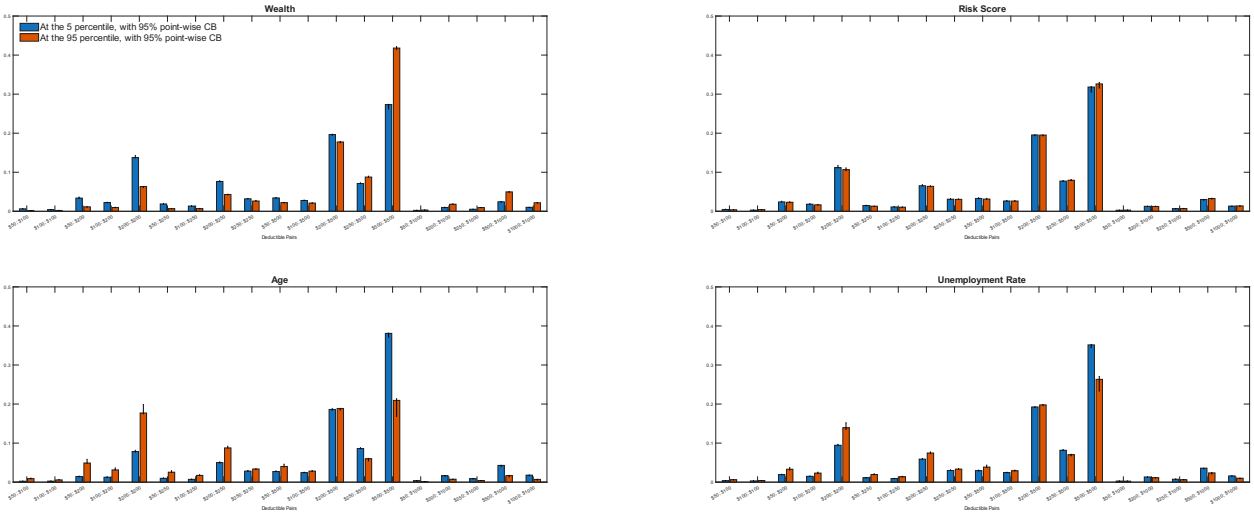
Preferences and consideration both vary substantially with observables. Age has the most pronounced cumulative effect: older DMs choose lower deductibles far more often (see the bottom left panel of Figure 8; estimates of γ_η and γ_ψ appear in Table A2 of the Appendix.) One possible explanation is that lower deductibles were more prevalent in the past and, once chosen, remained in the consideration set. The time trends in the data support this interpretation. From a behavioral economics perspective, this explanation is consistent with Malmendier and Nagel (2011) on the role of past experiences in shaping behavior, as well as with status quo biases (Samuelson and Zeckhauser 1988; Kempf and Ruenzi 2021) and inertia in consideration sets (Klemperer 1995; Honka 2014; Handel 2013).[42]

Higher unemployment rates or older DMs are associated with greater consideration parameters for lower deductibles and lower parameters for the highest deductible. These factors also increase the share of extreme preference types, though these – as noted in Section 3 – may alternatively reflect differences in consideration. The cumulative effect of unemploy-

[41] For example, suppose there are two alternatives, $A \prec B$, and only B is acceptable to the DM. If the DM searches sequentially with uniform priors, then, with sufficiently high search costs, the observed consideration sets will be, with equal probability, $\{A, B\}$ or $\{B\}$.

[42] To remind the reader, we consider only the DM's first choice with the company. Thus, auto-renewals do not explain the patterns described, though the limited consideration mechanism could also account for them.

Figure 8: The role of observables



ment is strong, see the top left panel of Figure 8. An increase from the 5th to the 95th percentile lowers the probability of choosing higher deductible pairs (i.e. those comprised of only \$500 or \$1000) by more than 15%. In the Chetty and Szeidl (2007) framework, this pattern is readily explained: as unemployment rises, a non-trivial fraction of households face liquidity constraints induced by consumption commitments, shifting both consideration and choices toward lower deductibles. Higher wealth has the opposite effect, plausibly by easing these constraints. (The effect of the remaining variables is displayed in Figure A2 of the Appendix).

Finally, we turn to the relationship between base prices and choices, echoing the discussion of liability coverage. In the data, the correlations between base prices and chosen deductible are -0.12 and -0.14 for collision and comprehensive coverage, respectively. The model reproduces these correlations nearly identically (-0.12 and -0.15, respectively), despite (1) our estimation procedure not directly targeting these moments, and (2) choices in the model being made over deductible pairs, so collision and comprehensive choices are not independent, whereas the reported moments are calculated separately for each coverage.

4.4 Alternative preference specifications

Our estimation strategy fixes ν at $1.1 \cdot 10^{-3}$ and estimates the distribution of the probability distortion parameter. Resetting the CARA coefficient to any value above 7×10^{-5} would keep our preference structure consistent with the results in Section 3. Doing so and re-estimating the model yields qualitative patterns identical to those described above, with one exception: the lower is the ν the more over-weighting the probability distortion function predicts: ceteris paribus, to generate the same average willingness to pay for lower deductibles, the loss events must carry higher weights. Yet another alternative is to fix the distortion function and estimate the distribution of the risk aversion parameter on a support bounded from below by 7×10^{-5}. Doing so, too, yields qualitative patterns identical to those described above.[43]

4.5 Probability Distortions: Interpretation and Comparison

A central premise of behavioral theories of choice under risk is that decision makers transform objective probabilities nonlinearly when forming valuations. This nonlinearity matters both for *levels* and for *slopes*. In terms of levels, the inverted S-shaped distortion function we estimate raises willingness to pay for insurance, especially at low claim probabilities. Because comprehensive claim probabilities are much lower than collision probabilities, this feature gives the preference component of the model the greatest scope to generate the triangular joint distribution of deductibles. In terms of slopes, it implies muted sensitivity of lottery valuations (i.e., certainty equivalents) to probability changes in the ranges relevant for our setting.[44] As shown in Barseghyan et al. (2013), when probability weighting

[43] A natural next step would be to allow for heterogeneity both in risk aversion and in probability distortions. However, combining such a saturated preference model with limited consideration would make semi–non-parametric identification tenuous at best. Barseghyan and Molinari (2023) show that a model with two types—one following expected utility and one following the dual theory of Yaari (1987) (which in our setting corresponds to linear utility with probability distortions) – can be identified, but only under the assumption that consideration is independent of preferences. They further find that most decision makers are of the Yaari (1987) type.

[44] For classic discussions, see Kahneman and Tversky (1979); Tversky and Kahneman (1992). For reverse-S shaped probability weighting functions, see Prelec (1998).

is combined with loss aversion (or, alternatively, disappointment aversion; see Gul 1991), the resulting probability distortion inherits the probability-insensitivity of the underlying weighting function—a pattern that loss aversion or disappointment aversion alone cannot generate. In this sense, probability weighting governs the *slope* of the distortion, whereas loss aversion (or disappointment aversion) primarily shifts its *level*, moving the distortion function upward (see Figure 3 of Barseghyan et al. 2013).

Empirically, we assess local sensitivity using conditional correlations between choices and the relevant claim probabilities. In the data, overall sensitivity to loss probabilities is small: the conditional correlations between choices and claim probabilities are 0.01 for collision and 0.03 for comprehensive (in the model, −0.01 and 0.01, respectively). Some of this insensitivity reflects limited consideration. However, under full consideration (i.e., the counterfactual in which all alternatives are considered), the corresponding correlations remain low—0.01 for collision and 0.10 for comprehensive. Comparing these objects implies that limited consideration materially attenuates sensitivity primarily for comprehensive, whereas collision choices remain essentially insensitive even absent consideration frictions. These patterns—while agnostic about the strength of loss aversion—are consistent with the classic probability-weighting interpretation: even under full consideration, the estimated probability distortion implies substantial insensitivity to probability changes in the empirically relevant ranges.[45]

We conclude by comparing the distortion function estimated here to those reported in Barseghyan et al. (2013). Notwithstanding the discussion in the previous section, the estimated distortion function (for non-extreme types) is comparable to—though toward the upper end of—the range reported in Barseghyan et al. (2013). Two offsetting forces contribute to this similarity. On the one hand, in our model higher deductibles are considered far more often than lower ones, which, all else equal, increases the willingness to pay for lower

[45]Differences in implied sensitivity across coverages arise mechanically from the functional form and the restriction that the same single-parameter (Prelec-type) weighting function applies in both domains. Because the slope of the weighting function varies with the probability level, and the two coverages operate over different probability ranges, the model can generate different local responsiveness even under a common parameterization.

deductibles required to match the data; in Barseghyan et al. (2013), by contrast, decision makers consider the entire choice set. On the other hand, we find that about a quarter of households rank deductibles from lowest to highest and therefore—subject to the pairwise choice restriction—tend to select the lowest deductible they consider. As a result, the remaining households account for most of the observed selection of higher deductibles, and matching their behavior requires less probability distortion.

5 Discussion of preference stability and types

So far, given the natural dichotomy created by the narrow bracketing assumption and the Kőszegi and Rabin (2007) framework, we have examined large- and small-stakes contexts separately. We now ask whether the joint distribution of all three coverage choices can yield additional insights. Following Einav et al. (2012), we study Spearman rank correlations between liability limits and deductibles constructed from residual variation, as follows. First, we regress the chosen (ordered) liability coverage limit, as well as collision and comprehensive deductibles, on observables $\mathbf{Z}$, base prices $(\bar{p}^{\mathcal{S}}, \bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})$, and loss probabilities $(\lambda^{\mathcal{S}}, \lambda^{\mathcal{L}}, \lambda^{\mathcal{M}})$.[46] Second, we take the residuals from these regressions and compute Spearman correlations. In the stability analysis of Einav et al. (2012), positive correlations across coverage levels suggest preference stability, while negative correlations suggest instability.[47] In general, these coefficients have no causal or structural interpretation, but they serve as useful diagnostics and may shed light on households' decision-making.

Consistent with Barseghyan et al. (2018), we find that rank correlations between deductible choices are positive, as expected. However, correlations between deductibles and liability coverage limits are negative (Table 6): households that choose higher liability limits

[46] Deductibles are ordered from highest to lowest, consistent with the notion that lower indices indicate less coverage.

[47] For a discussion of the correlations-based test, the assumptions required for its validity, and a general framework that does not impose a specific functional form on expected utility, see Barseghyan et al. (2025). For a model-based test, see Barseghyan et al. (2011). For an inference procedure that exploits stability, see Barseghyan et al. (2016).

tend to opt for higher deductibles (lower coverage) in collision and comprehensive. At first glance, this might indicate preference instability.[48] Yet both limited consideration and the Kőszegi and Rabin (2007) framework can generate such patterns if households that consider (or prefer) higher liability limits also tend to consider (or prefer) higher deductibles.

Our liability analysis identified two household types: $\mathcal{R}$ and $\mathcal{N}$. We now ask whether incorporating these types into the deductible-choice model can explain the observed correlations. Using 10,000 simulations of the benchmark liability model, we compute each household's expected type—the posterior probability of being an $\mathcal{R}$-type—conditional on $\mathbf{Z}$, liability coverage prices, claim probability, and the chosen limit. Including expected type as a control in the deductible regressions reduces the correlations with liability coverage to 0.01 (collision) and -0.01 (comprehensive)—as predicted by our benchmark models (both 0.00).

We also re-estimate the deductible choice model from Section 4 with expected type as an additional observable. We find that type affects choice via both consideration and preferences: $\mathcal{R}$-types consider the \$1,000 deductible more often, whereas $\mathcal{N}$-types more often consider lower deductibles. $\mathcal{R}$-types also exhibit less probability distortion than $\mathcal{N}$-types. We then simulate households' liability and deductible choices jointly, linking the two only through the drawn preference type: we first generate liability choices, and then feed the realized type (rather than expected type) into the deductible model to determine probability distortions, consideration probabilities, and deductible choices. When residuals are constructed controlling only for $\mathbf{Z}$, prices, and loss probabilities, simulated correlations are -0.05 for both collision and comprehensive—over half the empirical magnitude. When expected type is also controlled for, simulated correlations shrink to -0.02, nearly matching the empirical pattern.

In sum, negative correlations between large- and small-stakes choices do not imply preference instability. Instead, they arise from systematic differences between $\mathcal{R}$- and $\mathcal{N}$-type

[48] For why stable EUT preferences satisfying single crossing must yield non-negative cross-context correlations, see Barseghyan et al. (2025).

Table 6: Spearman rank correlations across coverages: data vs. model

Source	Controls	$\mathcal{S}$ vs. $\mathcal{L}$	$\mathcal{S}$ vs. $\mathcal{M}$	$\mathcal{L}$ vs. $\mathcal{M}$
Data	$\mathbf{Z}$, $\bar{p}$'s, λ's	-0.09	-0.07	0.50
Model	$\mathbf{Z}$, $\bar{p}$'s, λ's	-0.05	-0.05	0.46
Data	$\mathbf{Z}$, $\bar{p}$'s, λ's, $\mathcal{R}/\mathcal{N}$	0.01	-0.01	0.50
Model	$\mathbf{Z}$, $\bar{p}$'s, λ's, $\mathcal{R}/\mathcal{N}$	-0.02	-0.02	0.46

Notes: $\mathcal{S}$ = liability, $\mathcal{L}$ = collision, $\mathcal{M}$ = comprehensive. In the rows that include $\mathcal{R}/\mathcal{N}$, the deductible-choice model uses the expected type implied by the liability-choice model as an additional covariate augmenting $\mathbf{Z}$. Simulations use common draws of type to generate both liability and deductible-pair choices.

households. Accounting for type renders these correlations negligible and aligns the predictions of our model—taken as a whole—with observed choices across all three contexts. In particular, our preference structure, with a single parameterization, combined with limited consideration, can simultaneously (i) generate extreme preference rankings over liability coverages that are essentially invariant to observed price variation, (ii) match the high willingness to pay for lower deductibles that responds to price variation, and (iii) explain rank correlations across contexts. Importantly, this parameterization also admits a natural interpretation of $\mathcal{N}$-types: they should not be viewed as risk-neutral households (i.e., with $\nu \approx 0$). Rather, they are "nothing-to-lose" households (i.e., behaving as if $\omega \approx 0$). To see why, note that this interpretation is fully consistent with our analysis, and in particular with the assumed value of ν in the deductible-choice model. By contrast, interpreting $\mathcal{N}$-types as having $\nu \approx 0$ would imply that they prefer higher deductibles more often than $\mathcal{R}$-types, which runs counter to the data.[49]

We conclude this section by noting that random-utility specifications such as those in Sections 3.11 and 3.12 do not generate an economically interpretable notion of latent type. Conditional on observables, prices, and loss probabilities, any remaining heterogeneity in liability choices is absorbed by i.i.d. utility shocks. In principle, one can condition on an observed choice and recover the implied distribution of the vector of shocks consistent with

[49]Put differently, if one were to re-estimate both models with the sole modification that the risk-aversion parameter could take values $\nu = 0$ or $\nu = \bar{\nu} > 7 \times 10^{-5}$, while fixing ω at a non-trivial value, the resulting specification would fail to reproduce the rank correlations observed in the data.

that choice (an object in $\mathbb{R}^s$, where s is the number of liability limits). But this object is inherently alternative-specific and tied to normalization, and it does not yield a persistent, interpretable household-level type comparable to the $\mathcal{R}/\mathcal{N}$ distinction.

