The Nature of Risk Preferences: Evidence from Insurance Choices*

Levon Barseghyan†
Cornell University

Francesca Molinari
Cornell University

Ted O'Donoghue
Cornell University

Joshua C. Teitelbaum
Georgetown University

This version: June 30, 2011
Prior version: July 21, 2010

Abstract

We use data on insurance deductible choices to estimate a structural model of risky choice that permits "standard" risk aversion, loss aversion, and probability weighting. We show that loss aversion and probability weighting—though not separately identified without strong parametric assumptions—both imply a distortion of probabilities, and we demonstrate that such probability distortions are identified. We find that probability distortions—in the form of substantial overweighting of claim probabilities—play an important role in explaining the aversion to risk manifested in deductible choices. Once we allow for probability distortions, standard risk aversion is relatively small. (*JEL* D01, D03, D12, D81, G22)

*For helpful comments, we thank seminar and conference participants at Berkeley, UCLA, Georgetown University, the University of Pennsylvania, Princeton University, Michigan State University, the 2010 Behavioral Economics Annual Meeting, the Workshops on Behavioral and Institutional Research and Financial Regulation, FUR XIV, the 85th Annual Conference of the Western Economic Association International, the 2011 American Economic Association Annual Meeting, and the 21st Annual Meeting of the American Law and Economics Association. We also thank Darcy Steeg Morris for excellent research assistance. We acknowledge financial support from the National Science Foundation (SES-1031136). In addition, Barseghyan acknowledges financial support from the Institute for Social Sciences at Cornell University, and Molinari acknowledges financial support from National Science Foundation grants SES-0617482 and SES-0922330.

†Corresponding author: Levon Barseghyan, Department of Economics, Cornell University, 456 Uris Hall, Ithaca, NY 14853 (lb247@cornell.edu).

Electronic copy available at: http://ssrn.com/abstract=1646520

1 Introduction

Households are averse to risk—they require a premium to invest in equity and they purchase insurance at actuarially unfair rates. The standard expected utility model attributes risk aversion to a concave utility function defined over final wealth states (diminishing marginal utility for wealth). Indeed, many empirical studies of risk preferences assume expected utility and estimate such "standard" risk aversion (e.g., Cohen and Einav 2007).

A considerable body of research, however, suggests that the expected utility model may be inadequate. For instance, Rabin (2000) uses a calibration argument to demonstrate that reliance on the expected utility model to explain aversion to moderate-stakes risk implies an "absurd" degree of risk aversion over large-stakes risk. Sydnor (2010) applies a similar critique to argue that the level of standard risk aversion implied by most households' deductible choices in home insurance is implausibly large.

The leading alternatives to the expected utility model conjecture several features of risk preferences that may play a role in explaining aversion to risk. In terms of explaining aversion to moderate-stakes risk, the literature has focused on two features—loss aversion and probability weighting—both of which originate with prospect theory (Kahneman and Tversky 1979). For example, Kőszegi and Rabin (2007) and Sydnor (2010) argue that a form of "rational expectations" loss aversion proposed by Kőszegi and Rabin (2006) can explain the aversion to risk manifested in insurance deductible choices,[1] while Sydnor (2010) also mentions probability weighting and systematic risk misperceptions as possible explanations.[2]

In this paper, we investigate empirically the extent to which loss aversion and probability weighting can help explain aversion to moderate-stakes risk. We use data on households' deductible choices in auto and home insurance to estimate a structural model of risky choice that allows for loss aversion and probability weighting, as well as standard risk aversion. We explain that, without making strong assumptions regarding the form of the probability weighting function, one cannot separately identify loss aversion and probability weighting in our data. However, we show that loss aversion and probability weighting (alone and together) imply a distortion of probabilities relative to the expected utility model, and we demonstrate that such probability distortions are identified. Our estimates indicate that

[1] Under the original, "status quo" loss aversion proposed by Kahneman and Tversky (1979), gains and losses are defined relative to initial wealth. Under Kőszegi-Rabin loss aversion, gains and losses are defined relative to expectations about outcomes given choices. Unless we specify otherwise, we use "loss aversion" throughout the paper as shorthand for Kőszegi-Rabin loss aversion. On occasion, we say "KR loss aversion" to emphasize that we mean Kőszegi-Rabin loss aversion.

[2] As we discuss later, in our data literal probability weighting is indistinguishable from systematic risk misperceptions. Hence, we use "probability weighting" throughout the paper as shorthand for either literal probability weighting or systematic risk misperceptions.

Electronic copy available at: http://ssrn.com/abstract=1646520

probability distortions play a key role in explaining households' deductible choices. More specifically, we find large probability distortions in the form of substantial overweighting of claim probabilities. We also find that, once we allow for probability distortions, the mean estimate for standard risk aversion is relatively small—smaller than in Cohen and Einav (2007) or Sydnor (2010), but perhaps still larger than many economists typically assume.

In Section 2, we provide an overview of our data. The source of the data is a large U.S. property and casualty insurance company that offers multiple lines of insurance, including auto and home coverage. The full dataset comprises yearly information on more than 400,000 households who held auto or home policies between 1998 and 2006. For reasons we explain, we restrict attention in our main analysis to a core sample of 4170 households who hold both auto and home policies and who first purchased their auto and home policies from the company in the same year, in either 2005 or 2006. For each household, we observe the household's deductible choices in three lines of coverage—auto collision, auto comprehensive, and home all perils. We also observe the coverage-specific menus of premium-deductible combinations from which each household's choices were made. In addition, we observe each household's claims history for each coverage, as well as a rich set of demographic information. We utilize the data on claim realizations and demographics to assign each household a predicted claim probability for each coverage.

In Section 3, we develop our theoretical framework. We begin with an expected utility model of deductible choice, which incorporates standard risk aversion (Cohen and Einav 2007). To account for observationally equivalent households choosing different deductibles, and for individual households making "inconsistent" choices across coverages (Barseghyan et al. 2011; Einav et al. 2011), we follow McFadden (1974, 1981) and assume random utility with additively separable choice noise.

Next, we generalize the model to allow for loss aversion and probability weighting. We incorporate loss aversion by adopting a variant of the model of reference-dependent preferences proposed by Kőszegi and Rabin (2006, 2007). We show that loss aversion and probability weighting are not separately identified without strong assumptions regarding the form of the probability weighting function. Thus, although we could proceed to estimate the model by specifying one of the inverse-S-shaped probability weighting functions commonly used in the literature (e.g., Tversky and Kahneman 1992; Lattimore et al. 1992; Prelec 1998), we prefer not to rely on a parametric assumption to identify loss aversion. Instead, we show that loss aversion and probability weighting (alone and together) imply a distortion of probabilities relative to the expected utility model. Hence, we estimate a model that includes standard risk aversion and a probability distortion function (which is an amalgam of loss aversion and probability weighting). We conclude Section 3 by highlighting certain implications of our model, describing our estimation procedure, and discussing identification.

In Sections 4 and 5, we report the results of our main analysis. In Section 4, we assume homogenous preferences—i.e., we assume each household has the same coefficient of absolute risk aversion r and the same probability distortion function $\Omega(\mu)$—and take three nonparametric approaches to estimating $\Omega(\mu)$. Under each approach, we find large probability distortions in the form of substantial overweighting of claim probabilities on the relevant range (zero to twenty percent).[3] Under our primary approach, for example, the estimated probability distortion function implies $\Omega(0.02) = 0.08$, $\Omega(0.05) = 0.11$, and $\Omega(0.10) = 0.16$. We also find that, once we allow for probability distortions, the estimated degree of standard risk aversion declines substantially. At the end of Section 4, we highlight two key implications of our results for loss aversion and probability weighting. First, although we can say nothing about whether or not there is loss aversion, our results show that there indeed is probability weighting. Second, our results are suggestive of a probability weighting function that has the form originally posited by Kahneman and Tversky (1979), which is discontinuous at zero and trends toward a positive intercept.

In Section 5, we allow for heterogeneous preferences by permitting r and $\Omega(\mu)$ to depend on observables. The results are nearly identical. We also extend this model to account for household wealth or unobserved heterogeneity in risk, and little changes. Thus, whether we assume homogenous preferences or allow for heterogenous preferences, we find large probability distortions and concomitantly smaller standard risk aversion.

In Sections 6 and 7, we explore the robustness of our results. In Section 6, we investigate the sensitivity of our estimates to certain modeling assumptions, and we find that they are quite robust. In Section 7, we investigate whether unobserved heterogeneity in preferences (for which our model does not allow) might bias our results in favor of finding probability distortions. Specifically, we generate simulated deductible choices using models that include unobserved heterogeneity in preferences, and then estimate our empirical specification on such simulated data. The results lead us to conclude that unobserved heterogeneity in preferences cannot explain the large probability distortions we find in the actual data. We conclude in Section 8 by discussing certain implications and limitations of our study.

Numerous previous studies estimate risk preferences from observed choices, relying in most cases on survey and experimental data and in some cases on economic field data. Most of the studies in the literature—including two that use data on insurance deductible choices (Cohen and Einav 2007; Sydnor 2010)—estimate expected utility models, which permit only standard risk aversion. A handful of studies estimate models that allow for "status quo" loss aversion, probability weighting, or both—usually cumulative prospect theory models

[3] In the core sample, 99.4 percent of the predicted claim probabilities lie in the interval $(0, 0.20)$.

(Tversky and Kahneman 1992).[4] To our knowledge, no prior study estimates a model that permits standard risk aversion, KR loss aversion and probability weighting. Furthermore, the vast majority of the studies that estimate models with probability weighting take a parametric approach and assume one of the common inverse-S-shaped probability weighting functions. Hence, our paper contributes to the literature principally by estimating a model that allows for standard risk aversion, KR loss aversion, and probability weighting and that does not make strong parametric assumptions regarding the form of probability weighting.

Three recent studies echo our conclusion that probability weighting is important. Bruhin et al. (2010) use experimental data on subjects' choices over binary money lotteries to estimate a mixture model of cumulative prospect theory. They find that approximately 20 percent of subjects can essentially be characterized as expected value maximizers, while approximately 80 percent exhibit significant probability weighting. Snowberg and Wolfers (2010) use data on gamblers' bets on horse races to test the fit of two models—a model with standard risk aversion alone and a model with probability weighting alone—and find that the latter model better fits their data. Lastly, Kliger and Levy (2009) use data on call options on the S&P 500 index to estimate a cumulative prospect theory model, and they find that "status quo" loss aversion and probability weighting are manifested by their data. None of these studies, however, estimate models that combine standard risk aversion, KR loss aversion and probability weighting. Moreover, they all use typical inverse-S-shaped probability weighting functions. Finally, the latter two studies have only aggregate data, which necessitates that they take a representative agent approach and rely on equilibrium "ratio" conditions to identify the agent's utility function.

2 Data Description

2.1 Overview and Core Sample

We acquired the data from a large U.S. property and casualty insurance company. The company offers multiple lines of insurance, including auto, home, and umbrella policies. The full dataset comprises yearly information on more than 400,000 households who held auto or home policies between 1998 and 2006. For each household, the data contain all the information in the company's records regarding the household's characteristics (other than identifying information) and its policies (e.g., the limits on liability coverages, the limits and

[4] In addition to the studies discussed below, see, e.g., Tversky and Kahneman (1992), Cicchetti and Dubin (1994), Hey and Orme (1994), Jullien and Salanié (2000), Choi et al. (2007), Post et al. (2008), and Tanaka et al. (2010). Virtually all studies that estimate models with "status quo" loss aversion do not also permit standard risk aversion (over final wealth states), though some permit diminishing sensitivity.

deductibles on property damage coverages, and the premiums associated with each coverage). The data also record the number of claims that each household filed with the company under each of its policies during the period of observation.

In this paper, we restrict attention to households' deductible choices in three lines of coverage: (i) auto collision coverage; (ii) auto comprehensive coverage; and (iii) home all perils coverage.[5] In addition, we consider only the initial deductible choices of each household. This is meant to increase confidence that we are working with active choices; one might be concerned that some households renew their policies without actively reassessing their deductible choices. Finally, we restrict attention to households who hold both auto and home policies and who first purchased their auto and home policies from the company in the same year, in either 2005 or 2006. The latter restriction is meant to avoid temporal issues, such as changes in household characteristics and in the economic environment. In the end, we are left with a core sample of 4170 households.[6] Table 1 provides descriptive statistics for the variables we use later to estimate the households' utility parameters.