6 Conclusion

An overarching goal of this paper is to disentangle the forces that shape decision-making under risk. Two major takeaways emerge. First, a single preference framework can generate empirically plausible rankings of alternatives across contexts that differ sharply in stakes and probabilities. Second, limited consideration, together with latent types, is central to generating the choice variation needed to match the data; by contrast, full-consideration models deliver counterfactual choice patterns and biased preference inference.

In the liability domain, choices partially identify pairs of risk aversion and wealth exposure that imply extreme preference rankings over coverage limits—rankings that are essentially insensitive to in-sample price variation. In contrast, deductible choices are informative about probability distortions: holding the Bernoulli function fixed, roughly three quarters of households exhibit non-extreme preferences, with the degree of probability over-weighting in line with previous studies (Barseghyan et al. 2013). Limited consideration is crucial for explaining the remaining variation. In liability coverage, heterogeneity in consideration is sufficient to match the observed choice distribution. In collision and comprehensive coverages, limited consideration accounts for several robust features of the joint distribution of choices that are difficult to reconcile under full consideration. Moreover, there is evidence that consideration is "purposeful": households are more likely to consider deductible pairs that they prefer, and this channel is empirically important. A key source of unobserved heterogeneity is the presence of two latent liability preference types. Allowing type to influence behavior in both domains accounts for the correlations across contexts: households that behave as if they have little (or no) wealth at risk in the liability domain also behave as if liquidity constraints

bind more often in the deductible domain, making lower deductibles more attractive.

Taken together, these results provide empirical discipline on the contours of limited consideration in our setting—in particular, its connection to preferences, the limited role of simple budget-constraint accounts, the nuanced role of liquidity constraints, and the strong influence of observables.[50] They also point to a clear agenda: understanding how consideration sets form, and how households acquire, process, and act on information about alternatives across contexts and time. Limited consideration may also be a missing ingredient in behavioral accounts of cross-context variation: models such as Bordalo et al. (2013); Kőszegi and Szeidl (2012); Bushong et al. (2020) can be recast so that behavioral forces affect not only utilities but also which alternatives are considered. Given scant evidence of choice-set dependence in risk preferences (Chen et al. 2024) and recent theoretical and methodological advances on limited consideration,[51] this channel appears especially promising for understanding variation in risky choice both across and within contexts. It is also complementary to several other strands in the literature, including endogenous information acquisition (e.g., rational inattention), search frictions and incomplete product evaluation, information design and marketing, and household-level constraints.[52]

[50] In other settings, there is also suggestive evidence that third-party advice and industry practices can shape choices by compressing the effective choice set. Collier et al. (2022) analyze flood insurance (a large-stakes setting) and conclude that "some combination of industry practices that emphasize fully insuring and probability distortions in decision-making are likely explanations." We view practices that emphasize "full insurance" as a clean example of a mechanism that can operate through limited consideration.

[51] For early examples of limited-attention/limited-consideration models, see, among others, Manzini and Mariotti (2014) and Masatlioglu et al. (2012). On costly information acquisition and related revealed-preference restrictions, see Caplin and Dean (2015). Within econometric models of limited consideration, Barseghyan et al. (2021b) allow for dependence between preferences and consideration and obtain sharp bounds on the preference distribution under minimal restrictions on consideration formation; for related partial-identification results, see Cattaneo et al. (2020) and Cattaneo et al. (2024).

[52] For endogenous information acquisition and rational inattention, see Sims (2003) for a classic formulation; Matějka and McKay (2015) for the rational inattention–logit foundation; and Fosgerau et al. (2020) on the connection to additive random utility models. For search frictions and incomplete product evaluation, see, e.g., Honka (2014) and Abaluck et al. (forthcoming). For information design, advertising, and marketing, see, e.g., Goeree (2008) and Gabaix and Laibson (2006). For household-level constraints, see Ericson and Sydnor (forthcoming) for a theoretical analysis of liquidity constraints in insurance markets.

References

Abaluck, Jason and Abi Adams-Prassl, "What do Consumers Consider Before They Choose? Identification from Asymmetric Demand Responses," *The Quarterly Journal of Economics*, 03 2021, *136* (3), 1611–1663.

_ **and Jonathan Gruber**, "Choice inconsistencies among the elderly: evidence from plan choice in the Medicare Part D program," *American Economic Review*, 2011, *101* (4), 1180–1210.

_ **, Giovanni Compiani, and Fan Zhang**, "A Method to Estimate Discrete Choice Models That Is Robust to Consumer Search," *Journal of Political Economy*, forthcoming. Forthcoming.

Abdellaoui, Mohammed, Olivier L'Haridon, and Horst Zank, "Separating curvature and elevation: A parametric probability weighting function," *Journal of Risk and Uncertainty*, 2010, *41* (1), 39–65.

Apesteguia, Jose and Miguel A. Ballester, "Monotone Stochastic Choice Models: The Case of Risk and Time Preferences," *Journal of Political Economy*, 2018, *126* (1), 74 – 106.

Barseghyan, Levon and Francesca Molinari, "Risk Preference Types, Limited Consideration, and Welfare," *Journal of Business & Economic Statistics*, 2023, *41* (4), 1011–1029.

_ **and** _ , "Identification and Inference for Pure Random Coefficients Models with Limited Consideration," *Working Paper*, 2025.

_ **,** _ **, and Joshua C. Teitelbaum**, "Inference under stability of risk preferences," *Quantitative Economics*, 2016, *7* (2), 367–409.

_ **,** _ **, and** _ , "Learning about Stability of Risk Preferences," June 2025. Journal of Political Economy: Microeconomics.

_ **,** _ **, and Matthew Thirkettle**, "Discrete choice under risk with limited consideration," *American Economic Review*, 2021, *111* (6), 1972–2006.

_ **,** _ **, Ted O'Donoghue, and Joshua C Teitelbaum**, "The nature of risk preferences: Evidence from insurance choices," *American Economic Review*, 2013, *103* (6), 2499–2529.

_ **, Jeffrey Prince, and Joshua C. Teitelbaum**, "Are Risk Preferences Stable across Contexts? Evidence from Insurance Data," *American Economic Review*, April 2011, *101* (2), 591–631.

_ **, Joshua C. Teitelbaum, and Lin Xu**, "Different Contexts, Different Risk Preferences?," *Working paper*, 2018.

_ **, Maura Coughlin, Francesca Molinari, and Joshua C Teitelbaum**, "Heterogeneous choice sets and preferences," *Econometrica*, 2021, *89* (5), 2015–2048.

Bhargava, Saurabh, George Loewenstein, and Justin Sydnor, "Choose to lose: Health plan choices from a menu with dominated option," *The Quarterly Journal of Economics*, 2017, *132* (3), 1319–1372.

Bordalo, Pedro, Nicola Gennaioli, and Andrei Shleifer, "Salience and Consumer Choice," *Journal of Political Economy*, 2013, *121* (5), 803–843.

Bushong, Benjamin, Matthew Rabin, and Joshua Schwartzstein, "A Model of Relative Thinking," *The Review of Economic Studies*, 10 2020, *88* (1), 162–191.

Caplin, Andrew and Mark Dean, "Revealed Preference, Rational Inattention, and Costly Information Acquisition," *American Economic Review*, 2015, *105* (7), 2183–2203.

Cattaneo, Matias D., Paul Cheung, Xinwei Ma, and Yusufcan Masatlioglu, "Attention Overload," Technical Report 2024. available at https://arxiv.org/abs/2110.10650.

_ , Xinwei Ma, Yusufcan Masatlioglu, and Elchin Suleymanov, "A Random Attention Model," *Journal of Political Economy*, 2020, *128* (7), 2796–2836.

Chen, Zhuo, Russell Golman, and Jason Somerville, "Menu-dependent risk attitudes: Theory and evidence," *Journal of Risk and Uncertainty*, 02 2024, *68*, 77–105.

Chetty, Raj and Adam Szeidl, "Consumption Commitments and Risk Preferences," *The Quarterly Journal of Economics*, 05 2007, *122* (2), 831–877.

Cohen, Alma and Liran Einav, "Estimating Risk Preferences from Deductible Choice," *American Economic Review*, June 2007, *97* (3), 745–788.

Collier, Benjamin L., Daniel Schwartz, Howard C. Kunreuther, and Erwann O. Michel-Kerjan, "Insuring large stakes: A normative and descriptive analysis of households' flood insurance coverage," *Journal of Risk and Insurance*, 2022, *89* (2), 273–310.

Denuit, Michel, Xavier Marchal, Sandra Pitrebois, and Jean-Franois Walhin, *Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus-Malus Systems*, Chichester: John Wiley, 2007.

Einav, Liran, Amy Finkelstein, Iuliana Pascu, and Mark R. Cullen, "How General Are Risk Preferences? Choices under Uncertainty in Different Domains," *American Economic Review*, May 2012, *102* (6), 2606–38.

Ericson, Keith M. and Justin Sydnor, "Liquidity Constraints and the Value of Insurance," *American Economic Journal: Microeconomics*, forthcoming. Forthcoming.

Fosgerau, Mogens, Emerson Melo, André de Palma, and Matthew Shum, "Discrete Choice and Rational Inattention: A General Equivalence Result," *International Economic Review*, 2020, *61* (4), 1569–1599.

Gabaix, Xavier and David Laibson, "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *Quarterly Journal of Economics*, 2006, *121* (2), 505–540.

Goeree, Michelle Sovinsky, "Limited Information and Advertising in the U.S. Personal Computer Industry," *Econometrica*, 2008, *76* (5), 1017–1074.

Gul, Faruk, "A Theory of Disappointment Aversion," *Econometrica*, 1991, *59* (3), 667–686.

Handel, Benjamin R, "Adverse selection and inertia in health insurance markets: When nudging hurts," *American Economic Review*, 2013, *103* (7), 2643–2682.

Hansen, Bruce E., *Econometrics*, Princeton: Princeton University Press, 2022.

Honka, Elisabeth, "Quantifying search and switching costs in the U.S. auto insurance industry," *RAND Journal of Economics*, 2014, *45* (4), 847–884.

Jin, Yizhou, "Re-examining Moral Hazard under Inattention: New Evidence from Behavioral Data in Auto Insurance," in "Proceedings of the 24th ACM Conference on Economics and Computation (EC '23)" Association for Computing Machinery New York, NY, USA 2023.

_ **and Shoshana Vasserman**, "Buying Data from Consumers: The Impact of Monitoring Programs in U.S. Auto Insurance," NBER Working Paper 29096, National Bureau of Economic Research, Cambridge, MA July 2021. First version July 2021; last revised July 3, 2025. Also available at SSRN (DOI: 10.2139/ssrn.3897531).

Kahneman, Daniel and Amos Tversky, "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 1979, *47* (2), 263–291.

Kempf, Alexander and Stefan Ruenzi, "Status quo bias in financial markets," *Review of Financial Studies*, 2021, *34* (2), 852–896.

Kőszegi, Botond and Matthew Rabin, "Reference-Dependent Risk Attitudes," *American Economic Review*, 2007, *97* (4), 1047–1073.

Klemperer, Paul, "Competition when consumers have switching costs: An overview with applications to industrial organization, macroeconomics, and international trade," *Review of Economic Studies*, 1995, *62* (4), 515–539.

Kőszegi, Botond and Adam Szeidl, "A Model of Focusing in Economic Choice," *The Quarterly Journal of Economics*, 11 2012, *128* (1), 53–104.

Malmendier, Ulrike and Stefan Nagel, "Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?," *The Quarterly Journal of Economics*, 02 2011, *126* (1), 373–416.

Manski, Charles F, "The structure of random utility models," *Theory and decision*, 1977, *8* (3), 229.

Manzini, Paola and Marco Mariotti, "Stochastic choice and consideration sets," *Econometrica*, 2014, *82* (3), 1153–1176.

Masatlioglu, Yusufcan, Daisuke Nakajima, and Erkut Y. Ozbay, "Revealed Attention," *American Economic Review*, 2012, *102* (5), 2183–2205.

Matějka, Filip and Alisdair McKay, "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *American Economic Review*, 2015, *105* (1), 272–298.

Mortimer, Julie Holland, "Discussion of Levon Barseghyan and Francesca Molinari's "Risk Preference Types, Limited Consideration, and Welfare"," *Journal of Business & Economic Statistics*, 2023, *41* (4), 1042–1045.

Neumann, Nico, Ulf Böckenholt, and Ashish Sinha, "A meta-analysis of extremeness aversion," *Journal of Consumer Psychology*, 2016, *26* (2), 193–212.

Politis, Dimitris N., Joseph P. Romano, and Michael Wolf, *Subsampling* Springer Series in Statistics, Springer, 1999.