TABLE 1

2.2 Deductibles and Premiums

For each household in the core sample, we observe the household's deductible choices for auto collision, auto comprehensive, and home, as well as the premiums paid by the household for each type of coverage. In addition, the data contain the exact menus of premium-deductible combinations that were available to each household at the time it made its deductible choices. Table 2 summarizes the deductible choices of the households in the core sample. For each coverage, the most popular deductible choice is $500.

TABLE 2

Table 3 summarizes the premium menus. For each type of coverage, it summarizes the premium for coverage with a $500 deductible, as well as the marginal cost of decreasing the deductible from $500 to $250 and the marginal savings from increasing the deductible from $500 to $1000. The average annual premium for coverage with a $500 deductible is $180 for auto collision, $115 for auto comprehensive, and $679 for home. The average annual cost of decreasing the deductible from $500 to $250 is $54 for auto collision, $30 for auto

[5] A brief description of each type of coverage appears in the Appendix. For simplicity, we often refer to home all perils simply as home.

[6] As part of our sensitivity analysis in Section 6, we consider alternative samples with less restrictive inclusion criteria.

comprehensive, and \$56 for home. The average annual savings from increasing the deductible from \$500 to \$1000 is \$41 for auto collision, \$23 for auto comprehensive, and \$74 for home.[7]

TABLE 3

Because it is important to understand the sources of variation in premiums, we briefly describe the plan the company uses to rate a policy in each line of coverage. Upon observing the household's coverage-relevant characteristics, the company determines a base price $\bar{p}$ according to a coverage-specific rating function. Using the base price, the company then generates a household-specific menu $\{(p_d, d) : d \in \mathcal{D}\}$, which associates a premium p_d with each deductible d in the coverage-specific set of deductible options $\mathcal{D}$, according to a coverage-specific multiplication rule, $p_d = (g(d) \cdot \bar{p}) + c$, where $g(\cdot) > 0$ and $c > 0$. The multiplicative factors $\{g(d) : d \in \mathcal{D}\}$ are known in the industry as deductible factors, and c is known as an expense fee. The deductible factors and the expense fees are coverage specific but household invariant. Moreover, the expense fees are small markups that do not depend on the deductibles. The company's rating plan, including its rating function and multiplication rule, are subject to state regulation. Among other things, the regulations require that the company base its rating plan on actuarial considerations (losses and expenses) and prohibit the company and its agents from charging rates that depart from the company's rating plan.[8]

2.3 Claim Probabilities

For purposes of our analysis, we need to estimate for each household the likelihood of experiencing a claim for each coverage. We begin by estimating how claim rates depend on observables. In an effort to obtain the most precise estimates, we use the full dataset: 1,348,020 household-year records for auto and 1,265,229 household-year records for home. For each household-year record, the data record the number of claims filed by the household in that year. We assume that household i's claims under coverage j in year t follow a Poisson distribution with arrival rate λ_{ijt}. In addition, we assume that deductible choices do not influence claim rates, i.e., households do not suffer from moral hazard.[9] We treat the claim

[7] Tables A.1 through A.3 in the Appendix summarize the premium menus conditional on households' actual deductible choices.

[8] They also prohibit "excessive" rates and provide that insurers shall consider only "reasonable profits" in making rates. See, e.g., N.Y. Ins. Law §§ 2303, 2304 & 2314 (Consol. 2010), N.Y. Comp. Codes R. & Regis. tit. 11, § 160.2 (2010), and Dunham (2009, §§ 26.03 & 43.10).

[9] We follow Cohen and Einav (2007) and Barseghyan et al. (2011) in making this assumption. Note that it subsumes that there is neither ex ante moral hazard (deductible choice does not influence the frequency of claimable events) nor ex post moral hazard (deductible choice does not influence the decision to file a claim).

rates as latent random variables and assume that

$$\ln \lambda_{ijt} = \beta_j X_{ijt} + \epsilon_{ij},$$

where X_{ijt} is a vector of observables,[10] ϵ_{ij} is an unobserved iid error term, and $\exp(\epsilon_{ij})$ follows a gamma distribution with unit mean and variance ϕ_j. We perform standard Poisson panel regressions with random effects to obtain maximum likelihood estimates of β_j and ϕ_j for each coverage j. The results are reported in Tables A.4 and A.5 in the Appendix.

Next, we use the results of the claim rate regressions to generate predicted claim probabilities. Specifically, for each household i, we use the regression estimates to generate a predicted claim rate $\widehat{\lambda}_{ij}$ for each coverage j, conditional on the household's (ex ante) characteristics X_{ij} and (ex post) claims experience. In principle, during the policy period, a household may experience zero claims, one claim, two claims, and so forth. In the model, we assume that households disregard the possibility of more than one claim (see Section 3.1).[11] Given this assumption, we transform $\widehat{\lambda}_{ij}$ into a predicted claim probability $\widehat{\mu}_{ij}$ using[12]

$$\widehat{\mu}_{ij} = 1 - \exp(-\widehat{\lambda}_{ij}).$$

Table 4 summarizes the predicted claim probabilities for the core sample. The mean predicted claim probabilities for auto collision, auto comprehensive, and home are 0.069, 0.021, and 0.084, respectively, and there is substantial variation across households and coverages. Table 4 also reports pairwise correlations among the predicted claim probabilities and between the predicted claim probabilities and the premiums for coverage with a $500 deductible. Each of the pairwise correlations is positive, though none are large.

TABLE 4

2.4 Variation in Premiums and Claim Probabilities

Tables 3 and 4 reveal that, within each coverage, there is substantial variation in premiums and claim probabilities, and that premiums and claim probabilities are largely uncorrelated. A key identifying assumption is that there is variation in premiums and claim probabilities

[10] In addition to the variables in Table 1 (which we use later to estimate the households' utility parameters), X_{ijt} includes numerous other variables (see Tables A.4 and A.5 in the Appendix).

[11] Because claim rates are small (85 percent of the predicted claim rates in the core sample are less than 0.1, and 99 percent are less than 0.2), the likelihood of two or more claims is very small.

[12] The Poisson probability mass function is $f(x, \lambda) = \exp(-\lambda)\lambda^x/x!$ for $x = 0, 1, 2, ...$ and $\lambda \geq 0$. Thus, if the number of claims x follow a Poisson distribution with arrival rate λ, then the probability of experiencing at least one claim is $1 - \exp(-\lambda)$.

that is exogenous to the households' risk preferences. Hence, in our estimation procedure, we assume that a household's utility parameters depend on a vector of observables Z that is a strict subset of the variables that determine premiums and claim probabilities.[13] Many of the variables outside Z that determine premiums and claim probabilities, such as protection class and territory code,[14] are undoubtedly exogenous to the households' risk preferences. In addition, there are other variables outside Z that determine premiums but not claim probabilities, including numerous discount programs, which also are undoubtedly exogenous to the households' risk preferences.

Given our choice of Z, there is substantial variation in premiums and claim probabilities that is not explained by Z. In particular, regressions of premiums and predicted claim probabilities on Z yield low coefficients of determination. In the case of auto collision coverage, for example, regressions of premiums (for coverage with a \$500 deductible) on Z and predicted claim probabilities on Z yield coefficients of determination of 0.16 and 0.34, respectively.[15]

In addition to the substantial variation in premiums and claim probabilities within a coverage, there also is substantial variation in premiums and claim probabilities across coverages. A key feature of the data is that for each household we observe deductible choices for three coverages, and (even for a fixed Z) there is substantial variation in premiums and claim probabilities across the three coverages. Hence, even if the within-coverage variation in premiums and claim probabilities were insufficient in practice, we still might be able to estimate the model using across-coverage variation.

3 Theoretical Framework

3.1 Deductible Lotteries

We assume that a household treats its three deductible choices as independent decisions. This assumption is motivated in part by computational considerations,[16] but also by the literature on "narrow bracketing" (e.g., Read et al. 1999), which suggests that when people make multiple choices, they frequently do not assess the consequences in an integrated way, but rather tend to make each choice in isolation. Thus, we develop a model for how a house-

[13] In general, Z comprises the variables in Table 1.

[14] Protection class guages the effectiveness of local fire protection and building codes. Territory codes are based on actuarial risk factors, such as traffic and weather patterns, population demographics, wildlife density, and the cost of goods and services.

[15] They are even lower for auto comprehensive and home. In the case of auto comprehensive the coefficients of determination are 0.07 and 0.31, and in the case of home they are 0.04 and 0.15.

[16] If instead we were to assume that a household treats its deductible choices as a joint decision, then the household would face 120 options and the utility function would have several hundred terms.

hold chooses the deductible for a single type of insurance coverage. To simplify notation, we suppress the subscripts for household and coverage (though we remind the reader that premiums and claim probabilities are household and coverage specific).

The household faces a menu of premium-deductible pairs $\{(p_d, d) : d \in \mathcal{D}\}$, where p_d is the premium associated with deductible d and $\mathcal{D}$ is the coverage-specific set of deductible options. We assume that the household disregards the possibility of experiencing more than one claim during the policy period, and that the household believes the probability of experiencing one claim is μ. In addition, we assume that the household believes that its choice of deductible does not influence its claim probability, and that every claim exceeds the highest available deductible.[17] Under the foregoing assumptions, the choice of deductible involves a choice among lotteries of the form

$$L_d \equiv (-p_d, 1-\mu; -p_d - d, \mu),$$

to which we refer as *deductible lotteries.*

3.2 Standard Risk Aversion

To fix ideas, we initially assume that a household's preferences over deductible lotteries are influenced only by standard risk aversion—i.e., households are expected utility maximizers. The expected utility of deductible lottery L_d is

$$EU(L_d) = (1-\mu)u(w - p_d) + \mu u(w - p_d - d).$$

The function u represents standard utility defined over final wealth states, where w denotes initial wealth, and standard risk aversion is captured by the concavity of u.

To estimate the model, we first must specify utility u. We follow Cohen and Einav (2007) and Barseghyan et al. (2011) and consider a second-order Taylor expansion under the assumption that u has a negligible third derivative.[18] Also, because u is unique only up to an affine transformation, we normalize the scale of utility by dividing $u'(w)$. This yields

$$\frac{u(w+\Delta)}{u'(w)} - \frac{u(w)}{u'(w)} = \Delta - \frac{r}{2}\Delta^2,$$

[17] We make the latter assumption more plausible by excluding the \$2500 and \$5000 deductible options from the home menu. Only 1.6 percent of households in the core sample chose a home deductible of \$2500 or \$5000. We assign these households a home deductible of \$1000. In this respect, we follow Cohen and Einav (2007), who also exclude the two highest deductible options (chosen by 1 percent of the policyholders in their sample) and assign the third highest deductible to policyholders who chose the two highest options.

[18] While we use this utility specification in most of our analysis, we consider CRRA utility and CARA utility in Sections 5.2 and 6.1, respectively.

where $r \equiv -u''(w)/u'(w)$ is the coefficient of absolute risk aversion. Because the term $u(w)/u'(w)$ enters as an additive constant, it does not affect utility comparisons. With this specification, we have

$$U(L_d) \equiv \frac{EU(L_d)}{u'(w)} - \frac{u(w)}{u'(w)} = -\left[p_d + \mu d\right] - \frac{r}{2}\left[(1-\mu)\left(p_d\right)^2 + \mu\left(p_d + d\right)^2\right]. \tag{1}$$

The first term on the right-hand side of equation (1) reflects the expected value of deductible lottery L_d. The second term reflects the disutility from risk—it is the expected value of the squared losses, scaled by standard risk aversion.[19]

Next, we must account for observationally equivalent households choosing different deductibles, and for individual households making "inconsistent" choices across coverages (Barseghyan et al. 2011; Einav et al. 2011). We follow McFadden (1974, 1981) and assume random utility with additively separable choice noise. Specifically, we assume that the utility from deductible $d \in \mathcal{D}$ is given by

$$\mathcal{U}(d) \equiv U(L_d) + \varepsilon_d, \tag{2}$$

where ε_d is an iid random variable that represents error in evaluating utility. We initially assume that ε_d follows a type 1 extreme value distribution (also known as a Gumbel distribution) with scale parameter σ.[20] Hence, a household chooses deductible d when $\mathcal{U}(d) > \mathcal{U}(d')$ for all $d' \neq d$, and thus the probability that a household chooses deductible d is

$$\begin{aligned} \Pr(d) &= \Pr\left(\varepsilon_{d'} - \varepsilon_d < U(L_d) - U(L_{d'}) \text{ for all } d' \neq d\right) \\ &= \frac{\exp\left(U(L_d)/\sigma\right)}{\sum_{d' \in \mathcal{D}} \exp\left(U(L_{d'})/\sigma\right)}. \end{aligned} \tag{3}$$

In the estimation, we construct the likelihood function from these choice probabilities.