Prelec, Drazen, "The Probability Weighting Function," *Econometrica*, 1998, *66* (3), 497–527.

Rabin, Matthew, "Risk Aversion and Expected-Utility Theory: A Calibration Theorem," *Econometrica*, 2000, *68* (5), 1281–1292.

Read, Daniel, George Loewenstein, Matthew Rabin, Gideon Keren, and David Laibson, "Choice bracketing," *Elicitation of preferences*, 2000, pp. 171–202.

Samuelson, William and Richard Zeckhauser, "Status quo bias in decision making," *Journal of Risk and Uncertainty*, 1988, *1* (1), 7–59.

Simon, Herbert A., "Rational choice and the structure of the environment," *Psychological Review*, 1956, *63* (2), 129–138.

_ , "Theories of Decision-Making in Economics and Behavioral Science," *The American Economic Review*, 1959, *49* (3), 253–283.

Sims, Christopher A., "Implications of Rational Inattention," *Journal of Monetary Economics*, 2003, *50* (3), 665–690.

Sydnor, Justin, "(Over)Insuring Modest Risks," *American Economic Journal: Applied Economics*, 2010, *2* (4), 177–199.

Tversky, Amos, "Choice by elimination," *Journal of Mathematical Psychology*, 1972, *9* (4), 341 – 367.

_ **and Daniel Kahneman**, "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 1992, *5* (4), 297–323.

Weitzman, Martin L., "Optimal Search for the Best Alternative," *Econometrica*, 1979, *47* (3), 641–654.

WSJ, "Your Credit Score Affects Your Car Insurance Rate—Here's How to Save," June 17, 2024.

Yaari, Menahem E., "The Dual Theory of Choice under Risk," *Econometrica*, 1987, *55* (1), 95–115.

Appendix

A1 Additional Figures and Tables

Figure A1: The role of observables, cont.d

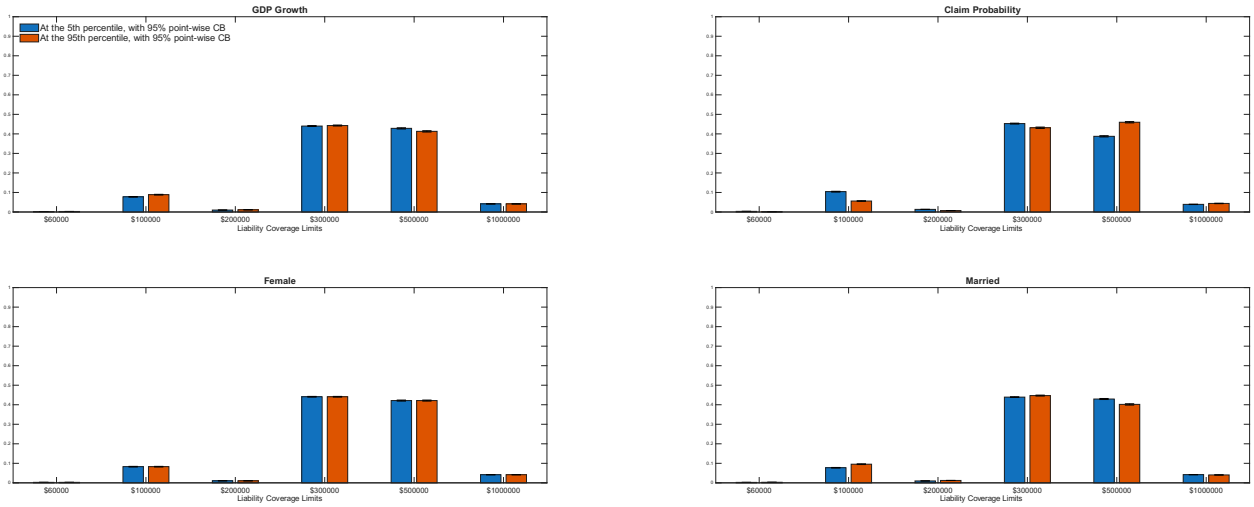


Figure A2: The role of observables, cont.d

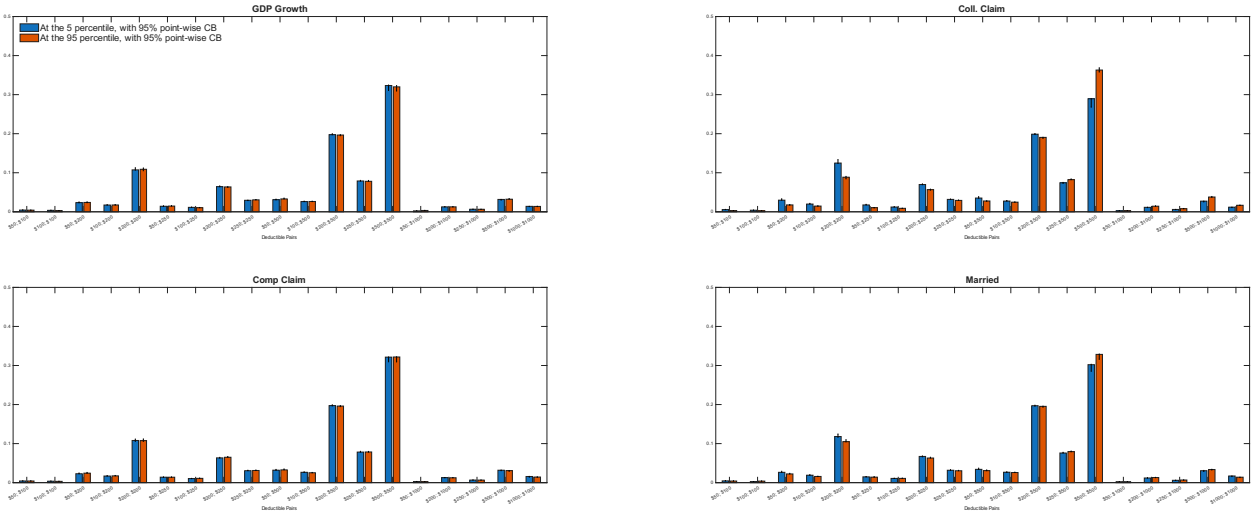


Figure A3: Subjective Beliefs Model: Choice distributions

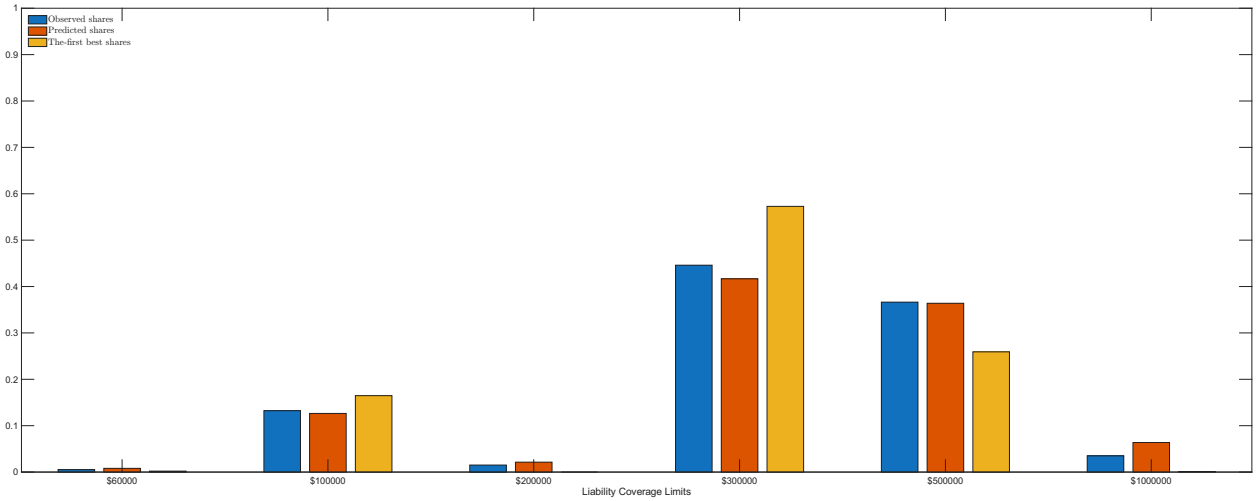


Figure A4: Random Utility Model with Fixed Effects: Choice distributions

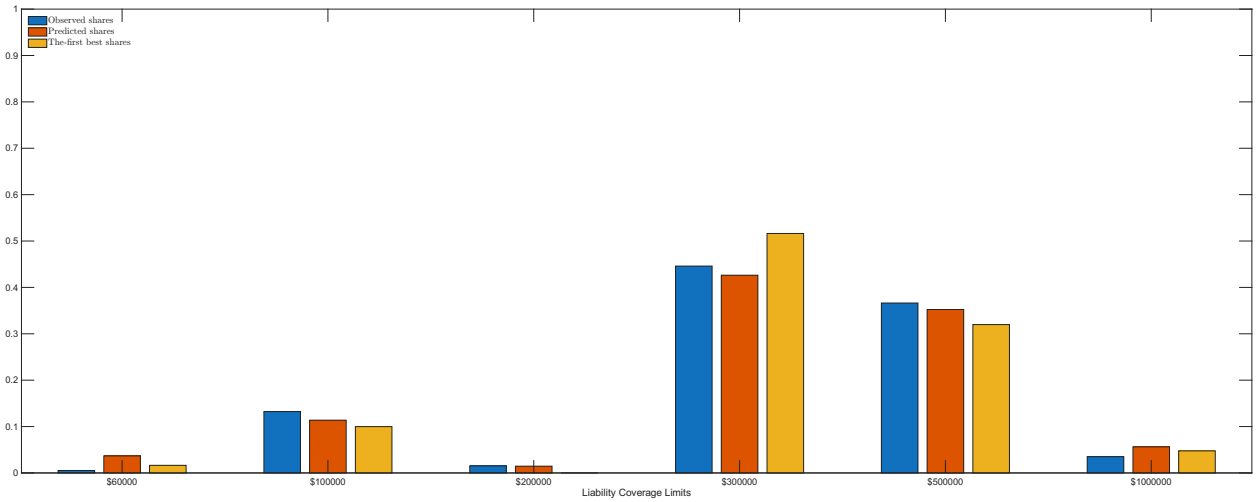


Figure A5: Choice distributions with $\bar{p}^{\mathcal{L}}$ and $\bar{p}^{\mathcal{M}}$ as additional controls

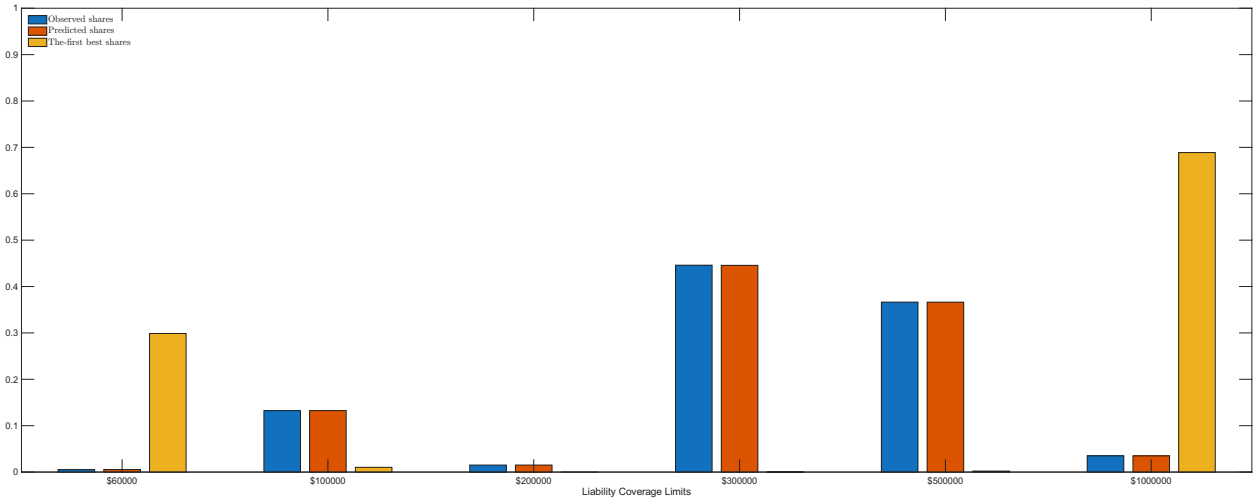


Figure A6: Choice distributions with $\bar{p}^{\mathcal{L}}$, $\bar{p}^{\mathcal{M}}$ and $\bar{p}^{\mathcal{S}}$ as additional controls

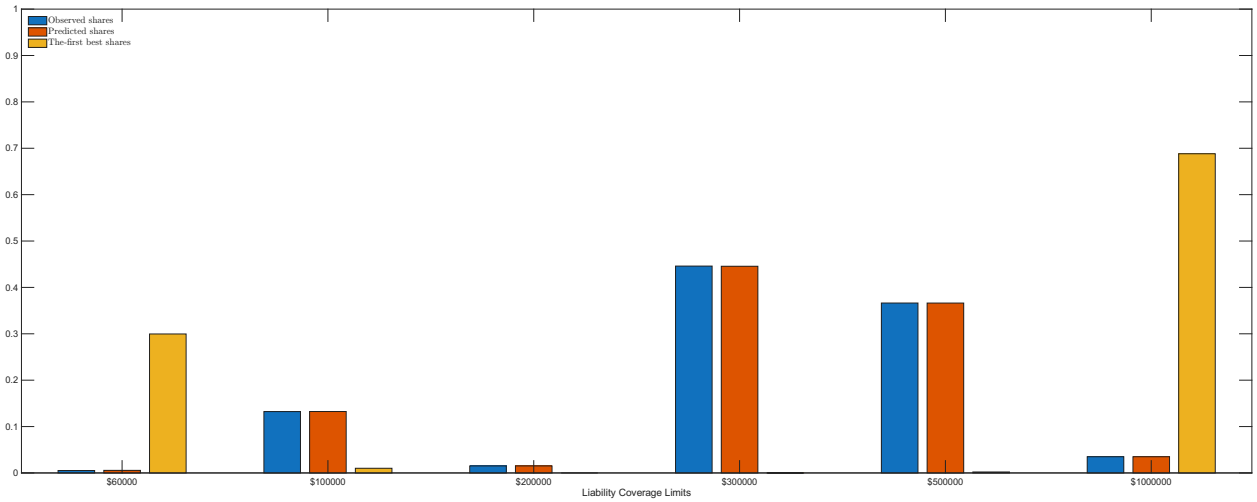


Figure A7: The effect of $\bar{p}^{\mathcal{L}}$, $\bar{p}^{\mathcal{M}}$ and $\lambda^{\mathcal{S}}$