At this point, we could estimate equation (2) assuming that utility is specified by equation (1) and recover an estimate for r. When we do so, the estimated degree of risk aversion is quite large (see Section 4). This is consistent with Sydnor's (2010) main result, namely that, under the hypothesis of expected utility, homeowners' deductible choices imply absurdly large risk aversion. However, we are interested in going beyond standard risk aversion, and assessing whether we can enrich the model to better explain households' deductible choices.

[19] Note that this specification differs slightly from Cohen and Einav (2007) and Barseghyan et al. (2011), who use $U(L_d) = -\left[p_d + \lambda d\right] - \frac{r}{2}\left[\lambda d^2\right]$ (where λ is the Poisson arrival rate). The difference derives from the fact that those papers additionally take the limit as the policy period becomes arbitrarily small.

[20] The scale parameter σ is a monotone transformation of the variance of ε_d, and thus a larger σ means larger variance. Our estimation procedure permits σ to vary across coverages (see Section 3.4).

3.3 Probability Distortions

The behavioral economics literature conjectures several features of risk preferences that may play a role in explaining aversion to risk. In terms of explaining aversion to moderate-stakes risk, the literature focuses on two features—loss aversion and probability weighting—both of which originate with prospect theory (Kahneman and Tversky 1979). As we explain below, both loss aversion and probability weighting imply a distortion of probabilities relative to the expected utility model. We enrich our model by allowing for such probability distortions.

3.3.1 Loss Aversion

The original, "status quo" loss aversion proposed by Kahneman and Tversky (1979)—wherein gains and losses are defined relative to initial wealth—cannot explain aversion to risk in the context of insurance deductible choices, because all outcomes are losses relative to initial wealth. More recently, however, Kőszegi and Rabin (2006, 2007) and Sydnor (2010) have suggested that a form of "rational expectations" loss aversion—wherein gains and losses are defined relative to expectations about outcomes given choices—can explain the aversion to moderate-stakes risk manifested in insurance deductible choices.

In the Kőszegi-Rabin (KR) model, the expected utility from choosing a lottery depends both on standard "intrinsic" utility, which is defined over final wealth states, and on "gain-loss" utility, which results from experiencing outcomes that are better or worse than expected. Using the KR specification for gain-loss utility, and applying "choice-acclimating personal equilibrium" (which KR suggest is the appropriate equilibrium concept for insurance choices), equation (1) becomes

$$\begin{aligned} U(L_d) &= -\left[p_d + \mu d\right] - \frac{r}{2}\left[(1-\mu)\left(p_d\right)^2 + \mu\left(p_d + d\right)^2\right] \\ &\quad -\Lambda\left(1-\mu\right)\mu\left[d + \frac{r}{2}\left[\left(p_d + d\right)^2 - \left(p_d\right)^2\right]\right]. \end{aligned} \tag{4}$$

For an overview of the KR model and a derivation of equation (4), see the Appendix. The first two terms of equation (4) are equivalent to equation (1) and reflect the standard expected utility component of the KR model. The third term reflects the expected gain-loss utility component, where Λ captures the degree of loss aversion.[21] For $\Lambda = 0$, the household is loss neutral and the model reduces to expected utility. For $\Lambda > 0$, the household is loss averse.

[21] The KR model actually contains two parameters, one (λ) that captures the degree of loss aversion and one (η) that captures the importance of gain-loss utility relative to standard utility. Under choice-acclimating personal equilibrium, however, these parameters always appear as the product $\eta(\lambda - 1)$, which we label Λ.

We can rewrite equation (4) as

$$U(L_d) = -\left[p_d + \delta(\mu) d\right] - \frac{r}{2}\left[(1-\delta(\mu))\,(p_d)^2 + \delta(\mu)\,(p_d + d)^2\right], \tag{5}$$

where $\delta(\mu) \equiv \mu\left[1 + \Lambda\,(1-\mu)\right]$. Equation (5) is equivalent to equation (1), except that μ has been replaced by $\delta(\mu)$. Thus, KR loss aversion effectively implies a distortion of probabilities relative to the expected utility model. Moreover, because $\Lambda > 0$ implies $\delta(\mu) > \mu$, KR loss aversion increases the household's willingness to pay for lower deductibles.

3.3.2 Probability Weighting

In their original prospect theory paper, Kahneman and Tversky (1979) introduce probability weighting, whereby individual probabilities are transformed into decision weights. Incorporating probability weighting, and adopting the rank-dependent approach of Quiggin (1982), equation (1) becomes

$$U(L_d) = -\left[p_d + \pi(\mu) d\right] - \frac{r}{2}\left[(1-\pi(\mu))(p_d)^2 + \pi(\mu)(p_d + d)^2\right], \tag{6}$$

where $\pi(\mu)$ is the probability weighting function. From equation (6), we can see that probability weighting directly implies a distortion of probabilities relative to expected utility.

Over the years, several functional forms for $\pi(\mu)$ have been proposed (e.g., Tversky and Kahneman 1992; Lattimore et al. 1992; Prelec 1998), all of which share the same basic properties originally posited by Kahneman and Tversky (1979): overweighting of small probabilities, underweighting of large probabilities, and some insensitivity to probability changes in the intermediate range. It follows that in the domain of auto and home insurance, where claim probabilities generally are small, such probability weighting would increase the household's willingness to pay for lower deductibles.

3.3.3 Probability Distortions

As we have seen, both KR loss aversion and probability weighting imply a distortion of probabilities relative to the expected utility model. Incorporating both KR loss aversion and probability weighting into the model, equation (1) becomes

$$U(L_d) = -\left[p_d + \Omega(\mu) d\right] - \frac{r}{2}\left[(1-\Omega(\mu))\,(p_d)^2 + \Omega(\mu)\,(p_d + d)^2\right], \tag{7}$$

where

$$\Omega(\mu) \equiv \pi(\mu)\left[1 + \Lambda\,(1-\pi(\mu))\right]. \tag{8}$$

We refer to $\Omega(\mu)$ as the *probability distortion function.*

From equation (8), it is clear that if we impose a functional form for $\pi(\mu)$, we (potentially) can separately identify Λ and $\pi(\mu)$. If, however, we impose no functional form for $\pi(\mu)$—which is our preferred approach—then we cannot separately identify Λ and $\pi(\mu)$. In particular, even though (as we demonstrate below) we can identify the probability distortion function $\Omega(\mu)$, this function can comprise many different combinations of Λ and $\pi(\mu)$. In our analysis, therefore, we focus on estimating $\Omega(\mu)$, and then we discuss the implications for Λ and $\pi(\mu)$ (see Section 4.4).

Of course, even if we separate Λ from $\pi(\mu)$, we still face the issue of interpreting $\pi(\mu)$. As we discuss in Section 4.4, our estimates for $\Omega(\mu)$ suggest that $\pi(\mu) \neq \mu$ (assuming $\Lambda \geq 0$). Given our data, however, we cannot say whether $\pi(\mu) \neq \mu$ implies that households are engaging in probability weighting per se or that their subjective beliefs deviate systematically from the objective probabilities (or some combination of both). Hence, although we use the label probability weighting, we caution the reader that we use it as shorthand for either literal probability weighting or systematic risk misperceptions.

3.3.4 Model Implications

It is worth highlighting certain implications of equation (7) that play a role in identifying r versus $\Omega(\mu)$. Take any three deductible options $a, b, c \in \mathcal{D}$, with $a > b > c$. For a household with premium p_a for deductible a and claim probability μ, define $\widetilde{p}_b(p_a, \mu)$ as the premium for deductible b that makes the household indifferent between a and b, and define $\widetilde{p}_c(p_a, \mu)$ as the the premium for deductible c that makes the household indifferent between a and c. In other words, $\widetilde{p}_b(p_a, \mu) - p_a$ reflects the household's maximum willingness to pay (WTP) to reduce its deductible from a to b, and $\widetilde{p}_c(p_a, \mu) - p_a$ reflects the household's WTP to reduce its deductible from a to c. To simplify notation, we let $\widetilde{p}_b \equiv \widetilde{p}_b(p_a, \mu)$ and $\widetilde{p}_c \equiv \widetilde{p}_c(p_a, \mu)$. In the Appendix, we prove the following properties of $\widetilde{p}_b$ and $\widetilde{p}_c$ as functions of r and $\Omega(\mu)$.[22]

Property 1. *Both $\widetilde{p}_b$ and $\widetilde{p}_c$ are strictly increasing in r and $\Omega(\mu)$.*

Property 2. *If $r = 0$ then $\frac{\widetilde{p}_b - p_a}{\widetilde{p}_c - \widetilde{p}_b} = \frac{a-b}{b-c}$. If $r > 0$ then $\frac{\widetilde{p}_b - p_a}{\widetilde{p}_c - \widetilde{p}_b} > \frac{a-b}{b-c}$.*

Property 3. *Holding $\Omega(\mu)$ fixed, the ratio $\frac{\widetilde{p}_b - p_a}{\widetilde{p}_c - \widetilde{p}_b}$ is strictly increasing in r.*

Property 4. *If $\widetilde{p}_c$ is the same for $(r, \Omega(\mu))$ and a different $(r', \Omega(\mu)')$, then $\widetilde{p}_b$ is different for $(r, \Omega(\mu))$ and $(r', \Omega(\mu)')$. In particular, if $r > r'$ (in which case $\Omega(\mu) < \Omega(\mu)'$), then $\widetilde{p}_b$ is greater for $(r, \Omega(\mu))$.*

[22] We supress the explicit dependence of $\widetilde{p}_b$ and $\widetilde{p}_c$ on r and $\Omega(\mu)$ to simplify notation.

Property 1 is straightforward. A household's WTP to reduce its deductible (and thereby reduce its exposure to risk) will be larger if either it is more risk averse (in the standard sense) or its (distorted) claim probability is larger.

Property 2 is an implication of standard risk aversion. For a given $\Omega(\mu)$, a risk neutral household is willing to pay, for instance, exactly twice as much to reduce its deductible from \$1000 to \$500 as it is willing to pay to reduce its deductible from \$500 to \$250. In contrast, a risk averse household is willing to pay more than twice as much. This is because a risk averse household's WTP to avoid an incremental loss depends positively on the magnitude of the absolute loss. Property 3 states that the higher is the household's standard risk aversion, the stronger is this effect.

Property 4 is the key property for identification. Given Property 1, any particular $\widetilde{p}_c$ is consistent with multiple pairs of r and $\Omega(\mu)$. In particular, if we observe that $\widetilde{p}_c - p_a > \mu(a-c)$ (i.e., WTP is larger than what is actuarially fair), this could be due to $r > 0$, $\Omega(\mu) > \mu$, or both. However, using the same underlying intuition behind Property 3, one can show that, holding $\widetilde{p}_c$ constant, the ratio $(\widetilde{p}_b - p_a)/(\widetilde{p}_c - \widetilde{p}_b)$ is increasing in r. That is, the decline in $\Omega(\mu)$ required to keep $\widetilde{p}_c$ constant in response to an increase in r does not outweigh the direct effect of an increase in r on the ratio $(\widetilde{p}_b - p_a)/(\widetilde{p}_c - \widetilde{p}_b)$ from Property 3. Property 4 then immediately follows—if two $(r, \Omega(\mu))$ combinations both are consistent with a particular $\widetilde{p}_c$, the combination with more standard risk aversion implies a greater $\widetilde{p}_b$.

3.4 Estimation Procedure

We observe data $\{D_{ij}, \Gamma_{ij}\}$, where D_{ij} is household i's deductible choice for coverage j and $\Gamma_{ij} \equiv (Z_i, \widehat{\mu}_{ij}, P_{ij})$. In Γ_{ij}, Z_i is a vector of household characteristics, $\widehat{\mu}_{ij}$ is household i's predicted claim probability for coverage j, and P_{ij} denotes household i's menu of premium-deductible pairs for coverage j. In our analysis in Section 4, in which we assume homogenous preferences, Z_i comprises only a constant. When we allow for heterogeneous preferences in Section 5, Z_i comprises a constant and the variables in Table 1 (except, in most specifications, home value).