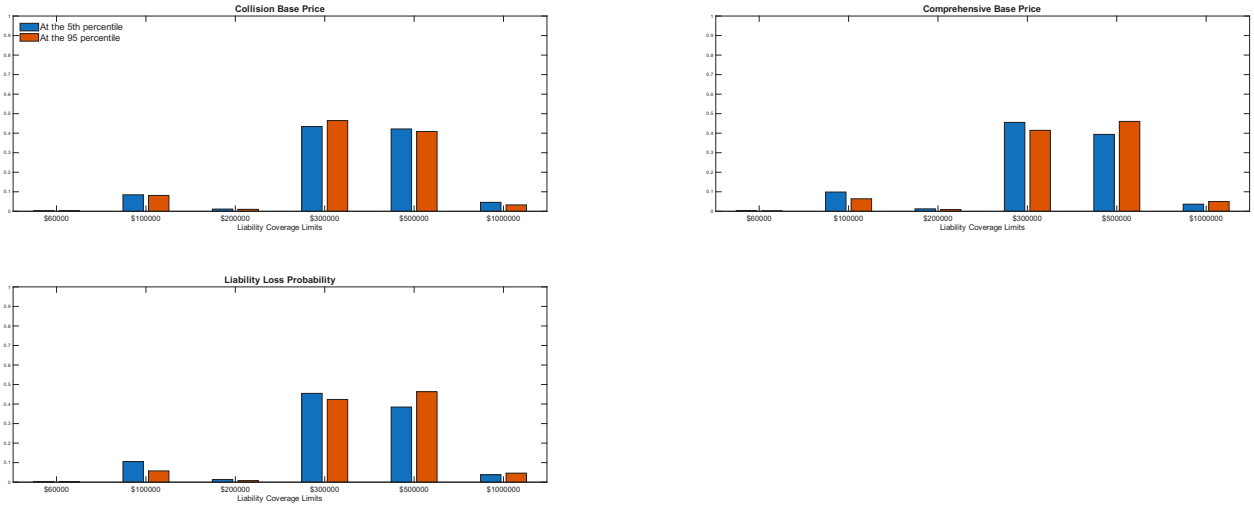


Figure A8: The effect of $\bar{p}^{\mathcal{L}}$, $\bar{p}^{\mathcal{M}}$, $\bar{p}^{\mathcal{S}}$ and $\lambda^{\mathcal{S}}$

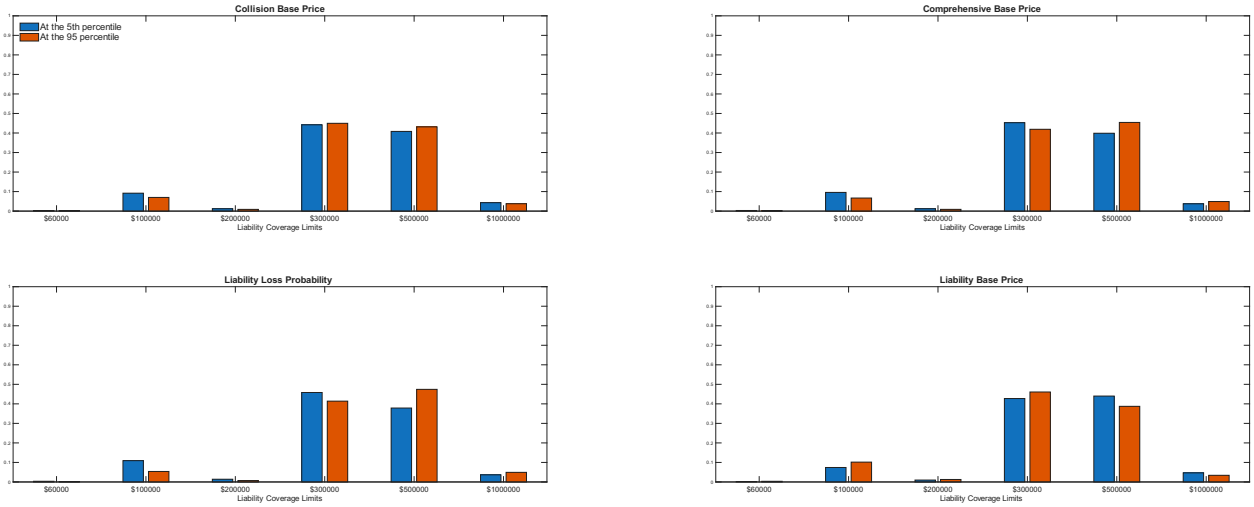


Table A1: Estimates and confidence intervals for consideration and preference models

Variable	Δ_{GDP}	**U.R.**	**R.S.**	**H.V.**	**H.V.**^{2}	**Age**	**Age**^{2}	**Fem.**	**Mar.**	**Year**	**Year**^{2}	**Year**^{3}	**Year**^{4}	*P.2*	*P.3*	$\lambda^{\mathcal{S}}$
Consideration model																
$LB_{95\%}$	0.01	0.30	0.21	1.25	-0.09	0.25	0.19	-0.06	-0.15	1.31	-0.95	-0.82	0.34	-0.1	5.91	-0.08
γ_{ψ}	0.04	0.33	0.23	1.31	-0.08	0.27	0.22	-0.02	-0.10	1.45	-0.82	-0.7	0.42	0.00	6.17	-0.06
$UB_{95\%}$	0.06	0.36	0.26	1.37	-0.06	0.3	0.24	0.02	-0.05	1.59	-0.69	-0.59	0.50	0.10	6.47	-0.03
Preference model																
$LB_{95\%}$	-0.09	-0.45	0.33	0.92	0.17	0.32	0.02	0.02	0.32	0.28	-0.46	-0.28	0.08	0.09	0.14	0.29
γ_{η}	-0.08	-0.43	0.34	0.95	0.2	0.33	0.04	0.04	0.34	0.33	-0.39	-0.24	0.12	0.15	0.20	0.30
$UB_{95\%}$	0.07	-0.41	0.35	0.98	0.22	0.35	0.05	0.06	0.37	0.39	-0.32	-0.19	0.16	0.20	0.26	0.32

Notes: Δ_{GDP} - GDP growth; U.R. - unemployment rate; H.V. - home value; Fem. - female; Mar. - married; P.2 and P.3 - program indicators. All non-binary variables are standardized before taking power. Estimates are bias-corrected and $LB_{95\%}$ and $UB_{95\%}$ are constructed using a 66% subsampled bootstrap procedure with 750 replications.

Table A2: Estimates and confidence intervals for consideration and preference models

Variable	Δ_{GDP}	**U.R.**	**R.S.**	**H.V.**	**H.V.**^{2}	**Age**	**Age**^{2}	**Fem.**	**Mar.**	**Year**	**Year**^{2}	**Year**^{3}	**Year**^{4}	*P.2*	*P.3*	$\lambda^{\mathcal{L}}$	$\lambda^{\mathcal{M}}$
Consideration model																	
$LB_{95\%}$	0.03	-0.15	0.05	0.61	-0.01	-0.57	-0.13	0.17	0.1	1.51	-0.75	-0.26	0.04	-0.17	0.57	0.14	0.19
γ_{ψ}	0.04	-0.12	0.07	0.67	0.01	-0.55	-0.11	0.2	0.13	1.62	-0.67	-0.15	0.09	-0.09	0.7	0.17	0.21
$UB_{95\%}$	0.06	-0.09	0.09	0.71	0.03	-0.52	-0.09	0.22	0.17	1.78	-0.55	-0.07	0.16	-0.04	0.76	0.19	0.23
Preference model																	
$LB_{95\%}$	0.09	0.42	0.01	-0.02	-0.08	0.34	0.48	-0.42	-0.51	-1.59	0.74	0.62	-0.41	0.08	-1.45	-0.42	0.06
γ_{η}	0.11	0.47	0.05	0.12	-0.03	0.36	0.53	-0.38	-0.43	-1.39	0.96	0.76	-0.28	0.32	-0.82	-0.37	0.09
$UB_{95\%}$	0.13	0.53	0.08	0.17	0.15	0.41	0.56	-0.32	-0.37	-1.09	1.20	0.88	-0.17	0.52	-0.6	-0.32	0.13

Notes: Δ_{GDP} - GDP growth; U.R. - unemployment rate; H.V. - home value; Fem. - female; Mar. - married; P.2 and P.3 - program indicators. All non-binary variables are standardized before taking power. Estimates are bias-corrected and $LB_{95\%}$ and $UB_{95\%}$ are constructed using a 66% subsampled bootstrap procedure with 750 replications.

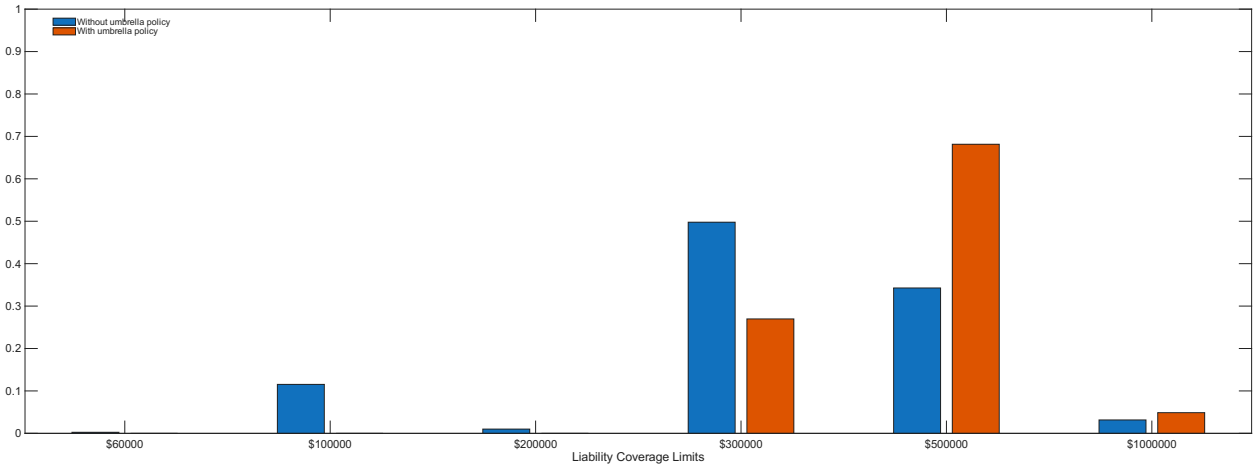
A2 Robustness: Umbrella Policy

We do not observe whether households hold additional excess liability coverage. Such coverage—typically referred to as an "umbrella" policy—is generally offered to high–net-worth consumers who carry multiple personal lines with the insurer (e.g., auto and homeowners) and is sold in $1 million increments. Purchasing an umbrella policy typically requires meeting minimum underlying liability limits on the base policies (e.g., $300,000 auto liability). Households with umbrella coverage have total liability protection that exceeds our highest observed auto limit of $1 million. In our framework, this is consistent with $\mathcal{R}$-type preferences over total liability protection. At the same time, because the umbrella sits on top of the underlying auto limit, these households may appear in our data as choosing lower auto limits (e.g., $300,000 or $500,000) even though their effective total coverage is higher (e.g., $1,300,000 or $1,500,000). This behavior can be interpreted as a form of limited consideration. Below we investigate its quantitative relevance.

First, out of an abundance of caution, and following an approach similar to Cohen and Einav (2007), we re-estimate our benchmark model on a coarsened version of the data in which we combine all households with liability limits above $300,000 into a single "high coverage" group and treat $300,000 as the highest available limit. By construction, all households who might potentially hold an umbrella policy are in this group. We again find that preferences are extreme and that choices are not sensitive to base prices. Ex post, this is unsurprising: in the data, the correlation between these coarsened liability choices and the base price is negligible (but *positive* 0.009).

The remaining issue is whether, conditional on choosing a limit of $300,000 or above, households with umbrella coverage select lower underlying auto limits (and face lower base prices) than households without umbrella coverage. If so, there could be price sensitivity that our data cannot capture. Figure A9 shows the distribution of liability limits in the Barseghyan et al. (2018) sample (2,960 households making first-time purchases sometime between 2004 and 2007, within six months of purchasing homeowners coverage), separately

Figure A9: Choice distributions



for households with and without umbrella policies. Umbrella households choose substantially higher underlying auto limits on average, even conditional on choosing $300,000 or above. Moreover, they face a slightly higher base price ($258 versus $250).

In sum, even if a nontrivial share of households in our data hold umbrella coverage, we find no evidence that this omission materially affects our results.