We estimate equation (2) assuming that utility is specified by equation (7). We generally assume that, for each coverage j, household i's predicted claim probability $\widehat{\mu}_{ij}$ corresponds to it subjective claim probability μ_{ij}.[23] The parameters to be estimated are:

[23] Of course, $\widehat{\mu}_{ij}$ may not correspond to μ_{ij} due to risk misperceptions, which could explain why we find probability distortions (see Section 3.3). It also is possible, however, that $\widehat{\mu}_{ij}$ may not correspond to μ_{ij} due to unobserved heterogeneity in risk. We address this issue in Section 5.3.

r_i	–	coefficient of absolute risk aversion ($r_i = 0$ means no standard risk aversion);
$\Omega_i(\mu)$	–	probability distortion function ($\Omega_i(\mu) = \mu$ means no probability distortions); and
σ_j	–	scale of choice noise for coverage j ($\sigma_j = 0$ means no choice noise).

We assume throughout that $\ln r_i = \beta_r Z_i$. We take several approaches to estimating the probability distortion function. For now, we write $\Omega_i(\mu|\beta_\Omega, Z_i)$, where β_Ω is a vector of parameters that will depend on the specific approach. Finally, we assume that σ_j does not vary across households but does vary across coverages. Hence, we estimate σ_L, σ_M, and σ_H for auto collision, auto comprehensive, and home, respectively.

We estimate the model via maximum likelihood using combined data for all three coverages. For each household i, the conditional loglikelihood function is

$$\ell_i(\theta) \equiv \sum_j \sum_{d \in \mathcal{D}_j} 1(D_{ij} = d) \ln [\Pr(D_{ij} = d | \Gamma_{ij}, \theta)],$$

where $\theta = (\beta_r, \beta_\Omega, \sigma_L, \sigma_M, \sigma_H)$, the indicator function selects the deductible chosen by household i for coverage j, and $\Pr(D_{ij} = d | \Gamma_{ij}, \theta)$ denotes the choice probability in equation (3). We estimate θ by maximizing $\sum_i \ell_i(\theta)$.

3.5 Identification

The random utility model in equation (2) comprises the sum of a utility function $U(L_d)$ and an error term ε_d. Using the results of Matzkin (1991), normalizations that fix scale and location, plus regularity conditions that are satisfied in our model, allow us to identify nonparametrically the utility function $U(L_d)$ within the class of monotone and concave utility functions. Identification of $U(L_d)$ allows us to identify r and $\Omega(\mu)$. To see this, note that identification of $U(L_d)$ allows us to identify—for any three deductible options $a, b, c \in \mathcal{D}$, with $a > b > c$, and any given premium p_a for deductible a—the pair of indifference premiums $\widetilde{p}_b(p_a, \mu)$ and $\widetilde{p}_c(p_a, \mu)$, where $\widetilde{p}_b(p_a, \mu)$ is the premium for deductible b that makes the household indifferent between a and b, and $\widetilde{p}_c(p_a, \mu)$ is the the premium for deductible c that makes the household indifferent between a and c. As we show in Section 3.3.4, different pairs of indifference premiums imply different combinations of r and $\Omega(\mu)$ (and vice versa). Thus, it is the variation in premiums for a fixed μ that allows us to pin down r and $\Omega(\mu)$. Variation in claim probabilities allows us to map out $\Omega(\mu)$ for all μ.

4 Analysis with Homogenous Preferences

We begin our analysis by assuming homogeneous preferences—i.e., r and $\Omega(\mu)$ are the same for all households. This permits us to take a nonparametric approach to estimating $\Omega(\mu)$ without facing a (prohibitive) curse of dimensionality. As a point of reference for our analysis, we note that if we do not allow for probability distortions (i.e., we restrict $\Omega(\mu) = \mu$), the estimate for r is 0.0129 (standard error: 0.0004).

4.1 Estimates

We take three nonparametric approaches to estimating $\Omega(\mu)$, none of which constrain $\Omega(\mu)$ to be continuous at $\mu = 0$. In Model 1a, we estimate a quadratic Chebyshev polynomial expansion of $\ln \Omega(\mu)$.[24] This approach naturally constrains $\Omega(\mu) > 0$. In Model 1b, we estimate a quadratic Chebyshev polynomial expansion of $\Omega(\mu)$ (and we restrict $\Omega(\mu) > 0$).[25] This approach naturally nests the case $\Omega(\mu) = \mu$. In Model 1c, we estimate $\Omega(\mu)$ using an 11-point cubic spline on the interval $(0, 0.20)$ (wherein lie 99.4 percent of the predicted claim probabilities in the core sample). Because it is a local approximation method, the cubic spline approach serves as a robustness check of both polynomial approaches.

Table 5 reports our results. The estimates for $\Omega(\mu)$ indicate large probability distortions. To illustrate, Figure 1 depicts the estimated $\Omega(\mu)$ for Models 1a, 1b, and 1c, along with the 95 percent pointwise confidence bands for Model 1c. In each model, there is substantial overweighting of claim probabilities. Moreover, all three models imply nearly identical distortions of claim probabilities between zero and 14 percent (wherein lie 96.7 percent of the predicted claim probabilities in the core sample), and even for claim probabilities greater than 14 percent the three models are statistically indistinguishable (Models 1a and 1b lie within the 95 percent confidence bands for Model 1c). Naturally, given this overweighting, the estimates for r are smaller than without probability distortions. Specifically, $\widehat{r}$ is 0.00064, 0.00063, and 0.00049 in Models 1a, 1b, and 1c, respectively.

TABLE 5 & FIGURE 1

[24] We considered expansions up to the 20th degree, and selected a quadratic on the basis of the Bayesian information criterion (BIC).

[25] As before, we considered expansions up to the 20th degree. Here, the BIC selected a cubic. However, because the BIC for the quadratic and cubic were essentially the same, we report results for the quadratic to faciliate direct comparisons with Model 1a.

4.2 Statistical Significance

To assess the relative statistical importance of probability distortions and standard risk aversion, we estimate restricted models and perform Vuong (1989) model selection tests.[26] We find that a model with probability distortions alone is "better" at the 1 percent level than a model with standard risk aversion alone. However, a likelihood ratio test rejects at the 1 percent level both (i) the null hypothesis of standard risk neutrality ($r = 0$) for Models 1a and 1b and (ii) the null hypothesis of no probability distortions ($\Omega(\mu) = \mu$) for Model 1b.[27] This suggests that probability distortions and standard risk aversion both play a statistically significant role.

4.3 Economic Significance

To give a sense of the economic significance of our estimates, we consider the implications for a household's maximum willingness to pay (WTP) for lower deductibles. Specifically, consider the household's WTP to reduce its deductible from \$1000 to \$500 when the premium for coverage with a \$1000 deductible is \$200. Note that this WTP corresponds to $\widetilde{p}_{\$500} - \$200$, where $\widetilde{p}_{\$500}$ is the premium for coverage with a \$500 deductible that makes the household indifferent between coverage with a \$1000 deductible and coverage with a \$500 deductible when the premium for coverage with a \$1000 deductible is \$200. Table 6 displays WTP for selected claim probabilities μ and various preference combinations, using the estimates for r and $\Omega(\mu)$ from Model 1a. It reveals that our estimated probability distortions and our estimated standard risk aversion both have an economically significant impact on a household's WTP for lower deductibles. More specifically, Table 6 illustrates two main points. First, in general, our estimated probability distortions have a large effect on WTP, while our estimated standard risk aversion has a relatively moderate effect on WTP. Second, the relative disparity in the effects of probability distortions and standard risk aversion diminishes at high claim probabilities, where both generate substantial aversion to risk.

TABLE 6

[26] Vuong's (1989) test allows one to select between two nonnested models on the basis of which best fits the data. Neither model is assumed to be correctly specified. Vuong (1989) shows that testing whether one model is significantly closer to the truth (its loglikelihood value is significantly greater) than another model amounts to testing the null hypothesis that the loglikelihoods have the same expected value.

[27] We do not perform a likelihood ratio test of the null hypothesis of no probability distortions for Model 1a because it does not nest the case $\Omega(\mu) = \mu$.

4.4 Loss Aversion and Probability Weighting

As we discuss in Section 3.3, the probability distortions we estimate could derive from KR loss aversion, probability weighting, or both. Moreover, without specifying a functional form for the probability weighting function, we cannot separately identify the degree of loss aversion Λ and the probability weighting function $\pi(\mu)$. Nevertheless, we can investigate which combinations of Λ and $\pi(\mu)$ are consistent with our estimated $\Omega(\mu)$.

Specifically, one can use equation (8) to derive an implied $\pi(\mu)$ as a function of Λ and the estimated $\Omega(\mu)$. Figure 2 performs this exercise using our estimated $\Omega(\mu)$ from Model 1a. As is clear from equation (8), if there is no loss aversion ($\Lambda = 0$), then the implied $\pi(\mu)$ is exactly equal to the estimated $\Omega(\mu)$. Furthermore, if households are loss averse ($\Lambda > 0$), the implied $\pi(\mu)$ is lower than the estimated $\Omega(\mu)$, and the larger is Λ the lower is the implied $\pi(\mu)$ (as illustrated in Figure 2). Notice the conclusion that emerges Figure 2. Our analysis can say nothing about whether or not there is any KR loss aversion. However, if households are either loss neutral or loss averse (i.e., $\Lambda \geq 0$)—as is typically assumed in the literature—our analysis implies that there indeed is probability weighting (i.e., $\pi(\mu) \neq \mu$).

There is another way to see that KR loss aversion alone cannot explain the $\Omega(\mu)$ that we estimate. From equation (8), we readily can see if households are loss averse but do not weight probabilities ($\pi(\mu) = \mu$), then (i) $\Omega(\mu) > \mu$ for all μ and (ii) $\Omega'(\mu) > 1$ for $\mu < \frac{1}{2}$. Neither, however, is consistent with our estimated $\Omega(\mu)$.

Finally, note that among the probability weighting functions that have been proposed in the literature, our estimated $\Omega(\mu)$ is suggestive of the form originally posited by Kahneman and Tversky (1979). In particular, it is consistent with a probability weighting function that is discontinuous at $\mu = 0$ and trends toward a positive intercept as μ approaches zero (although we have relatively little data for $\mu < 0.01$). In contrast, the functional forms suggested by Tversky and Kahneman (1992), Lattimore et al. (1992), and Prelec (1998), among others, require $\pi(\mu)$ to be continuous at $\mu = 0$ and $\pi(0) = 0$. Figure 2, however, suggests that such functional forms will not fit our data well, because they imply that $\pi(\mu)$ becomes very steep as μ approaches zero.[28]

FIGURE 2

[28] In Barseghyan et al. (2010), we estimate a similar model asuming one of these functional forms for $\pi(\mu)$ (in which case Λ and the relevant probability weighting parameter are separately identified), and we find $\Lambda \approx 0$. In light of Figure 2, this result is not surprising, because $\Lambda > 0$ would only exacerbate the excess steepness that these functions imply for small probabilities.

5 Analysis with Heterogenous Preferences

In this section, we allow for heterogenous preferences. Because Models 1a, 1b, and 1c yield nearly identical results, and because it naturally constrains $\Omega(\mu) > 0$, throughout this section we estimate a quadratic Chebyshev polynomial expansion of $\ln \Omega(\mu)$ (which is the best fit in Model 1a). Thus, we permit preferences to depend on observables Z_i, as follows:

$$\ln r_i = \beta_r Z_i \qquad \text{and} \qquad \ln \Omega_i(\mu) = \beta_{\Omega,1} Z_i + \left(\beta_{\Omega,2} Z_i\right) \mu + \left(\beta_{\Omega,3} Z_i\right) \mu^2.$$

As explained in Section 3.4, we estimate $\theta = (\beta_r, \beta_{\Omega,1}, \beta_{\Omega,2}, \beta_{\Omega,3}, \sigma_L, \sigma_M, \sigma_H)$ via maximum likelihood. We then use $\widehat{\theta}$ to assign fitted values of r_i and $\Omega_i(\mu)$ to each household i.

5.1 Benchmark Estimates

Table 7 reports the estimates of our benchmark specification with heterogenous preferences, which we label Model 2. In Model 2, Z_i includes a constant and the variables in Table 1, except for home value. We view home value primarily as a proxy for wealth, and thus we introduce it below when we endeavor to account for wealth. The top panel presents the coefficient estimates, $\widehat{\beta}_r$, $\widehat{\beta}_{\Omega,1}$, $\widehat{\beta}_{\Omega,2}$, and $\widehat{\beta}_{\Omega,3}$, and the bottom panel presents the estimates of the scale of choice noise, $\widehat{\sigma}_L$, $\widehat{\sigma}_M$, and $\widehat{\sigma}_H$. These estimates imply nontrivial heterogeneity in preferences and nonzero choice noise. The middle panel presents the mean and median fitted values for r and $\Omega(\mu)$. The estimates for $\Omega(\mu)$ are virtually identical to the estimates in Model 1a. Figure 3 depicts the mean estimated $\Omega(\mu)$, along with the 2.5th, 5th, 95th, and 97.5th percentiles. For comparison, it also depicts the mean estimated $\Omega(\mu)$ in Model 1a. Hence, whether we assume homogenous preferences or allow for heterogenous preferences, the main message is the same: large probability distortions in the form of substantial overweighting of claim probabilities. For r, the mean estimate is 0.00073, slightly higher than in Model 1a, while the median estimate is 0.00056, slightly lower than in Model 1a.[29]

TABLE 7 & FIGURE 3

5.2 Accounting for Wealth

Models 1a, 1b, 1c, and 2 all estimate a local approximation of standard risk aversion. Economists generally believe, however, that standard risk aversion depends on wealth. In this section, we endeavor to account for household wealth by using home value as a proxy. We pursue three approaches.

[29] The 2.5th, 5th, 95th, and 97.5th percentiles are 0.00028, 0.00028, 0.00177, and 0.00223, respectively.

In Model 3, we merely reestimate Model 2 when we add home value to the vector of observables Z_i upon which a household's utility parameters depend. However, economists typically do not assume that the parameters of utility functions depend on wealth, but rather that utility is a function of wealth (i.e., wealth is the domain of the utility function). Hence, Model 3 is perhaps a misspecified model.

In Model 4, we continue to specify utility by equation (7)—so we still are estimating a local approximation of standard risk aversion—but we assume that the underlying utility function exhibits constant relative risk aversion (CRRA), i.e., $u(w) = w^{1-\rho}/(1-\rho)$, where $\rho > 0$ is the coefficient of relative risk aversion. Under CRRA utility, $r = \rho/w$. For each household i, we assume $\ln \rho_i = \beta_\rho Z_i$ and take home value as a proxy for wealth, to wit $r_i = \rho_i/(\text{home value})_i$.[30] We also include home value in Z_i, because in addition to being a proxy for wealth, home value might also be a signal of household type.

In Model 5, we directly specify CRRA utility. That is, we replace equation (7) with

$$U(L_d) = \frac{EU(L_d)}{u'(w)} = (1-\Omega(\mu))\,\frac{(w-p_d)^{1-\rho}}{(1-\rho)\,w^{-\rho}} + \Omega(\mu)\frac{(w-p_d-d)^{1-\rho}}{(1-\rho)\,w^{-\rho}}.$$

This approach requires that we specify wealth for each household. We assume that (i) wealth is proportional to home value and (ii) average wealth is \$33,000 (2010 U.S. per capita disposable personal income), viz. $w = (33/191) \times (\text{home value})$.[31] As before, we assume $\ln \rho_i = \beta_\rho Z_i$ and include home value in Z_i. Once we estimate ρ_i for each household i, we recover an estimate for r_i using $r_i = \rho_i/w_i$.

Table 8 reports the mean and median of the utility parameter estimates for Models 3, 4, and 5.[32] For comparison, it also restates the benchmark estimates from Model 2. Models 3, 4, and 5 all yield estimates for $\Omega(\mu)$ that are very similar to the benchmark estimates. Hence, our main result—large probability distortions in the form of substantial overweighting of claim probabilities—seems robust to various methods of accounting for wealth.

The estimates for standard risk aversion also are noteworthy. In Model 3, the estimates for r are larger than the benchmark estimates, but again we believe this is a misspecified model. In Models 4 and 5, the estimates for r are somewhat smaller than the benchmark estimates—the mean and median estimates for r are 0.00050 and 0.00028 in Model 4 and 0.00056 and 0.00044 in Model 5. Finally, in Model 5, where we directly estimate relative risk

[30] Given an underlying assumption that $w_i = \beta_w \times (\text{home value})_i$, in this model we cannot separately identify β_w and the constant in β_ρ, as these parameters appear always via the ratio $r_i = \rho_i/w_i$ and $\rho_i = \exp(\beta_\rho Z_i)$. Hence, we can estimate ρ_i only up to a scale. On the other hand, by the same logic the scale of β_w (which we fix to $\beta_w = 1$) does not affect our estimates of r_i and $\Omega_i(\mu)$.

[31] Recall from Table 1 that the average home value in the core sample is approximately \$191,000.

[32] For the sake of brevity, Table 8 does not report the coefficient estimates or the estimates of the scale of choice noise. The complete results, however, are reported in Tables A.6 through A.8 in the Appendix.

aversion, the mean and median estimates for ρ are roughly 18 and 13. These estimates for ρ reflect that, in general, our estimates for standard risk aversion, though smaller than in Cohen and Einav (2007) or Sydnor (2010),[33] are perhaps still larger than many economists typically assume. We return to this issue in our concluding remarks in Section 8.

TABLE 8

5.3 Accounting for Unobserved Heterogeneity in Risk

In our analysis up to this point, we assign to each household in the core sample a predicted claim probability $\widehat{\mu}$ for each coverage. While this approach allows for heterogeneity in risk based on observable characteristics, it does not permit unobserved heterogeneity. Such unobserved heterogeneity, however, is potentially important (Cohen and Einav 2007). In order to account for unobserved heterogeneity in risk, we expand our approach and assign to each household a predicted *distribution* of claim probabilities for each coverage.

More specifically, in Section 3 we derive a household's choice probability as a function of its subjective claim probability μ. Up to this point, we assume that μ corresponds to the predicted claim probability $\widehat{\mu}$ derived from the claim rate regressions. Of course, the regressions yield not only the conditional expectation, but also the conditional distribution of claim rates. Hence, we can use the regression estimates to assign to each household, not just a predicted claim probability $\widehat{\mu}$, but also predicted claim probability distribution $\widehat{F}(\mu)$. We can then construct the likelihood function by integrating over $\widehat{F}(\mu)$.[34]

Table 9 reports the mean and median of the utility parameter estimates for Models 2, 4, and 5—relabeled as Models 2u, 4u, and 5u—when we allow for unobserved heterogeneity in risk.[35] As compared to Models 2, 4, and 5, respectively, Models 2u, 4u, and 5u indicate similar probability distortions, except at high claim probabilities where the overweighting is more pronounced. Models 2u, 4u, and 5u also indicate somewhat higher levels of standard risk aversion, with mean and median estimates for r ranging from 0.00081 to 0.00097 and from 0.00055 to 0.00082, respectively.[36]

[33] To be precise, our estimates for r are smaller than the mean estimates in Cohen and Einav (2007) and the median estimates in Sydnor (2010) (which does not report mean estimates).

[34] Just as the predicted claim rate underlying $\widehat{\mu}$ is conditional on the household's (ex ante) characteristics and (ex post) claims experience, the predicted claim rate distribution underlying $\widehat{F}(\mu)$ also is conditional on the household's (ex ante) characteristics and (ex post) claims experience. When integrating over $\widehat{F}(\mu)$ to construct the likelihood function, we compute the integral using the Gauss-Laguerre quadrature method.

[35] The complete results, with the coefficient estimates and the estimates of the scale of choice noise, are reported in Tables A.9, A.10, and A.11 in the Appendix.

[36] When we estimate Models 1a and 1b allowing for unobserved heterogeneity in risk, the estimates for r and $\Omega(\mu)$ are similar to the estimates reported in Section 4.

TABLE 9

6 Sensitivity Analysis

The results of our main analysis yield a clear main message: large probability distortions in the form of substantial overweighting of claim probabilities. In this section, we investigate the sensitivity of this message, and we find that it is quite robust to a variety of alternative modeling assumptions. In addition, we find that standard risk aversion remains relatively small, though the estimates for r vary somewhat across specifications. To conserve space, we only summarize the results of the sensitivity analysis below. The complete results are available in the Appendix (Tables A.12 through A.20).

6.1 CARA Utility

In our main analysis, we consider a second-order Taylor expansion of the utility function, and also CRRA utility. Here we take yet another approach: we assume constant absolute risk aversion (CARA) utility, $u(w) = -\exp(-rw)$. That is, we specify utility as

$$U(L_d) = \frac{EU(L_d)}{u'(w)} = (1 - \Omega(\mu)) \frac{-\exp(rp_d)}{r} + \Omega(\mu) \frac{-\exp(r(p_d + d))}{r},$$

which we note is independent of wealth w. When we estimate Model 2 with CARA utility, the main message is the same. The estimates for $\Omega(\mu)$ indicate similar probability distortions, albeit somewhat less pronounced than the benchmark. Consequently, the mean and median estimates for r are higher than the benchmark, at 0.00113 and 0.00103, respectively.

6.2 Alternative Samples

In the core sample, we restrict attention to households who hold both auto and home policies and who first purchased their auto and home policies from the company in the same year, in either 2005 or 2006. Here we estimate Model 2 using two less restrictive samples: (1) households who hold auto policies and who first purchased their auto policies from the company in the same year, in either 2005 or 2006; and (2) households who hold both auto and home policies and who first purchased their auto and home policies from the company in the same year, in either 2004, 2005, or 2006. Again, the main message is the same. In both samples, the estimates for $\Omega(\mu)$ indicate probability distortions that are very similar to the benchmark. As for standard risk aversion, in sample 1 the mean and median estimates for r are higher than the benchmark estimates, at 0.00113 and 0.00112, while in sample 2 they are lower than the benchmark estimates, at 0.00060 and 0.00048.

6.3 Restricted Menus

In our main analysis, we use the full menu of deductible options for each coverage, up to \$1000. This raises two potential concerns. The first concern arises from the fact that, in each coverage, the vast majority of households choose one of three deductibles: 92.3 percent of households choose a deductible of \$200, \$250, or \$500 in auto collision; 87.1 percent of households choose a deductible of \$200, \$250, or \$500 in auto comprehensive; and 97.5 percent of households choose a deductible of \$250, \$500, or \$1000 in home. Given these choice patterns, one might worry that households do not really consider the other deductible options, which could bias our estimates.[37] The second concern arises from our assumption that every claim exceeds the highest available deductible (\$1000). One might worry that this assumption is too strong, which would imply that the model (specifically, the form of the deductible lottery) is misspecified. To address these concerns, we estimate Model 2 for two cases of restricted menus: (I) we restrict the menu of deductible options to $\{\$200, \$250, \$500\}$ for each auto coverage and to $\{\$250, \$500, \$1000\}$ for home coverage; and (II) we restrict the menu of deductible options to $\{\$200, \$250, \$500\}$ for each auto coverage and to $\{\$100, \$250, \$500\}$ for home coverage.[38] In each case, the estimates for $\Omega(\mu)$ indicate probability distortions that are similar to the benchmark. Indeed, in case I the overweighting is more pronounced at high claim probabilities, while in case II it is more pronounced at low and high claim probabilities. The mean and median estimates for r are lower than the benchmark estimates, at 0.00029 and 0.00013 in case I and 0.00066 and 0.00048 in case II.

6.4 Alternative Error Structures

In our main analysis, we assume that the utility from every deductible $d \in \mathcal{D}$ is given by $\mathcal{U}(d) = U(L_d) + \varepsilon_d$, where ε_d is an iid Gumbel random variable. Here, we estimate Model 2 under two alternative assumptions (and, for computational and theoretical reasons, using

[37] For instance, when a household chooses a \$250 deductible in home, we are using the fact that it did not choose a \$100 deductible to infer an upper bound on its aversion to risk. But if the household in fact does not even consider the \$100 deductible as an option, our inference would be invalid. Similarly, when a household chooses a \$500 deductible in auto comprehensive, we are using the fact that it did not choose a \$1000 deductible to infer a lower bound on its aversion to risk. Again, if the household in fact does not even consider the \$1000 deductible as an option, our inference would be invalid.