A3 Identification

A3.1 The Set Up

In this section we develop an argument for semi-non parametric identification of the model with a generic preference distribution function and a generic specification of $\psi_j^{\mathcal{L}}(\eta; \mathbf{Z})$ and $\psi_k^{\mathcal{M}}(\eta; \mathbf{Z})$, with a single normalization that $\psi_1^{\mathcal{L}}(\eta; \mathbf{Z}) = \psi_1^{\mathcal{M}}(\eta; \mathbf{Z})$. The argument requires new notation as well as a formal statement of underlying assumptions. Most of these are borrowed from Barseghyan and Molinari (2023) or Barseghyan et al. (2021a). All statements are conditional on $\mathbf{Z}$ as well as λ's. For clarity of exposition the dependence on $\{\mathbf{Z}, \lambda^{\mathcal{L}}, \lambda^{\mathcal{M}}\}$ is suppressed. Finally, instead of writing $CE(p_j^{\mathcal{L}}, \mathcal{L}_j, \lambda^{\mathcal{L}} | \nu, \Omega(\cdot))$ or $CE(p_j^{\mathcal{M}}, \mathcal{M}_j, \lambda^{\mathcal{M}} | \nu, \Omega(\cdot))$ we will write $CE_\eta(p_j^{\mathcal{L}})$, or $CE_\eta(p_j^{\mathcal{M}})$, respectively. This is unambiguous because the superscript of the price, e.g., $p_j^{\mathcal{L}}$ indicates that the coverage is $\mathcal{L}$, its subscript shows the coverage level, and $\lambda^{\mathcal{L}}$ is taken fixed. It is also without loss of generality since in our setting we keep ν and δ fixed and claim identification for the case when η is the only unobserved heterogeneous preference parameter. The proofs below apply to any model with uni-dimensional unobserved heterogeneity that yield CE satisfying conditions below.

Assumption 4 (The distribution of the preference parameter)**.** *The random preference parameter η is distributed according to a cumulative distribution function $G(\cdot)$ that satisfies the properties of CDFs, and admits a density function $g(\cdot)$ that is continuous and strictly positive on $[0, \bar{\eta}]$ and is zero everywhere else.*

Assumption 5 (Stability)**.** *The utility function of each DM is* context-invariant*.*

Assumption 6 (Narrow Bracketing, Read et al. 2000)**.** *For a DM with a preference parameter η, the certainty equivalent of the bundle $\mathcal{I}_{j,k} \equiv \{\mathcal{L}_j, \mathcal{M}_k\}$ is equal to $CE_\eta(p_j^{\mathcal{L}}) + CE_\eta(p_k^{\mathcal{M}})$.*

Assumption 7 (Single Crossing Property)**.** *There exists a continuously differentiable and*

strictly monotone function $\boldsymbol{\eta}_a^b : \text{supp}(\bar{p}^{\mathcal{L}}) \to \mathbf{R}$, *such that*

$$\begin{aligned} CE_\eta(p_a^{\mathcal{L}}) < CE_\eta(p_b^{\mathcal{L}}) \quad & \forall \eta < \boldsymbol{\eta}_a^b(\bar{p}^{\mathcal{L}}), \\ CE_\eta(p_a^{\mathcal{L}}) = CE_\eta(p_b^{\mathcal{L}}) \quad & \eta = \boldsymbol{\eta}_a^b(\bar{p}^{\mathcal{L}}), \\ CE_\eta(p_a^{\mathcal{L}}) > CE_\eta(p_b^{\mathcal{L}}) \quad & \forall \eta > \boldsymbol{\eta}_a^b(\bar{p}^{\mathcal{L}}). \end{aligned}$$

where $a < b$. *We refer to* $\boldsymbol{\eta}_a^b(\bar{p}^{\mathcal{L}})$ *as the cutoff between coverage levels* a *and* b *in context* $\mathcal{L}$.

We assume that the SCP holds also in context $\mathcal{M}$ with cutoffs $\boldsymbol{\eta}_c^d(\bar{p}^{\mathcal{M}})$, $c < d$, defined similarly. The SCP implies that *within a single context*, the household's ranking of (any two) coverage levels is monotone in η, yielding vertical differentiation of (any two) coverage levels. Whenever the SCP holds for each context, it also holds for bundles $\mathcal{I}_{a,c}$ and $\mathcal{I}_{b,d}$ such that $a \leqslant b$ and $c \leqslant d$, with at least one inequality being strict. For such bundles, the cutoff $\boldsymbol{\eta}_{a,c}^{b,d}(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})$ is the preference parameter η such that

$$CE_\eta(p_a^{\mathcal{L}}) + CE_\eta(p_c^{\mathcal{M}}) = CE_\eta(p_b^{\mathcal{L}}) + CE_\eta(p_d^{\mathcal{M}}).$$

Note that $\boldsymbol{\eta}_{a,c}^{b,c}(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}}) = \boldsymbol{\eta}_a^b(\bar{p}^{\mathcal{L}})$ and, similarly, $\boldsymbol{\eta}_{a,c}^{a,d}(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}}) = \boldsymbol{\eta}_c^d(\bar{p}^{\mathcal{M}})$. Finally, not only the SCP is a verifiable property, the cutoffs are known functions implicitly defined by the indifference condition between two alternatives.[53]

Assumption 8 (Distinct Contexts)**.** *Contexts* $\mathcal{L}$ *and* $\mathcal{M}$ *are distinct if the following condition holds: If* $\boldsymbol{\eta}_a^b(\bar{p}^{\mathcal{L}}) = \boldsymbol{\eta}_c^d(\bar{p}^{\mathcal{M}})$, *then* $\boldsymbol{\eta}_{a'}^{b'}(\bar{p}^{\mathcal{L}}) \neq \boldsymbol{\eta}_{c'}^{d'}(\bar{p}^{\mathcal{M}})$ *a.e.* $\forall (a', b', c', d') \neq (a, b, c, d)$.

Assumption 9 (Natural Order of the Cutoffs)**.** *For any triplet of alternatives,* $1 < i < j$, $\boldsymbol{\eta}_1^i(\bar{p}^{\mathcal{L}}) < \boldsymbol{\eta}_1^j(\bar{p}^{\mathcal{L}}), \quad \forall \bar{p}^{\mathcal{L}}$, *and* $\boldsymbol{\eta}_1^i(\bar{p}^{\mathcal{M}}) < \boldsymbol{\eta}_1^j(\bar{p}^{\mathcal{M}}), \quad \forall \bar{p}^{\mathcal{M}}$.

The assumption above is equivalent to $\boldsymbol{\eta}_1^i(\cdot) < \boldsymbol{\eta}_i^j(\cdot)$, $i < j$. That is, coverage levels $1 < i < j$ split the parameter space into three regions. DMs with η in the left region would prefer the lowest coverage, those in the middle region – the intermediate coverage, and those

[53] In some cases, one can solve for it explicitly. E.g., the expected utility specification of Cohen and Einav (2007) is linear in risk aversion, and hence the indifference condition between two coverage levels is a linear equation in the risk aversion parameter.

in the right region – the highest coverage. The assumption above is written with the identity of the lowest coverage level fixed at 1 and holds in our data. For any three coverage levels, $k < i < j$, it does not necessarily hold. It is possible that the intermediate coverage level is never the first best. As explained in, e.g., Barseghyan et al. (2016), this can happen depending on how prices are set, and indeed occurs in our data for context $\mathcal{L}$: the \$200 deductible can only be the second best because it is always dominated either by the \$250 or by the \$100 deductible.

Assumption 10 (No Three-Way-Ties, The First and Second Bests)**.** *At all base prices:*

- *In each context, for any triplet of coverage levels, $k < i < j$, $\boldsymbol{\eta}_k^i(\cdot) \neq \boldsymbol{\eta}_i^j(\cdot)$.*
- *In context $\mathcal{M}$, every coverage level is potentially the first best: $\forall i$, there exists η s.t. $i = \arg\max_j CE_\eta(p_j^{\mathcal{M}})$.*
- *In context $\mathcal{L}$, every coverage level is potentially the first best except for $\hat{l} \geqslant 3$, which can be at most the second best: there exists η s.t. $\hat{l} = \arg\max_{j \in \{1,\ldots,l\} \setminus l^{\max}} CE_\eta(p_j^{\mathcal{L}})$ where $l^{\max} = \arg\max_j CE_\eta(p_j^{\mathcal{L}})$.*

The first part of this assumption requires that no DM is indifferent between three coverage levels at a single base price.[54] The second part states that in context $\mathcal{M}$ the coverage levels are well behaved. The third part allows for a never-the-first-best alternative in context $\mathcal{L}$.

Assumption 11 (Independence and Large Support)**.** *For each $\eta \in [0, \bar{\eta}]$, the random vector $(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})$ has a strictly positive density on an open ball with radius $\epsilon > 0$ centered at each of the following values:*

A. $(\bar{p}_c^{\mathcal{L}}, \bar{p}_c^{\mathcal{M}})$ s.t. $\boldsymbol{\eta}_1^2(\bar{p}_c^{\mathcal{L}}) = \boldsymbol{\eta}_i^{i+1}(\bar{p}_c^{\mathcal{M}})$, $\forall i$;

B. $(\bar{p}_c^{\mathcal{L}}, \bar{p}_c^{\mathcal{M}})$ s.t. $\boldsymbol{\eta}_i^{i+1}(\bar{p}_c^{\mathcal{L}}) = \boldsymbol{\eta}_i^{i+1}(\bar{p}_c^{\mathcal{M}})$, $\forall i$, and $(\bar{p}_c^{\mathcal{L}}, \bar{p}_c^{\mathcal{M}})$ s.t. $\boldsymbol{\eta}_{\hat{l}-1}^{\hat{l}+1}(\bar{p}_c^{\mathcal{L}}) = \boldsymbol{\eta}_{\hat{l}-1}^{\hat{l}+1}(\bar{p}_c^{\mathcal{M}})$.

[54] Algebraically, this amounts to stating that a system of two equations, $CE_\eta(p_k^{\mathcal{L}}) = CE_\eta(p_i^{\mathcal{L}}) = CE_\eta(p_j^{\mathcal{L}})$, in one unknown, η, has no solution.

Definition 3 (Consideration Probabilities)**.** *For any $\mathcal{K}_1, \mathcal{K}_2 \subseteq \mathcal{L} \otimes \mathcal{M}$, $\mathcal{K}_1 \cap \mathcal{K}_2 = \varnothing$, the probability that all elements of $\mathcal{K}_1$ are included in the consideration set while all elements of $\mathcal{K}_2$ are excluded from it is denoted by*

$$\mathcal{O}_\eta(\mathcal{K}_1; \mathcal{K}_2) \equiv \sum_{\mathcal{K}:\ \mathcal{K}_1 \subset \mathcal{K},\ \mathcal{K}_2 \cap \mathcal{K} = \varnothing} \mathcal{Q}_\eta(\mathcal{K}),$$

where $\mathcal{Q}_\eta(\mathcal{K})$ is the probability that DMs with the preference parameter η draw consideration set $\mathcal{K}$.

Assumption 12 (Consideration is continuous in η)**.** *All consideration probabilities $\mathcal{Q}_\eta(\cdot)$ vary continuously with the preference parameter η.*

A3.2 Identification: Two coverage levels

Suppose there are only two coverage levels per context. Since under the $ARC^{\triangle}$ model the bundle $\mathcal{I}_{2,1}$ is never considered, only these three bundles, $\{\mathcal{I}_{1,1}, \mathcal{I}_{1,2}, \mathcal{I}_{2,2}\}$, can be chosen. The SCP implies that the cutoff $\boldsymbol{\eta}^{2,2}_{1,1}(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})$ must lie between $\boldsymbol{\eta}^{2}_{1}(\bar{p}^{\mathcal{L}})$ and $\boldsymbol{\eta}^{2}_{1}(\bar{p}^{\mathcal{M}})$, or, equivalently, between $\boldsymbol{\eta}^{2,1}_{1,1}(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})$ and $\boldsymbol{\eta}^{1,2}_{1,1}(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})$.

Figure A10: Preference order of bundles $\{\mathcal{I}_{1,1}, \mathcal{I}_{1,2}, \mathcal{I}_{2,2}\}$, $\boldsymbol{\eta}^{2,1}_{1,1} < \boldsymbol{\eta}^{2,2}_{1,1} < \boldsymbol{\eta}^{1,2}_{1,1}$.

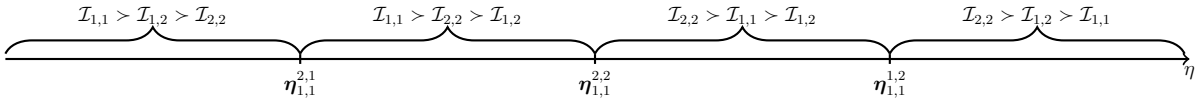
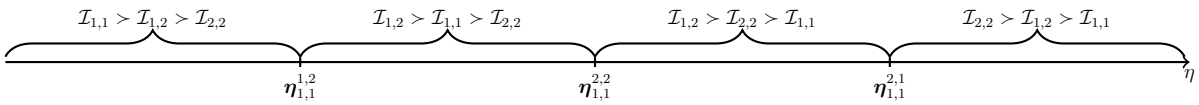


Figure A11: Preference order of bundles $\{\mathcal{I}_{1,1}, \mathcal{I}_{1,2}, \mathcal{I}_{2,2}\}$, $\boldsymbol{\eta}^{1,2}_{1,1} < \boldsymbol{\eta}^{2,2}_{1,1} < \boldsymbol{\eta}^{2,1}_{1,1}$.



Let $\boldsymbol{\eta}^{2,1}_{1,1}(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}}) < \boldsymbol{\eta}^{2,2}_{1,1}(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}}) < \boldsymbol{\eta}^{1,2}_{1,1}(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})$. Then the cutoffs split the real line into four intervals, each with a unique preference order over bundles, as depicted in Figure A10. In the first two regions (from the left), $\mathcal{I}_{1,1}$ is chosen whenever it is considered. In the third

region, $\mathcal{I}_{1,1}$ is chosen whenever it is considered and $\mathcal{I}_{2,2}$ is not. In the fourth region, $\mathcal{I}_{1,1}$ is chosen only when it is considered and neither $\mathcal{I}_{1,2}$ nor $\mathcal{I}_{2,2}$ is considered. Consider how the share of bundle $\mathcal{I}_{1,1}$ responds to price changes:

$$\frac{\partial \Pr(\mathcal{I}_{1,1}|\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})}{\partial \bar{p}^{\mathcal{L}}} = \left[\mathcal{O}_\eta(\{\mathcal{I}_{1,1}, \mathcal{I}_{2,2}\}; \varnothing) g(\eta)\right]_{\eta=\boldsymbol{\eta}_{1,1}^{2,2}} \frac{\partial \boldsymbol{\eta}_{1,1}^{2,2}}{\partial \bar{p}^{\mathcal{L}}}.$$

$$\frac{\partial \Pr(\mathcal{I}_{1,1}|\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})}{\partial \bar{p}^{\mathcal{M}}} = \left[\mathcal{O}_\eta(\{\mathcal{I}_{1,1}, \mathcal{I}_{2,2}\}; \varnothing) g(\eta)\right]_{\eta=\boldsymbol{\eta}_{1,1}^{2,2}} \frac{\partial \boldsymbol{\eta}_{1,1}^{2,2}}{\partial \bar{p}^{\mathcal{M}}} + \left[\mathcal{O}_\eta(\{\mathcal{I}_{1,1}, \mathcal{I}_{1,2}\}; \mathcal{I}_{2,2}) g(\eta)\right]_{\eta=\boldsymbol{\eta}_{1,1}^{1,2}} \frac{\partial \boldsymbol{\eta}_{1,1}^{1,2}}{\partial \bar{p}^{\mathcal{M}}}.$$

Similarly, the share of bundle $\mathcal{I}_{1,2}$ responds to price changes as follows:

$$\frac{\partial \Pr(\mathcal{I}_{1,2}|\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})}{\partial \bar{p}^{\mathcal{L}}} = \left[\mathcal{O}_\eta(\{\mathcal{I}_{1,2}, \mathcal{I}_{2,2}\}; \mathcal{I}_{1,1}) g(\eta)\right]_{\eta=\boldsymbol{\eta}_{1,1}^{2,1}} \frac{\partial \boldsymbol{\eta}_{1,1}^{2,1}}{\partial \bar{p}^{\mathcal{L}}}$$

If instead the cutoff ordering is reversed, as in Figure A11, i.e. $\boldsymbol{\eta}_{1,1}^{1,2}(\cdot) < \boldsymbol{\eta}_{1,1}^{2,2}(\cdot) < \boldsymbol{\eta}_{1,1}^{2,1}(\cdot)$, analogous derivative expressions obtain. In particular,

$$\frac{\partial \Pr(\mathcal{I}_{1,1}|\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})}{\partial \bar{p}^{\mathcal{L}}} = \left[\mathcal{O}_\eta(\{\mathcal{I}_{1,1}, \mathcal{I}_{2,2}\}; \mathcal{I}_{1,2}) g(\eta)\right]_{\eta=\boldsymbol{\eta}_{1,1}^{2,2}} \frac{\partial \boldsymbol{\eta}_{1,1}^{2,2}}{\partial \bar{p}^{\mathcal{L}}}.$$

Fix a base-price vector $(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})$ at which the single-context indifference points coincide, i.e. $\boldsymbol{\eta}_1^2(\bar{p}^{\mathcal{L}}) = \boldsymbol{\eta}_1^2(\bar{p}^{\mathcal{M}})$. Consider local one-sided perturbations $(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}}_{+o})$ and $(\bar{p}^{\mathcal{L}}_{+o}, \bar{p}^{\mathcal{M}})$ and take limits as $o \searrow 0$. Under these perturbations, the derivative identities displayed above (those under the ordering in Figure A10, together with the one under the ordering in Figure A11) identify the following objects up to scale $g(\eta)$:

$$\mathcal{O}_\eta(\{\mathcal{I}_{1,1}, \mathcal{I}_{2,2}\}; \varnothing),\ \mathcal{O}_\eta(\{\mathcal{I}_{1,1}, \mathcal{I}_{2,2}\}; \mathcal{I}_{1,2}),\ \mathcal{O}_\eta(\{\mathcal{I}_{1,1}, \mathcal{I}_{1,2}\}; \mathcal{I}_{2,2}),\ \text{and}\ \mathcal{O}_\eta(\{\mathcal{I}_{1,2}, \mathcal{I}_{2,2}\}; \mathcal{I}_{1,1}).$$

For the $ARC^{\triangle}$ Model, the quantities translate to:

$$\Theta(\eta)\Psi_{1,1}(\eta)\Psi_{2,2}(\eta);$$

$$\Theta(\eta)\Psi_{1,1}(\eta)\Psi_{2,2}(\eta)(1-\Psi_{1,2}(\eta));$$

$$\Theta(\eta)\Psi_{1,1}(\eta)\Psi_{1,2}(\eta)(1-\Psi_{2,2}(\eta));$$

$$\Theta(\eta)\Psi_{1,2}(\eta)\Psi_{2,2}(\eta)(1-\Psi_{1,1}(\eta)).$$

where $\Theta(\eta)$ depending on the completion rule is either one or $\Theta(\eta) = \left[1 - \prod_{i \leqslant j}\left(1 - \Psi_{i,j}\right)\right]^{-1}$. Identification follows in four steps:

1. $\frac{\Psi_{1,1}(\eta)\cdot\Psi_{2,2}(\eta)\cdot(1-\Psi_{1,2}(\eta))}{\Psi_{1,1}(\eta)\cdot\Psi_{2,2}(\eta)} = 1 - \Psi_{1,2}(\eta)$ which identifies $\Psi_{1,2}(\eta)$;

2. $\frac{\Psi_{1,1}(\eta)\cdot\Psi_{2,2}(\eta)}{\Psi_{1,1}(\eta)\cdot\Psi_{1,2}(\eta)\cdot(1-\Psi_{2,2}(\eta))} = \frac{\Psi_{2,2}(\eta)}{\Psi_{1,2}(\eta)\cdot(1-\Psi_{2,2}(\eta))}$; which identifies $\Psi_{2,2}(\eta)$;

3. $\frac{\Psi_{1,1}(\eta)\cdot\Psi_{2,2}(\eta)}{\Psi_{1,2}(\eta)\cdot\Psi_{2,2}(\eta)\cdot(1-\Psi_{1,1}(\eta))} = \frac{\Psi_{1,1}(\eta)}{\Psi_{1,2}(\eta)\cdot(1-\Psi_{1,1}(\eta))}$ which identifies $\Psi_{1,1}(\eta)$;

4. Once the Ψ's are known, so is $\Theta(\eta)g(\eta)$ and consequently $\Theta(\eta)$ and $g(\eta)$.

The argument gives rise to the following Theorem:

Theorem 1. *Suppose Assumptions 4 to 12 hold and there are only two coverage levels per context. Then, $\Psi_{1,1}(\eta)$, $\Psi_{1,2}(\eta)$, $\Psi_{2,2}(\eta)$, and $g(\eta)$, are identified $\forall\eta \in [0,\bar{\eta}]$. Hence, the model is identified.*

A3.3 Identification: Multiple coverage levels

We extend the argument to more than two coverage levels via a sequence of results.

Lemma 1. *Suppose Assumptions 7, 9, and 10 hold. Then, $\min\{\boldsymbol{\eta}_1^3(\bar{p}^{\mathcal{L}}), \boldsymbol{\eta}_1^3(\bar{p}^{\mathcal{M}})\} \leqslant \boldsymbol{\eta}_{1,1}^{i,j}(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})$, $\forall(i,j)$ s.t. $j \geqslant i \geqslant 3$. Similarly, $\min\{\boldsymbol{\eta}_1^3(\bar{p}^{\mathcal{L}}), \boldsymbol{\eta}_2^3(\bar{p}^{\mathcal{M}})\} \leqslant \boldsymbol{\eta}_{1,2}^{i,j}(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})$, $\forall(i,j)$ s.t. $j \geqslant i \geqslant 3$.*

Proof. It must be the case that $\boldsymbol{\eta}_{1,1}^{i,j}(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})$ lies between $\boldsymbol{\eta}_1^i(\bar{p}^{\mathcal{L}})$ and $\boldsymbol{\eta}_1^j(\bar{p}^{\mathcal{M}})$. Since $\boldsymbol{\eta}_1^3(\bar{p}^{\mathcal{L}}) < \boldsymbol{\eta}_1^i(\bar{p}^{\mathcal{L}})$ and $\boldsymbol{\eta}_1^3(\bar{p}^{\mathcal{M}}) < \boldsymbol{\eta}_1^j(\bar{p}^{\mathcal{M}})$, the first statement follows. Next, it must be the case that $\boldsymbol{\eta}_{1,2}^{i,j}(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})$ lies between $\boldsymbol{\eta}_1^i(\bar{p}^{\mathcal{L}})$ and $\boldsymbol{\eta}_2^j(\bar{p}^{\mathcal{M}})$. Since $\boldsymbol{\eta}_1^3(\bar{p}^{\mathcal{L}}) < \boldsymbol{\eta}_1^i(\bar{p}^{\mathcal{L}})$, and $\boldsymbol{\eta}_2^3(\bar{p}^{\mathcal{M}}) < \boldsymbol{\eta}_2^j(\bar{p}^{\mathcal{M}})$, the second statement follows.

□

Lemma 2. *Suppose Assumptions 4 to 12 hold. Then, there exists $\eta_0 < \bar{\eta}$ such that $\Psi_{1,1}(\eta)$, $\Psi_{1,2}(\eta)$, $\Psi_{2,2}(\eta)$, and $\Theta(\eta)g(\eta)$ are identified $\forall\eta \in [\eta_0,\bar{\eta}]$.*

Proof. By Assumption 9, in each context, $\boldsymbol{\eta}_1^2 \leqslant \boldsymbol{\eta}_1^3 < \boldsymbol{\eta}_1^j < \boldsymbol{\eta}_2^j$ for all $j \geqslant 3$. Consider $(\bar{p}_0^{\mathcal{L}}, \bar{p}_0^{\mathcal{M}})$ such that $\boldsymbol{\eta}_{1,1}^{3,3}(\bar{p}_0^{\mathcal{L}}, \bar{p}_0^{\mathcal{M}}) = \boldsymbol{\eta}_1^3(\bar{p}_0^{\mathcal{L}}) = \boldsymbol{\eta}_1^3(\bar{p}_0^{\mathcal{M}}) = \bar{\eta}$ and $\eta_0 = \max\{\boldsymbol{\eta}_1^2(\bar{p}_0^{\mathcal{L}}), \boldsymbol{\eta}_1^2(\bar{p}_0^{\mathcal{M}})\}$. For any pair $(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}}) > (\bar{p}_0^{\mathcal{L}}, \bar{p}_0^{\mathcal{M}})$. bundles $\mathcal{I}_{1,1}, \mathcal{I}_{1,2}, \mathcal{I}_{2,1}$ and $\mathcal{I}_{2,2}$ dominate all other bundles on the entire support. It follows that on interval $[\eta_0,\bar{\eta}]$, the argument of Theorem 1 applies and $\Psi_{1,1}(\eta)$, $\Psi_{1,2}(\eta)$, $\Psi_{2,2}(\eta)$, and $\Theta(\eta)g(\eta)$ are identified. □

Theorem 2. *Suppose Assumptions 4-12 hold. Then, the following objects are identified* $\forall \eta$*:*

1. $\Psi_{1,i}(\eta)\Psi_{2,i+1}(\eta)\Psi_{1,i+1}(\eta)\Theta(\eta)g(\eta)$, $\forall i$,
2. $\Psi_{i,i}(\eta)\Psi_{i,i+1}(\eta)\Psi_{i+1,i+1}(\eta)\Theta(\eta)g(\eta)$, $\forall i \notin \{\hat{l}-1, \hat{l}\}$,
3. $\Psi_{\hat{l}-1,\hat{l}-1}(\eta)\Psi_{\hat{l}+1,\hat{l}+1}(\eta)\Psi_{\hat{l}-1,\hat{l}+1}(\eta)\Theta(\eta)g(\eta)$.

Proof.

1. Fix an η.

2.a Consider a pair of base prices $\bar{p}^{\mathcal{L}}$ and $\bar{p}^{\mathcal{M}}$ such that the DM endowed with this η is indifferent between bundles $\{\mathcal{I}_{1,i}, \mathcal{I}_{1,i+1}, \mathcal{I}_{2,i+1}\}$ and prefers them to all others.

2.b Consider a pair of base prices $\bar{p}^{\mathcal{L}}$ and $\bar{p}^{\mathcal{M}}$ such that the DM endowed with this η is indifferent between bundles $\{\mathcal{I}_{i,i}, \mathcal{I}_{i,i+1}, \mathcal{I}_{i+1,i+1}\}$ and prefers them to all others.

2.c Consider a pair of base prices $\bar{p}^{\mathcal{L}}$ and $\bar{p}^{\mathcal{M}}$ such that the DM endowed with η is indifferent between bundles $\{\mathcal{I}_{\hat{l}-1,\hat{l}-1}, \mathcal{I}_{\hat{l}-1,\hat{l}+1}, \mathcal{I}_{\hat{l}+1,\hat{l}+1}\}$ and prefers them to all others.

3 For each of the cases, 2.a, 2.b and 2.c., above repeat the steps of the proof of Theorem 3.1 in Barseghyan and Molinari (2023).

4 Repeat for all η.

□

Definition 4. *Let* $\mathcal{B}^{b,d}_{a,c}(\eta)$ *be the collection of all alternatives strictly preferred to* $\mathcal{I}_{a,c}$ *and* $\mathcal{I}_{b,d}$ *by the DM with the preference parameter* η *at the base price vector* $(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})$ *s.t.* $\boldsymbol{\eta}^b_a(\bar{p}^{\mathcal{L}}) = \boldsymbol{\eta}^d_c(\bar{p}^{\mathcal{M}})$.

Definition 5. *Let* $\mathcal{BB}^{b,d}_{a,c}(\eta)$ *be the collection of all alternatives strictly preferred to* $\mathcal{I}_{a,c}$ *and* $\mathcal{I}_{b,d}$ *by the DM with the preference parameter* η *at the base price vector* $(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})$ *s.t.* $\eta = \boldsymbol{\eta}^{b,d}_{a,c}(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})$.