[38] In each case, if a household's actual deductible choice is outside the restricted menu, we assign to the household the deductible option from the restricted menu that is closest to their actual deductible choice. In this respect, we follow Cohen and Einav (2007).

the restricted menus from Section 6.3): (A) we assume (as before) that the utility from every deductible $d \in \mathcal{D}$ is given by $\mathcal{U}(d) = U(L_d) + \varepsilon_d$, but we assume that ε_d is an iid normal random variable; and (B) we assume that the utility from the maximum deductible, D, is given by $\mathcal{U}(D) = U(L_D) + \varepsilon_D$, where ε_D is an iid normal random variable, but that the utility from the other deductibles are given by $\mathcal{U}(d) = U(L_d) + \zeta_d$, where $\zeta_d = -\varepsilon_D$ for the minimum deductible and $\zeta_d = 0$ for the intermediate deductible. Alternative A provides a check of the Gumbel error assumption. Alternative B adds a check of the iid assumption. More specifically, we consider alternative B to address concerns arising from the fact that in principle the iid assumption allows for nonmonotonic ranking of deductibles. Once again, the main message is the same. Under both alternatives, the estimates for $\Omega(\mu)$ indicate similar probability distortions, though generally somewhat more pronounced. In addition, under alternative A the mean and median estimates for r are lower than the benchmark estimates—0.00022 and 0.00008 under restricted menu I and 0.00035 and 0.00015 under restricted menu II—while under alternative B they are higher than the benchmark estimates—0.00101 and 0.00081 under restricted menu I and 0.00093 and 0.00082 under restricted menu II.

7 Robustness to Unobserved Heterogeneity in Preferences

In our main analysis, we do not allow for unobserved heterogeneity in preferences. In this section, we investigate whether this modeling choice might bias our results in favor of our main finding—namely, large probability distortions. Specifically, we perform robustness exercises in which we generate simulated deductible choices using alternative models and then estimate our model on the "simulated data"—i.e., the actual data but with the simulated deductible choices substituted for the actual deductible choices. We pursue this approach instead of a sensitivity analysis in which we directly allow for unobserved heterogeneity in preferences because the latter would require estimating two distributions of unobserved heterogeneity, each of which might be correlated with unobserved heterogeneity in risk, and doing so would impose an undue computational burden and put a strain on identification.[39]

We emphasize that at the simulation stage of each exercise, we allow for both observed and unobserved heterogeneity in preferences and risk. However, because of severe computational burden, at the estimation stage we use Models 1a and 1b, which assume homogenous preferences and do not allow for unobserved heterogeneity in risk (although in one case in which we find some small bias in favor of probability distortions, we also estimate Model 1b

[39] Moreover, these problems would be amplified if we were to permit a flexible correlation structure among the various sources of unobserved heterogeneity.

allowing for unobserved heterogeneity in risk).[40] We report the full details of each exercise in the Appendix. Here, we summarize our methods and conclusions.

7.1 Unobserved Heterogeneity in Standard Risk Aversion

We first consider whether unobserved heterogeneity in standard risk aversion could lead us to find large probability distortions when in fact none exist. Cohen and Einav (2007) estimate an expected utility model that permits unobserved heterogeneity in standard risk aversion, and they conclude that it is substantial. They further conclude that it is strongly positively correlated with unobserved heterogeneity in risk. Using a variant of the Cohen-Einav model, we generate simulated deductible choices for each household i and coverage j, as follows. First, we generate the subjective claim probability $\mu_{ij} = 1 - \exp(-\lambda_{ij})$ using the claim rate $\lambda_{ij} = \widehat{\lambda}_{ij} \exp(\varepsilon_{ij}^{\lambda})$, where $\widehat{\lambda}_{ij}$ is household i's predicted claim rate for coverage j and $\varepsilon_{ij}^{\lambda}$ is a draw from a gamma distribution with unit mean and variance $\widehat{\phi}_{ij}$ (the estimated variance from the claim rate regression for coverage j updated for household i's (ex post) claims experience). Second, we generate standard risk aversion $r_{ij} = \exp(\widehat{\beta}_r Z_i + \varepsilon_{ij}^{r})$, where $\widehat{\beta}_r$ is the vector of coefficient estimates assuming only standard risk aversion and ε_{ij}^{r} is a draw from a standard normal distribution.[41] We consider several correlations between $\varepsilon_{ij}^{\lambda}$ and ε_{ij}^{r}. Finally, we generate the simulated deductible choice $\widetilde{D}_{ij}$ by applying the expected utility model (equation (1)) when the household faces menu P_{ij}.[42] In the end, the simulated data comprise $\{\widetilde{D}_{ij}, \widehat{\mu}_{ij}, P_{ij}\}$, where $\widetilde{D}_{ij}$ is household i's simulated deductible choice for coverage j and $\widehat{\mu}_{ij}$ and P_{ij} come from the actual data.[43]

Using this simulated data, we estimate Models 1a and 1b—i.e., homogenous preferences with a quadratic Chebyshev polynomial expansion of $\ln \Omega(\mu)$ (Model 1a) or $\Omega(\mu)$ (Model 1b). Figure 4 depicts, for each model, the mean estimated $\Omega(\mu)$ using the simulated data (along with the 2.5th and 97.5th percentiles) when there is a strong positive correlation between $\varepsilon_{ij}^{\lambda}$ and ε_{ij}^{r}, which is the case where we find the greatest bias.[44] For comparison, it

[40] In each exercise, we generate 100 independent simulated datasets of 4170 households making the three deductible choices. We then estimate our model on each dataset, and report the mean estimated $\Omega(\mu)$ along with the 2.5th and 97.5th percentiles.

[41] Increasing the variance of ε_{ij}^{r} increases the potential for bias, but at the same time it increases the frequency with which the alternative model predicts extreme deductible choices (because it increases the number of households with a very low or very high r_i). We limit the variance of ε_{ij}^{r} to keep this frequency from becoming excessive relative to the actual data, and with a variance of one the simulated data generating process already implies a substantially higher frequency than we observe in the data.

[42] Note that our procedure permits household i's risk aversion r_{ij} to be coverage specific. Thus, it is as if each household-coverage observation is a distinct household. If instead we assume that household i's risk aversion r_i is the same across coverages, the results are nearly identical.

[43] Recall that $\widehat{\mu}_{ij}$ is household i's predicted claim probability for coverage j and P_{ij} denotes household i's menu of premium-deductible pairs for coverage j.

[44] In the Appendix, we report in detail the results of many simulations in which we vary the correlation

also depicts the estimated $\Omega(\mu)$ using the actual data. We find some bias in favor of finding probability distortions ($\Omega(\mu) > \mu$); however, the distortions we find using the simulated data are rather small and significantly less than the large distortions we find using the actual data. In order to investigate the source of the bias, we also estimate Model 1b when we allow for unobserved heterogeneity in risk (Model 1bu). As Figure 4 illustrates, when we estimate Model 1bu on the simulated data, we now find no bias in favor of finding probability distortions. By contrast, when we estimate Model 1bu on the actual data, the probability distortions do not go away—indeed, they are very similar to the probability distortions we find when we estimate Model 1b on the actual data. Moreover, recall that in our analysis in Section 5 (where we allow for observed heterogeneity in preferences) the estimated probability distortions also persist when we allow for unobserved heterogeneity in risk (Models 2u, 4u, and 5u). These results lead us to conclude that unobserved heterogeneity in standard risk aversion cannot explain the large probability distortions we find in the data.

FIGURE 4

7.2 Unobserved Heterogeneity in Standard Risk Aversion and Probability Distortions

We next consider whether unobserved heterogeneity in both r and $\Omega(\mu)$ could bias our results in favor of finding probability distortions. To investigate this possibility, we generate simulated deductible choices for each household i and coverage j, as follows. First, we generate the subjective claim probability μ_{ij} exactly as above. Second, we generate standard risk aversion $r_{ij} = \exp(\widehat{\beta}_r Z_i + \varepsilon_{ij}^r)$ as above, except that now $\widehat{\beta}_r$ is the vector of coefficient estimates from Model 2. Third, we generate probability distortions $\Omega_{ij}(\mu) = \exp(\widehat{\beta}_{\Omega,1} Z_i + \widehat{\beta}_{\Omega,2} Z_i \mu + \widehat{\beta}_{\Omega,3} Z_i \mu^2 + \varepsilon_{ij}^{\Omega})$, where $\widehat{\beta}_{\Omega,1}$, $\widehat{\beta}_{\Omega,2}$, and $\widehat{\beta}_{\Omega,3}$ are the vectors of coefficient estimates from Model 2 and $\varepsilon_{ij}^{\Omega}$ is a draw from a standard normal distribution. We consider several correlation structures among $\varepsilon_{ij}^{\lambda}$, ε_{ij}^{r}, and $\varepsilon_{ij}^{\Omega}$. Finally, we generate the simulated deductible choice $\widetilde{D}_{ij}$ by applying our model (equation (7)) when the household faces menu P_{ij}. In the end, the simulated data comprise $\{\widetilde{D}_{ij}, \widehat{\mu}_{ij}, P_{ij}\}$, where $\widetilde{D}_{ij}$ is household i's simulated deductible choice for coverage j and $\widehat{\mu}_{ij}$ and P_{ij} come from the actual data.

Using this simulated data, we estimate Model 1a—homogenous preferences with a quadratic Chebyshev polynomial expansion of $\ln \Omega(\mu)$.[45] Figure 5 depicts (i) the mean $\Omega(\mu)$ used to generate the simulated data (along with the 2.5th and 97.5th percentiles) and (ii) the mean

between $\varepsilon_{ij}^{\lambda}$ and ε_{ij}^{r} as well as the variance of ε_{ij}^{r}.

[45] Here, we estimate only Model 1a because the process by which we generate the simulated data assumes the log form for $\Omega(\mu)$.

estimated $\Omega(\mu)$ using the simulated data (along with the 2.5th and 97.5th percentiles) for the cases where $\varepsilon_{ij}^{\lambda}$, ε_{ij}^{r}, and $\varepsilon_{ij}^{\Omega}$ are uncorrelated (left panel) and where $\varepsilon_{ij}^{\lambda}$, ε_{ij}^{r}, and $\varepsilon_{ij}^{\Omega}$ are strongly positively correlated (right panel). In each case, disregarding unobserved heterogeneity in r and $\Omega(\mu)$ biases *against* finding probability distortions. Moreover, when ε_{ij}^{r} and $\varepsilon_{ij}^{\Omega}$ are strongly negatively correlated (not depicted), there appears to be no bias. Intuitively, with strong negative correlation, the two sources of unobserved heterogeneity roughly offset each other in terms of their effects on choices. If anything, therefore, this exercise indicates that disregarding unobserved heterogeneity in both standard risk aversion and probability distortions might lead us to *underestimate* the magnitude of probability distortions.

FIGURE 5

7.3 Correlated Unobserved Heterogeneity in Risk

The exercises reported in Sections 7.1 and 7.2 assume that the unobserved heterogeneity in risk is not correlated across coverages (i.e., $\varepsilon_{iH}^{\lambda}$, $\varepsilon_{iL}^{\lambda}$, and $\varepsilon_{iM}^{\lambda}$ are uncorrelated). In fact, the results of these exercises are robust to various correlation structures. For instance, whether we assume $\varepsilon_{iH}^{\lambda}$, $\varepsilon_{iL}^{\lambda}$, and $\varepsilon_{iM}^{\lambda}$ are perfectly correlated or have pairwise correlations of -0.33 (the strongest possible given three pairs), our conclusions in Sections 7.1 and 7.2 do not change in any noticeable way. It is worth noting that, in some instances, allowing for correlation among $\varepsilon_{iH}^{\lambda}$, $\varepsilon_{iL}^{\lambda}$, and $\varepsilon_{iM}^{\lambda}$ causes unobserved heterogeneity in preferences to be correlated across coverages. Finally, we note that if we assume the same correlation structures for unobserved heterogeneity in risk, but without permitting unobserved heterogeneity in preferences, we again find essentially no bias in favor of finding probability distortions.

8 Discussion

We develop a structural model of risky choice that permits standard risk aversion and probability distortions, where the latter can arise from loss aversion or probability weighting. We estimate the model using data on households' deductible choices in auto and home insurance. We find that large probability distortions—in the form of substantial overweighting of small probabilities—play a statistically and economically significant role in explaining households' deductible choices. Given our data, we cannot say whether or to what extent loss aversion accounts for these probability distortions. However, our analysis provides clear evidence of probability weighting.

Perhaps the main takeaway of the paper is that economists should pay greater attention to the question of how people evaluate risk. Prospect theory incorporates two key features: a value function that describes how people evaluate outcomes and a probability weighting function that describes how people evaluate risk. The literature, however, has focused primarily on the value function, and there has been relatively little focus on probability weighting.[46] In light of our work, as well as the work discussed in the introduction that reaches similar conclusions using different methods (Bruhin et al. 2010; Snowberg and Wolfers 2010; Kliger and Levy 2009), it seems clear that future research on decision making under uncertainty should focus more attention on probability weighting.[47]

That said, it is worth highlighting certain limitations of our analysis. An important limitation is that, while our analysis clearly indicates that a lot "action" lies in how people evaluate risk, it does not enable us to say whether households are engaging in probability weighting per se—i.e., they know the probabilities but weight them nonlinearly—or whether their subjective beliefs simply do not correspond to the objective probabilities. Relatedly, although probability weighting is a natural candidate for explaining the probability distortions that we find, other mechanisms, such as ambiguity aversion, also could give rise to similar probability distortions. An important avenue of future research, therefore, is to investigate different accounts of how people evaluate risk and uncertainty.

Another limitation is that our analysis relies exclusively on insurance deductible choices, and hence we urge caution when generalizing our conclusions to other choices or settings. In particular, the vast majority of the claim probabilities we observe lie between zero and twenty percent, and thus our analysis implies little about what probability distortions might look like outside that range. While we suspect that our main message would resonate in many domains beyond insurance deductible choices that involve similar probabilities, we hesitate to make any conjectures about contexts where larger probabilities are involved.

Finally, it is worth discussing the magnitude of our estimates of standard risk aversion. As we note in the introduction, an important critique of the standard expected utility model, offered by Rabin (2000), is that reliance on the expected utility model to explain aversion to moderate-stakes risk implies an absurd degree of risk aversion over large-stakes risk. When we estimate an expected utility model—which does not permit probability distortions—our estimate of absolute risk aversion is 0.0129. Assuming wealth of $33,000 (2010 U.S. per capita disposable personal income), this implies relative risk aversion in the triple digits.[48] In light of

[46] Two prominent review papers—an early paper that helped set the agenda for behavioral economics (Rabin 1998) and a recent paper that surveys the current state of empirical behavioral economics (DellaVigna 2009)—contain almost no discussion of probability weighting. The behavioral finance literature has paid more attention to probability weighting (see, e.g., Barberis and Huang 2008; Barberis 2010)

[47] Indeed, Prelec (2000) conjectured that "probability nonlinearity will eventually be recognized as a more important determinant of risk attitudes than money nonlinearity."

[48] Using data similar to ours and assuming expected utility, Sydnor (2010) also estimates triple-digit coefficients of relative risk aversion.

the Rabin critique, triple-digit relative risk aversion is implausibly large. Indeed, economists frequently argue that a reasonable value is in the low single digits. By comparison, when we estimate our model—which permits probability distortions—there is far less standard risk aversion. Under our benchmark model, our mean estimate of absolute risk aversion is 0.0007, which, assuming wealth of $33,000, implies relative risk aversion of 23. Moreover, our mean estimates of absolute risk aversion in other specifications imply coefficients of relative risk aversion that range from 7 to 42, and when we directly specify CRRA utility our mean estimate of relative risk aversion is 18. Clearly, these estimates come much closer to what economists generally view as reasonable. However, they are perhaps still "too large." Hence, more work is necessary to fully understand the aversion to moderate-stakes risk manifested in insurance deductible choices.

References

BARBERIS, N. (2010): "A Model of Casino Gambling," Mimeo, Yale University, http://badger.som.yale.edu/faculty/ncb25/gb23d.pdf.

BARBERIS, N. AND M. HUANG (2008): "Stocks as Lotteries: The Implications of Probability Weighting for Security Prices," *American Economic Review*, 98, 2066–2100.

BARSEGHYAN, L., F. MOLINARI, T. O'DONOGHUE, AND J. C. TEITELBAUM (2010): "The Nature of Risk Preferences: Evidence from Insurance Choices," Mimeo, Cornell University, http://www.arts.cornell.edu/econ/lbarseghyan/bmotjuly2010.pdf.

BARSEGHYAN, L., J. PRINCE, AND J. C. TEITELBAUM (2011): "Are Risk Preferences Stable Across Contexts? Evidence from Insurance Data," *American Economic Review*, 101, 591–631.

BRUHIN, A., H. FEHR-DUDA, AND T. EPPER (2010): "Risk and Rationality: Uncovering Heterogeneity in Probability Distortion," *Econometrica*, 78, 1375–1412.

CHOI, S., R. FISMAN, D. GALE, AND S. KARIV (2007): "Consistency and Heterogeneity of Individual Behavior under Uncertainty," *American Economic Review*, 97, 1921–1938.

CICCHETTI, C. J. AND J. A. DUBIN (1994): "A Microeconometric Analysis of Risk Aversion and the Decision to Self-Insure," *Journal of Political Economy*, 102, 169–186.

COHEN, A. AND L. EINAV (2007): "Estimating Risk Preferences from Deductible Choice," *American Economic Review*, 97, 745–788.

DellaVigna, S. (2009): "Psychology and Economics: Evidence from the Field," *Journal of Economic Literature*, 47, 315–372.

Dunham, W. B., ed. (2009): *New Appleman New York Insurance Law, Second Edition*, New Providence, NJ: LexisNexis.

Einav, L., A. Finkelstein, I. Pascu, and M. R. Cullen (2011): "How General are Risk Preferences? Choice under Uncertainty in Different Domains," Mimeo, Stanford University, http://www.stanford.edu/~leinav/Domains.pdf.

Hey, J. D. and C. Orme (1994): "Investigating Generalizations of Expected Utility Theory Using Experimental Data," *Econometrica*, 62, 1291–1326.

Jullien, B. and B. Salanié (2000): "Estimating Preferences under Risk: The Case of Racetrack Bettors," *Journal of Political Economy*, 108, 503–530.

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

Kőszegi, B. and M. Rabin (2006): "A Model of Reference-Dependent Preferences," *Quarterly Journal of Economics*, 121, 1133–1166.

——— (2007): "Reference-Dependent Risk Attitudes," *American Economic Review*, 97, 1047–1073.

Kliger, D. and O. Levy (2009): "Theories of Choice under Risk: Insights from Financial Martkets," *Journal of Economic Behavior and Organization*, 71, 330–346.

Lattimore, P. K., J. R. Baker, and A. D. Witte (1992): "The Influence of Probability on Risky Choice: A Parametric Examination," *Journal of Economic Behavior and Organization*, 17, 377–400.

Matzkin, R. L. (1991): "Semiparametric Estimation of Monotone and Concave Utility Functions for Polychotomous Choice Models," *Econometrica*, 59, 1315–1327.

McFadden, D. (1974): "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics*, ed. by P. Zarembka, New York: Academic Press.

——— (1981): "Econometric Models of Probabilistic Choice," in *Structural Analysis of Discrete Data with Econometric Applications*, ed. by C. F. Manski, Cambridge, MA: MIT Press.

Post, T., M. J. van den Assem, G. Baltussen, and R. H. Thaler (2008): "Deal or No Deal? Decision Making under Risk in a Large-Payoff Game Show," *American Economic Review*, 98, 38–71.

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

——— (2000): "Compound Invariant Weighting Functions in Prospect Theory," in *Choices, Values, and Frames*, ed. by D. Kahneman and A. Tversky, Cambridge: Cambridge University Press.

Quiggin, J. (1982): "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*, 3, 323–343.

Rabin, M. (1998): "Psychology and Economics," *Journal of Economic Literature*, 36, 11–46.

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

Read, D., G. Loewenstein, and M. Rabin (1999): "Choice Bracketing," *Journal of Risk and Uncertainty*, 19, 171–197.

Snowberg, E. and J. Wolfers (2010): "Explaining the Favorite-Longshot Bias: Is It Risk-Love or Misperceptions?" *Journal of Political Economy*, 118, 723–746.

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

Tanaka, T., C. F. Camerer, and Q. Nguyen (2010): "Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam," *American Economic Review*, 100, 557–571.

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

Vuong, Q. H. (1989): "Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses," *Econometrica*, 57, 307–333.

Table 1: Descriptive Statistics

Core Sample (4170 Households)

Variable	Mean	Std Dev	1st Pctl	99th Pctl
Driver 1 age (years)	54.5	15.4	26	84
Driver 1 female	0.37			
Driver 1 single	0.24			
Driver 1 married	0.51			
Driver 1 credit score	766	113	530	987
Driver 2 indicator	0.42			
Home value (thousands of dollars)	191	125	10	619

Note: Omitted category for driver 1 marital status is divorced or separated.

Table 2: Summary of Deductible Choices

Core Sample (4170 Households)

Deductible	Collision	Comp	Home
$50		5.2	
$100	1.0	4.1	0.9
$200	13.4	33.5	
$250	11.2	10.6	29.7
$500	67.7	43.0	51.9
$1000	6.7	3.6	15.9
$2500			1.2
$5000			0.4

Note: Values are percent of households.

Table 3: Summary of Premium Menus

Core Sample (4170 Households)

Coverage	**Mean**	**Std Dev**	**1st Pctl**	**99th Pctl**
Auto collision premium for $500 deductible	180	100	50	555
Auto comprehensive premium for $500 deductible	115	81	26	403
Home all perils premium for $500 deductible	679	519	216	2511
Cost of decreasing deductible from $500 to $250:				
Auto collision	54	31	14	169
Auto comprehensive	30	22	6	107
Home all perils	56	43	11	220
Savings from increasing deductible from $500 to $1000:				
Auto collision	41	23	11	127
Auto comprehensive	23	16	5	80
Home all perils	74	58	15	294

Note: Annual amounts in dollars.

Table 4: Predicted Claim Probabilities (Annual)

Core Sample (4170 Households)

	Collision	**Comp**	**Home**
Mean	0.069	0.021	0.084
Standard deviation	0.024	0.011	0.044
1st percentile	0.026	0.004	0.024
5th percentile	0.035	0.007	0.034
25th percentile	0.052	0.013	0.053
Median	0.066	0.019	0.076
75th percentile	0.083	0.027	0.104
95th percentile	0.114	0.041	0.163
99th percentile	0.139	0.054	0.233
Correlations	**Collision**	**Comp**	**Home**
Auto collision	1		
Auto comprehensive	0.13	1	
Home all perils	0.28	0.21	1
Premium for coverage with $500 deductible	0.35	0.15	0.18

Note: Each of the correlations is significant at the 1 percent level.

Table 5: Estimates (Model 1)

Core Sample (4170 Households)

	Model 1a: Log Ω(μ)			Model 1b: Ω(μ)			Model 1c: Cubic Spline		
	Estimate		Std Err	Estimate		Std Err	Estimate	95% Bootstrap CI	
r	0.00064	***	0.00010	0.00063	***	0.00004	0.00049	0.0000	0.0009
Ω(μ): constant	-2.71	***	0.03	0.061	***	0.002			
Ω(μ): linear coef	12.03	***	0.30	1.186	***	0.078			
Ω(μ): quadratic coef	-35.15	***	2.17	-2.634	***	0.498			
σ_L	26.31	***	1.14	26.32	***	0.44	30.00	25.18	32.84
σ_M	17.50	***	0.50	17.49	***	0.69	25.20	20.88	27.80
σ_H	68.53	***	5.76	66.89	***	2.11	169.40	112.39	217.38

Note: In Model 1a, we estimate a quadratic Chebyshev polynomial expansion of log Ω(μ). In Model 1b, we estimate a quadraric Chebyshev polynomial expansion of Ω(μ). In Model 1c, we estimate Ω(μ) using an 11-point cubic spline on the interval (0,0.20).