In words $\mathcal{B}_{a,c}^{b,d}(\eta)$ is the set of bundles that dominate $\mathcal{I}_{a,c}$ and $\mathcal{I}_{b,d}$ for a DM indifferent between alternatives $\mathcal{L}_a$ and $\mathcal{L}_b$ **and** alternatives $\mathcal{M}_c$ and $\mathcal{M}_d$. Such a DM may or may not exist, depending on the base prices. $\mathcal{BB}_{a,c}^{b,d}(\eta)$ is the set of bundles that dominate $\mathcal{I}_{a,c}$ and $\mathcal{I}_{b,d}$ for a DM that is indifferent between bundles $\mathcal{I}_{a,c}$ and $\mathcal{I}_{b,d}$. Such a DM exists at all base prices. Whenever $\mathcal{B}_{a,c}^{b,d}(\eta)$ exists, it coincides with $\mathcal{BB}_{a,c}^{b,d}(\eta)$.

Lemma 3. *Suppose Assumptions 4 to 12 hold. Then,*

$$\left[\prod_{\mathcal{I}_{j,k}\in\mathcal{B}_{i,i}^{i+1,i+1}(\eta)}(1-\Psi_{j,k}(\eta))\right]\Psi_{i,i}(\eta)\Psi_{i+1,i+1}(\eta)\Psi_{i,i+1}(\eta)\Theta(\eta)g(\eta),\quad\forall i,$$

are identified.

Lemma 4. *If $\Psi_{1,1}(\eta)$, $\Psi_{1,2}(\eta)$, $\Psi_{2,2}(\eta)$, $\Theta(\eta)g(\eta)$, and $\Psi_{1,i}(\eta)\Psi_{2,i+1}(\eta)\Psi_{1,i+1}(\eta)\,\Theta(\eta)g(\eta)$ for all i, are identified, then so are $\Psi_{1,i}(\eta)$ and $\Psi_{2,i}(\eta)$ for all i.*

Proof. By $ARC^{\triangle}$, $\Psi_{j,k}(\eta)=\psi_j^{\mathcal{L}}(\eta)\psi_k^{\mathcal{M}}(\eta)$ for all relevant (j,k). Under the normalization $\psi_1^{\mathcal{L}}(\eta)=\psi_1^{\mathcal{M}}(\eta)$ we have

$$\Psi_{1,1}(\eta)=\psi_1^{\mathcal{L}}(\eta)\psi_1^{\mathcal{M}}(\eta)=\left(\psi_1^{\mathcal{L}}(\eta)\right)^2,$$

so $\psi_1^{\mathcal{L}}(\eta)$ (and hence $\psi_1^{\mathcal{M}}(\eta)$) is identified. Then

$$\Psi_{1,2}(\eta)=\psi_1^{\mathcal{L}}(\eta)\psi_2^{\mathcal{M}}(\eta)\quad\Rightarrow\quad\psi_2^{\mathcal{M}}(\eta)=\frac{\Psi_{1,2}(\eta)}{\psi_1^{\mathcal{L}}(\eta)},$$

and

$$\Psi_{2,2}(\eta)=\psi_2^{\mathcal{L}}(\eta)\psi_2^{\mathcal{M}}(\eta)\quad\Rightarrow\quad\psi_2^{\mathcal{L}}(\eta)=\frac{\Psi_{2,2}(\eta)}{\psi_2^{\mathcal{M}}(\eta)}.$$

Next, for each i define the identified object

$$H_i(\eta)\equiv\Psi_{1,i}(\eta)\Psi_{2,i+1}(\eta)\Psi_{1,i+1}(\eta)\,\Theta(\eta)g(\eta).$$

Since $\Theta(\eta)g(\eta)$ is identified, so is $H_i(\eta)/(\Theta(\eta)g(\eta))$. Moreover,

$$\begin{aligned}\frac{H_i(\eta)}{\Theta(\eta)g(\eta)}&=\Psi_{1,i}(\eta)\Psi_{2,i+1}(\eta)\Psi_{1,i+1}(\eta)\\&=\left(\psi_1^{\mathcal{L}}(\eta)\psi_i^{\mathcal{M}}(\eta)\right)\left(\psi_2^{\mathcal{L}}(\eta)\psi_{i+1}^{\mathcal{M}}(\eta)\right)\left(\psi_1^{\mathcal{L}}(\eta)\psi_{i+1}^{\mathcal{M}}(\eta)\right)\\&=\left(\psi_1^{\mathcal{L}}(\eta)\right)^2\psi_2^{\mathcal{L}}(\eta)\,\psi_i^{\mathcal{M}}(\eta)\left(\psi_{i+1}^{\mathcal{M}}(\eta)\right)^2.\end{aligned}$$

For $i = 2$, all terms on the RHS are identified except $\psi_3^{\mathcal{M}}(\eta)$, so $\psi_3^{\mathcal{M}}(\eta)$ is identified. Inductively, given $\psi_i^{\mathcal{M}}(\eta)$, the same display identifies $\psi_{i+1}^{\mathcal{M}}(\eta)$ for $i \geqslant 2$.

Finally, for each i,

$$\Psi_{1,i}(\eta) = \psi_1^{\mathcal{L}}(\eta)\psi_i^{\mathcal{M}}(\eta) \quad \text{and} \quad \Psi_{2,i}(\eta) = \psi_2^{\mathcal{L}}(\eta)\psi_i^{\mathcal{M}}(\eta),$$

so $\Psi_{1,i}(\eta)$ and $\Psi_{2,i}(\eta)$ are identified for all i. □

Lemma 5. *If the following objects are identified:*

1. $\Psi_{1,i}(\eta),\ \Psi_{2,i}(\eta),\ \forall i,\ \Theta(\eta)g(\eta),$
2. $\Psi_{i,i}(\eta)\Psi_{i+1,i+1}(\eta)\Psi_{i,i+1}(\eta)\Theta(\eta)g(\eta),\ \forall i \notin \{\hat{l}-1, \hat{l}\},$
3. $\Psi_{\hat{l}-1,\hat{l}-1}(\eta)\Psi_{\hat{l}+1,\hat{l}+1}(\eta)\Psi_{\hat{l}-1,\hat{l}+1}(\eta)\Theta(\eta)g(\eta),$
4. *as well as*

$$\left[\prod_{\mathcal{I}_{j,k} \in \mathcal{B}_{\hat{l}-1,\hat{l}-1}^{\hat{l},\hat{l}}(\eta)} (1 - \Psi_{j,k}(\eta))\right] \Psi_{\hat{l}-1,\hat{l}-1}(\eta)\Psi_{\hat{l},\hat{l}}(\eta)\Psi_{\hat{l}-1,\hat{l}}(\eta)\Theta(\eta)g(\eta),$$

then $\Psi_{i,j}(\eta),\ \forall i \leqslant j$, *are identified.*

Proof. Under $ARC^{\triangle}$ we have $\Psi_{j,k}(\eta) = \psi_j^{\mathcal{L}}(\eta)\psi_k^{\mathcal{M}}(\eta)$. From $\Psi_{1,1}(\eta) = (\psi_1^{\mathcal{L}}(\eta))^2$ and the normalization $\psi_1^{\mathcal{L}}(\eta) = \psi_1^{\mathcal{M}}(\eta)$, $\psi_1^{\mathcal{L}}(\eta)$ (and hence $\psi_1^{\mathcal{M}}(\eta)$) is identified. Given item 1, for each i,

$$\psi_i^{\mathcal{M}}(\eta) = \frac{\Psi_{1,i}(\eta)}{\psi_1^{\mathcal{L}}(\eta)},$$

so $\{\psi_i^{\mathcal{M}}(\eta)\}_i$ are identified.

Next, fix any $i \notin \{\hat{l}-1, \hat{l}\}$. By item 2, the product

$$S_i(\eta) \equiv \Psi_{i,i}(\eta)\Psi_{i,i+1}(\eta)\Psi_{i+1,i+1}(\eta)\,\Theta(\eta)g(\eta)$$

is identified, and since $\Theta(\eta)g(\eta)$ is identified (item 1), so is $S_i(\eta)/(\Theta(\eta)g(\eta))$. Using the factorization of Ψ,

$$\begin{aligned}\frac{S_i(\eta)}{\Theta(\eta)g(\eta)} &= \Psi_{i,i}(\eta)\Psi_{i,i+1}(\eta)\Psi_{i+1,i+1}(\eta)\\ &= \big(\psi_i^{\mathcal{L}}(\eta)\psi_i^{\mathcal{M}}(\eta)\big)\big(\psi_i^{\mathcal{L}}(\eta)\psi_{i+1}^{\mathcal{M}}(\eta)\big)\big(\psi_{i+1}^{\mathcal{L}}(\eta)\psi_{i+1}^{\mathcal{M}}(\eta)\big)\\ &= \big(\psi_i^{\mathcal{L}}(\eta)\big)^2\,\psi_{i+1}^{\mathcal{L}}(\eta)\,\psi_i^{\mathcal{M}}(\eta)\,\big(\psi_{i+1}^{\mathcal{M}}(\eta)\big)^2.\end{aligned}$$

Since $\psi_i^{\mathcal{M}}(\eta)$ and $\psi_{i+1}^{\mathcal{M}}(\eta)$ are identified, this identifies

$$\big(\psi_i^{\mathcal{L}}(\eta)\big)^2\,\psi_{i+1}^{\mathcal{L}}(\eta) = \frac{S_i(\eta)}{\Theta(\eta)g(\eta)\,\psi_i^{\mathcal{M}}(\eta)\,(\psi_{i+1}^{\mathcal{M}}(\eta))^2}.$$

Starting from the identified $\psi_1^{\mathcal{L}}(\eta)$, this recursion yields $\psi_2^{\mathcal{L}}(\eta)$, then $\psi_3^{\mathcal{L}}(\eta)$, and so on, for all indices that do not require the excluded cases $i \in \{\hat{l}-1, \hat{l}\}$.

For the missing index $\hat{l}$, item 3 identifies

$$R(\eta) \equiv \Psi_{\hat{l}-1,\hat{l}-1}(\eta)\Psi_{\hat{l}-1,\hat{l}+1}(\eta)\Psi_{\hat{l}+1,\hat{l}+1}(\eta)\,\Theta(\eta)g(\eta),$$

hence $R(\eta)/(\Theta(\eta)g(\eta))$ is identified. Expanding,

$$\begin{aligned}\frac{R(\eta)}{\Theta(\eta)g(\eta)} &= \Psi_{\hat{l}-1,\hat{l}-1}(\eta)\Psi_{\hat{l}-1,\hat{l}+1}(\eta)\Psi_{\hat{l}+1,\hat{l}+1}(\eta)\\ &= \big(\psi_{\hat{l}-1}^{\mathcal{L}}(\eta)\psi_{\hat{l}-1}^{\mathcal{M}}(\eta)\big)\big(\psi_{\hat{l}-1}^{\mathcal{L}}(\eta)\psi_{\hat{l}+1}^{\mathcal{M}}(\eta)\big)\big(\psi_{\hat{l}+1}^{\mathcal{L}}(\eta)\psi_{\hat{l}+1}^{\mathcal{M}}(\eta)\big)\\ &= \big(\psi_{\hat{l}-1}^{\mathcal{L}}(\eta)\big)^2\,\psi_{\hat{l}+1}^{\mathcal{L}}(\eta)\,\psi_{\hat{l}-1}^{\mathcal{M}}(\eta)\,\big(\psi_{\hat{l}+1}^{\mathcal{M}}(\eta)\big)^2,\end{aligned}$$

which is fully known by the previous steps and serves as a consistency/bridge condition around the $\hat{l}$ gap.

Finally, item 4 identifies

$$T(\eta) \equiv \left[\prod_{\mathcal{I}_{j,k}\in\mathcal{B}_{\hat{l}-1,\hat{l}-1}^{\hat{l},\hat{l}}(\eta)} (1-\Psi_{j,k}(\eta))\right]\Psi_{\hat{l}-1,\hat{l}-1}(\eta)\Psi_{\hat{l},\hat{l}}(\eta)\Psi_{\hat{l}-1,\hat{l}}(\eta)\,\Theta(\eta)g(\eta).$$

Dividing by the identified $\Theta(\eta)g(\eta)$ and by the identified factor $\prod_{\mathcal{I}_{j,k}\in\mathcal{B}_{\hat{l}-1,\hat{l}-1}^{\hat{l},\hat{l}}(\eta)}(1-\Psi_{j,k}(\eta))$ (since all $\Psi_{j,k}(\eta)$ in the product are identified from the previous steps), we identify

$\Psi_{\hat{l}-1,\hat{l}-1}(\eta)\Psi_{\hat{l},\hat{l}}(\eta)\Psi_{\hat{l}-1,\hat{l}}(\eta)$. Expanding and using identified $\psi^{\mathcal{L}}_{\hat{l}-1}(\eta)$, $\psi^{\mathcal{M}}_{\hat{l}-1}(\eta)$, and $\psi^{\mathcal{M}}_{\hat{l}}(\eta)$ yields $\psi^{\mathcal{L}}_{\hat{l}}(\eta)$.

Therefore all $\psi^{\mathcal{L}}_{i}(\eta)$ and $\psi^{\mathcal{M}}_{i}(\eta)$ are identified, and hence $\Psi_{i,j}(\eta) = \psi^{\mathcal{L}}_{i}(\eta)\psi^{\mathcal{M}}_{j}(\eta)$ is identified for all $i \leqslant j$. □

Lemma 6. *Suppose Assumptions 4 to 12 hold. Then, the model is identified on the interval* $[\eta_0, \bar{\eta}]$*, where* η_0 *is defined as in the Proof of Lemma 2.*

Proof. Per Lemma 2 we have an interval $[\eta_0, \bar{\eta}]$ on which $\Psi_{1,1}(\eta)$, $\Psi_{1,2}(\eta)$, $\Psi_{2,2}(\eta)$, and $\Theta(\eta)g(\eta)$ are identified. Per Theorem 2 and Lemma 4, $\Psi_{1,i}(\eta)$ and $\Psi_{2,i}(\eta)$ are identified on $[\eta_0, \bar{\eta}]$. Combining these with Lemma 5 yields identification of $\Psi_{i,j}(\eta)$ for all $i \leqslant j$ on $[\eta_0, \bar{\eta}]$. Given $\{\Psi_{i,j}(\eta)\}_{i\leqslant j}$, $\Theta(\eta)$ is known (by the completion rule), so $g(\eta)$ is recovered from $\Theta(\eta)g(\eta)$. □

Theorem 3. *Suppose Assumptions 4 to 12 hold. Then, the model is identified.*

Proof. The proof combines the results established above as follows.

Step 1. Per Lemma 6 we have an interval $[\eta_0, \bar{\eta}]$ on which the model is identified.

Step 2. Consider now an interval $[\eta_1, \bar{\eta}]$ such that $\boldsymbol{\eta}^{3,3}_{1,1}(\bar{p}^{\mathcal{L}}_1, \bar{p}^{\mathcal{M}}_1) = \boldsymbol{\eta}^{3}_{1}(\bar{p}^{\mathcal{L}}_1) = \boldsymbol{\eta}^{3}_{1}(\bar{p}^{\mathcal{M}}_1) = \eta_0$ and $\eta_1 = \max\{\boldsymbol{\eta}^{2}_{1}(\bar{p}^{\mathcal{L}}_1), \boldsymbol{\eta}^{2}_{1}(\bar{p}^{\mathcal{M}}_1)\}$. Recall that the argument behind the proof of Theorem 1 relies on $\frac{\partial \Pr(\mathcal{I}_{1,1}|\bar{p}^{\mathcal{L}},\bar{p}^{\mathcal{M}})}{\partial \bar{p}^{\mathcal{L}}}$, $\frac{\partial \Pr(\mathcal{I}_{1,1}|\bar{p}^{\mathcal{L}},\bar{p}^{\mathcal{M}})}{\partial \bar{p}^{\mathcal{M}}}$, and $\frac{\partial \Pr(\mathcal{I}_{1,2}|\bar{p}^{\mathcal{L}},\bar{p}^{\mathcal{M}})}{\partial \bar{p}^{\mathcal{L}}}$. Taking these derivatives for any base price vector $(\bar{p}^{\mathcal{L}}, \bar{p}^{\mathcal{M}})$, $\bar{p}^{\mathcal{L}} \geqslant \bar{p}^{\mathcal{L}}_1, \bar{p}^{\mathcal{M}} \geqslant \bar{p}^{\mathcal{M}}_1$, returns exactly the same expressions with possible additional terms on the RHS, which consist of evaluating expressions

$$\left[\prod_{\mathcal{I}_{j,k}\in\mathcal{BB}^{q,r}_{1,1}(\eta)} (1-\Psi_{j,k}(\eta))\right]\Psi_{1,1}(\eta)\Psi_{q,r}(\eta)\Theta(\eta)g(\eta),$$

and

$$\left[\prod_{\mathcal{I}_{j,k}\in\mathcal{BB}^{q,r}_{1,2}(\eta)} (1-\Psi_{j,k}(\eta))\right]\Psi_{1,2}(\eta)\Psi_{q,r}(\eta)\Theta(\eta)g(\eta),$$

at values of η exceeding η_0. Since these terms are evaluated at values higher than η_0, they are known. In turn this implies that Lemma 2 can be deployed on the interval $[\eta_1, \eta_0]$.

Step 3. Relabel $\eta_0 = \eta_1$ and repeat Steps 1 and 2 until the entire support is covered. □

A4 Claim regressions

Table A3: Poisson estimation of claim risks

	Comprehensive		Collision		Liability	
	Beta	Std. Err.	Beta	Std. Err.	Beta	Std. Err.
Primary policyholder demographics						
Gender (1 = Male)	0.0354	(0.0099)	-0.0519	(0.0059)	-0.0530	(0.0059)
Age	-0.0161	(0.0033)	-0.0123	(0.0019)	-0.0046	(0.0018)
$(\text{Age})^2$	0.0137	(0.0035)	0.0168	(0.0020)	0.0084	(0.0019)
Married*	-0.0121	(0.0321)	-0.0210	(0.0185)	-0.0227	(0.0171)
Divorced	0.0317	(0.0441)	-0.0022	(0.0255)	-0.0013	(0.0231)
Separated	0.1038	(0.0733)	0.0331	(0.0441)	0.0075	(0.0383)
Widowed	-0.0017	(0.0711)	-0.0046	(0.0392)	-0.0056	(0.0372)
Rated driver demographics						
Gender (1 = Male)	0.0260	(0.0088)	-0.0037	(0.0053)	-0.0069	(0.0053)
Rated Driver Age	0.0067	(0.0031)	-0.0037	(0.0018)	-0.0046	(0.0017)
$(\text{Age})^2$	-0.0014	(0.0034)	0.0324	(0.0019)	0.0409	(0.0019)
Married*	-0.0059	(0.0320)	-0.0025	(0.0185)	-0.0025	(0.0171)
Divorced	-0.0002	(0.0457)	0.0178	(0.0265)	-0.0011	(0.0242)
Widowed	-0.0084	(0.0732)	-0.0032	(0.0401)	-0.0043	(0.0383)
Separated	-0.0039	(0.0764)	0.0173	(0.0461)	0.0260	(0.0402)
Vehicle information						
Number of drivers						
Many drivers (1[$\geqslant$ 3 drivers])	0.0246	(0.0093)	0.0422	(0.0056)	0.0575	(0.0056)
Number of young drivers	0.0330	(0.0053)	0.0600	(0.0031)	0.0499	(0.0030)
Vehicle age	-0.0029	(0.0026)	0.0016	(0.0017)	0.0047	(0.0013)
$(\text{Vehicle age})^2$	-0.0181	(0.0168)	-0.0243	(0.0119)	-0.0173	(0.0066)
Primary use†						
... for work	0.0961	(0.0081)	0.0087	(0.0048)	0.0283	(0.0048)
... for business	0.1868	(0.0341)	0.1485	(0.0203)	0.1541	(0.0203)
... for farm	-0.0015	(0.0989)	-0.0241	(0.0773)	-0.0290	(0.0758)
... for carpool	0.1670	(0.0445)	-0.0004	(0.0281)	0.0112	(0.0292)
Passive restraint system	-0.0007	(0.0471)	0.1790	(0.0363)	0.3100	(0.0244)
Anti-theft device	0.0612	(0.0128)	0.0498	(0.0076)	0.0735	(0.0072)
Anti-lock brakes	0.0039	(0.0148)	0.0016	(0.0087)	0.0470	(0.0083)
Motor vehicle records						
Accidents	0.0126	(0.0063)	0.0042	(0.0038)	-0.0007	(0.0038)
Suspensions	0.0173	(0.0125)	0.0236	(0.0071)	0.0200	(0.0066)
Revocations	-0.0087	(0.0569)	-0.0024	(0.0272)	0.0009	(0.0247)
Reinstatements	0.0206	(0.0125)	0.0147	(0.0074)	0.0194	(0.0068)
Misc. violations	0.0015	(0.0135)	-0.0004	(0.0083)	-0.0002	(0.0076)
Risk-related information						
Proprietary risk score	-0.0001	(0.0000)	-0.00017	(0.0000)	-0.0002	(0.0000)
County's annual mean unemployment	0.0053	(0.0072)	-0.0024	(0.0050)	-0.0025	(0.0049)
County's annual precip. (in.)	0.0410	(0.0118)	-0.0016	(0.0073)	-0.0014	(0.0067)
County's annual mean temp. (°F)	0.0014	(0.0028)	-0.0011	(0.0021)	-0.0013	(0.0021)
Estimated random effects						
$\alpha_j = Var(\exp(\epsilon_{ij}))$	0.92	(0.0242)	0.46	(0.0076)	1.47	(0.0112)
High-dimensional fixed effects						
Territory FE	Yes		Yes		Yes	
Year FE	Yes		Yes		Yes	
County FE	No		No		No	
Household-vehicle pairs	1,528,235		1,393,974		1,903,521	
Household-vehicle-years	5,130,679		4,639,503		6,058,169	
Log-likelihood	-398,747		-921,765		-1,141,065	

Notes: Random effects are at the household-vehicle level. Standard errors are not robust-clustered.
* Single is the omitted category for marital status.
† Pleasure is the omitted category for primary vehicle use.

A5 Liability loss distribution

The data record households' liability coverage limits and the loss amounts from submitted liability claims. A potential censoring concern is that losses that exceed the coverage limit may be renegotiated downward by the insurer; in the data we then observe only the renegotiated (and hence censored) amounts. The model below is designed to account for this feature of the claims process.

A5.1 Model of liability loss amounts

Households choose a combined single-limit liability coverage level $\int$ from the menu $\mathcal{S}$. Consider a decision maker with coverage $\int \in \mathcal{S}$. Conditional on a liability claim, the *true* (pre-negotiation) loss amount l is drawn from a distribution with density $h(l)$. The data record the *renegotiated* loss amount, denoted by ℓ, which is top-coded at the coverage limit. There are two cases:

- If $l < \int$, the insurer pays the full amount and no renegotiation occurs. Hence, $\ell = l$.
- If $l > \int$, the loss exceeds the coverage limit and the household is potentially liable for the amount above $\int$. In this case, renegotiation occurs with probability $\alpha \in (0, 1)$:
    1. If renegotiation occurs, the loss is negotiated down to an observed amount $\ell \in (\int - \underline{K}, \int)$, drawn independently with density $g(\ell)$.
    2. If renegotiation does not occur, the observed loss amount equals the coverage limit, $\ell = \int$.

A5.2 Parametrization and estimation

1. The distribution of pre-negotiation liability losses is parametrized as a mixture of two exponential distributions:

$$h(l; \theta^h) = m_\Gamma \Gamma_1 e^{-\Gamma_1 l} + (1 - m_\Gamma)\Gamma_2 e^{-\Gamma_2 l}, \qquad l \in (0, \infty), \tag{A2}$$

where $m_\Gamma \in [0,1]$, $\Gamma_1, \Gamma_2 \in (0,\infty)$, and $\theta^h = \{m_\Gamma, \Gamma_1, \Gamma_2\}$. Let $H(l;\theta^h)$ denote the corresponding CDF.

2. A loss drawn from $h(l;\theta^h)$ that exceeds the coverage limit $\int$ is renegotiated with probability α. When renegotiation occurs, the observed loss amount is redrawn independently from the interval $(\int - \underline{K}, \int)$. We set $\underline{K} = 20{,}000$. The redraw is distributed over this interval with density $g(\ell;\theta^g)$, which we parameterize as a beta distribution on $(0,1)$. Specifically, letting $u = (\ell - (\int - \underline{K}))/\underline{K}$,

$$g(u;\theta^g) = \frac{1}{B(a,b)}\, u^{a-1}(1-u)^{b-1}, \qquad u \in (0,1), \tag{A3}$$

where $B(a,b)$ is the beta function and $a, b \in (0,\infty)$.

For $\int \in \mathcal{S}$, the model-implied distribution of the observed loss amount is mixed (discrete at $\ell = \int$ and continuous on $(0, \int)$). The continuous density is

$$f_c(\ell \mid \int;\Theta) = \begin{cases} h(\ell;\theta^h), & \text{if } \ell < \int - \underline{K}, \\ h(\ell;\theta^h) + \alpha\big(1 - H(\int;\theta^h)\big)\dfrac{1}{\underline{K}}\, g\left(\dfrac{\ell - (\int - \underline{K})}{\underline{K}};\theta^g\right), & \text{if } \int - \underline{K} \leqslant \ell < \int. \end{cases} \tag{A4}$$

and the point mass at the limit is

$$\Pr(\ell = \int \mid \int;\Theta) = (1-\alpha)\big(1 - H(\int;\theta^h)\big).$$

In total, we estimate six parameters: $\theta^h = \{m_\Gamma, \Gamma_1, \Gamma_2\}$, $\theta^g = \{a, b\}$, and α. The maximum-likelihood estimates, reported in Table A4 result in

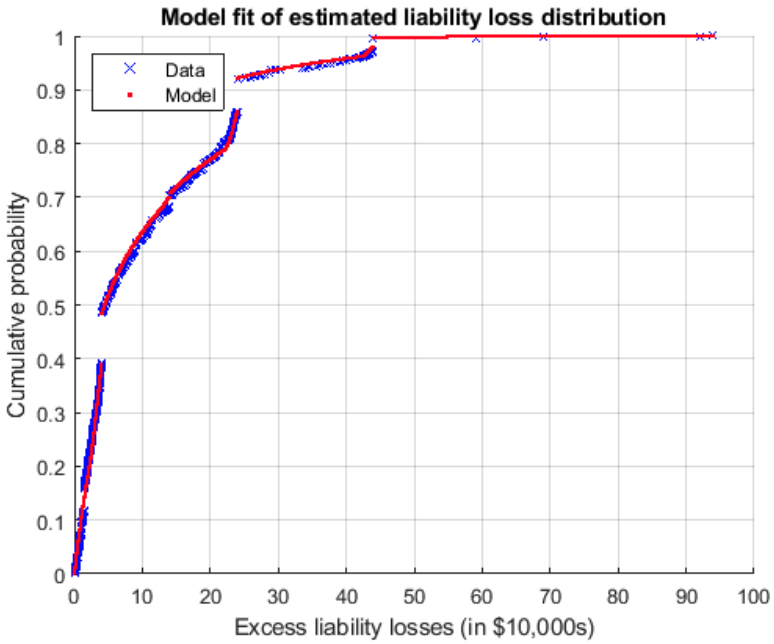
$$\Phi(l) \propto 0.3004 \cdot \left[0.2831 \cdot e^{-\frac{0.2831}{10000}\cdot l}\right] + 0.6996 \cdot \left[0.0389 \cdot e^{-\frac{0.0389}{10000}\cdot l}\right] \tag{A5}$$

Figure A12 compares the empirical claim distribution to the model-implied distribution.

Table A4: Estimated liability loss distribution

Parameter	Value
Γ_1	0.2831×10^{-4}
Γ_2	0.0389×10^{-4}
m_Γ	0.3004
α	0.4824
a	1.9226
b	1.1304

Figure A12: Model fit of the liability loss distribution



Notes: Plot shows liability loss amounts in excess of \$60,000 for individuals with coverages $\int > 0$.