*** Significant at 1 percent level.

Table 6: Economic Significance of Estimates (Model 1a)

	(1)	(2)	(3)	(4)
Standard risk aversion	**r=0**	**r=0.00064**	**r=0**	**r=0.00064**
Probability distortions?	**No**	**No**	**Yes**	**Yes**
μ	**WTP**	**WTP**	**WTP**	**WTP**
0.020	10.00	14.12	41.73	57.20
0.050	25.00	34.80	55.60	75.28
0.075	37.50	51.60	67.30	90.19
0.100	50.00	68.03	77.95	103.51
0.150	75.00	99.84	91.67	120.32

Note: WTP denotes--for a household with claim rate μ, the utility function in equation (7), and the specified utility parameters--the household's maximum willingness to pay to reduce its deductible from \$1000 to \$500 when the premium for coverage with a \$1000 deductible is \$200.

Table 7: Benchmark Estimates (Model 2)

Core Sample (4170 Households)

	r		**Log Ω(μ): constant**		**Log Ω(μ): linear**		**Log Ω(μ): quadratic**	
	Coef	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**
Constant	-7.39 **	0.09	-2.73 **	0.02	12.40 **	0.41	-35.61 **	2.47
Driver 1 age	-1.47 **	0.14	0.18 **	0.05	2.52 **	0.48	1.00 **	0.31
Driver 1 age squared	1.09 **	0.15	0.00	0.05	-4.94 **	0.43	9.92 **	1.82
Driver 1 female	0.15 **	0.04	-0.05 **	0.01	1.57 **	0.28	-12.29 **	1.66
Driver 1 single	0.08	0.05	-0.01	0.01	0.77 **	0.26	-5.93 **	1.91
Driver 1 married	0.09	0.06	-0.03	0.02	1.40 **	0.27	-9.34 **	1.40
Driver 1 credit score	-0.15 **	0.05	-0.02	0.01	2.31 **	0.21	-11.00 **	1.23
Driver 2 indicator	-0.04	0.06	0.00	0.02	-1.44 **	0.38	7.63 **	2.19
Parameter mean	0.00073		-2.73		12.40		-35.61	
Parameter median	0.00056		-2.73		12.42		-34.46	
	Collision		**Comprehensive**		**Home**		**Loglikelihood value**	
σ	27.22 **	0.76	17.91 **	0.48	65.45 **	2.68	-14936.00	

Note: Each variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at 5 percent level.

Table 8: Accounting for Wealth (Models 3-5)

Core Sample (4170 Households)

	r	**Log Ω(μ): constant**	**Log Ω(μ): linear**	**Log Ω(μ): quadratic**
Model 2:				
Parameter mean	0.00073	-2.73	12.40	-35.61
Parameter median	0.00056	-2.73	12.42	-34.46
Model 3:				
Parameter mean	0.00126	-2.83	12.55	-37.61
Parameter median	0.00110	-2.85	12.66	-37.29
Model 4:				
Parameter mean	0.00050	-2.65	12.22	-36.50
Parameter median	0.00028	-2.65	12.24	-35.90
Model 5:				
Parameter mean	0.00056	-2.75	12.00	-35.52
Parameter median	0.00044	-2.75	12.01	-35.42

Note: In Model 5, the mean and median estimates for ρ are 17.84 and 12.81, respectively.

Table 9: Accounting for Unobserved Heterogeneity (Models 2u, 4u, and 5u)

Core Sample (4170 Households)

	r	**Log Ω(μ): constant**	**Log Ω(μ): linear**	**Log Ω(μ): quadratic**
Model 2u:				
Parameter mean	0.00097	-2.73	11.04	-21.78
Parameter median	0.00076	-2.74	11.05	-19.49
Model 4u:				
Parameter mean	0.00081	-2.69	11.68	-26.98
Parameter median	0.00055	-2.71	11.65	-27.08
Model 5u:				
Parameter mean	0.00093	-2.96	13.06	-33.78
Parameter median	0.00082	-2.96	12.81	-31.62

Note: In Model 5u, the mean and median estimates for ρ are 31.73 and 22.72, respectively.

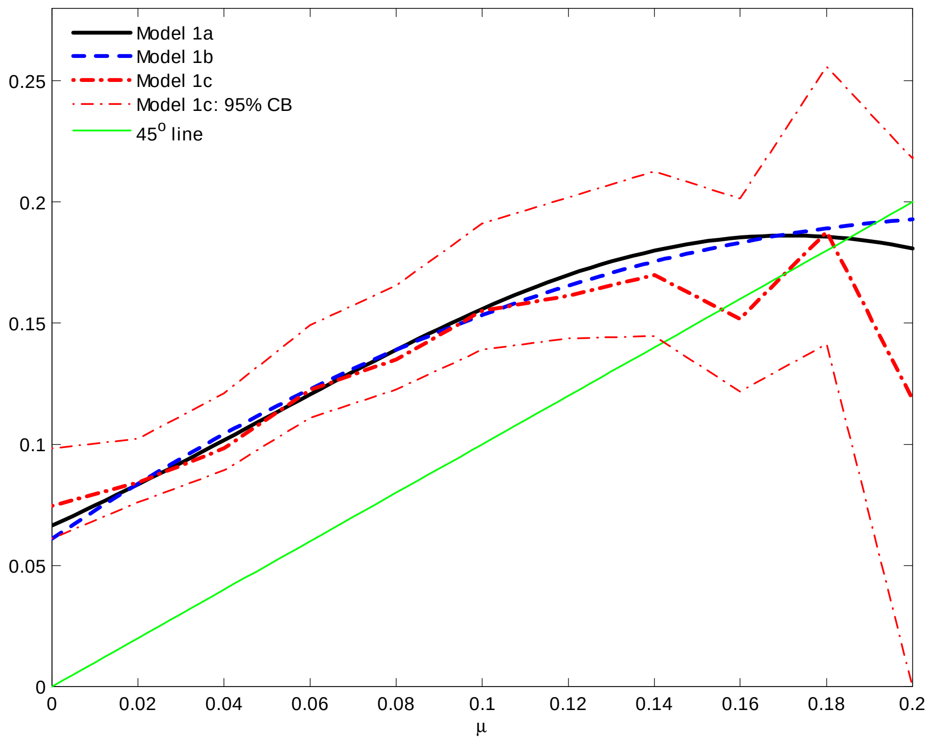

Figure 1: Estimated $\Omega(\mu)$ – Model 1

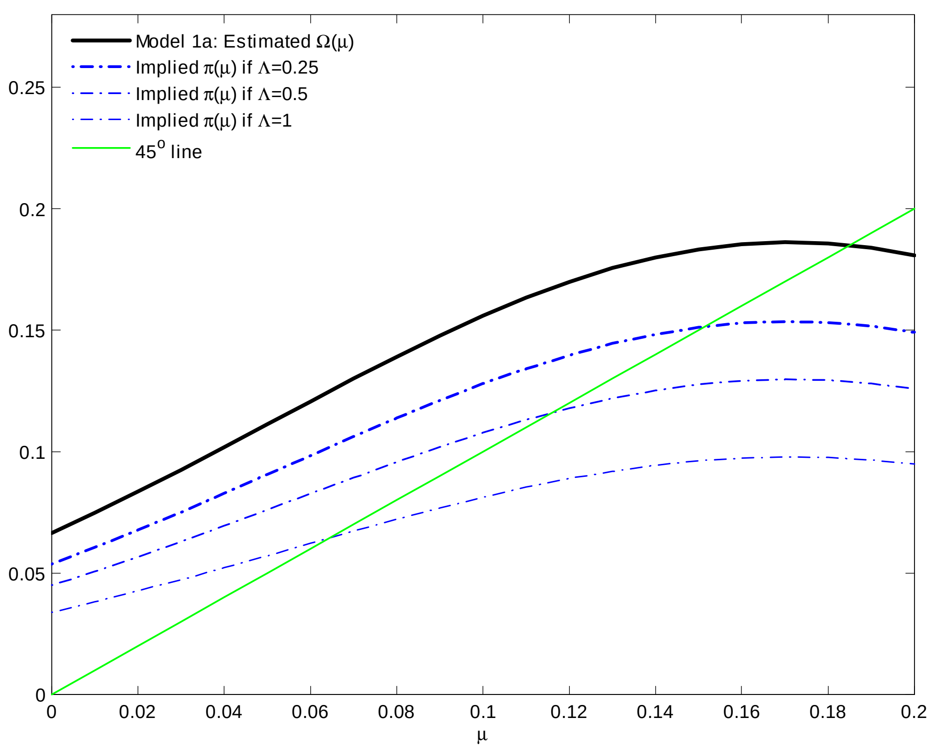

Figure 2: Loss Aversion and Probability Weighting

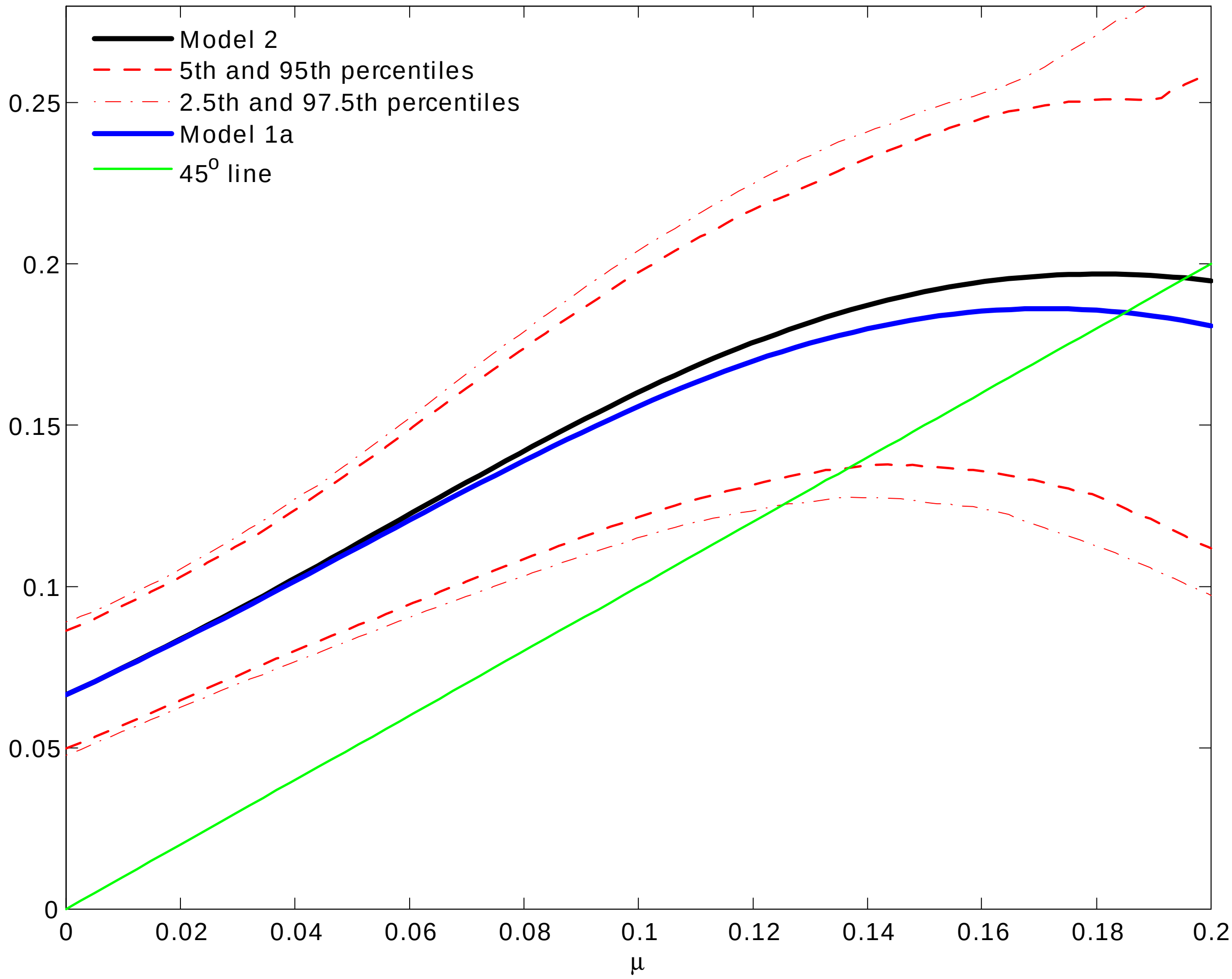

Figure 3: Mean Estimated $\Omega(\mu)$ – Model 2

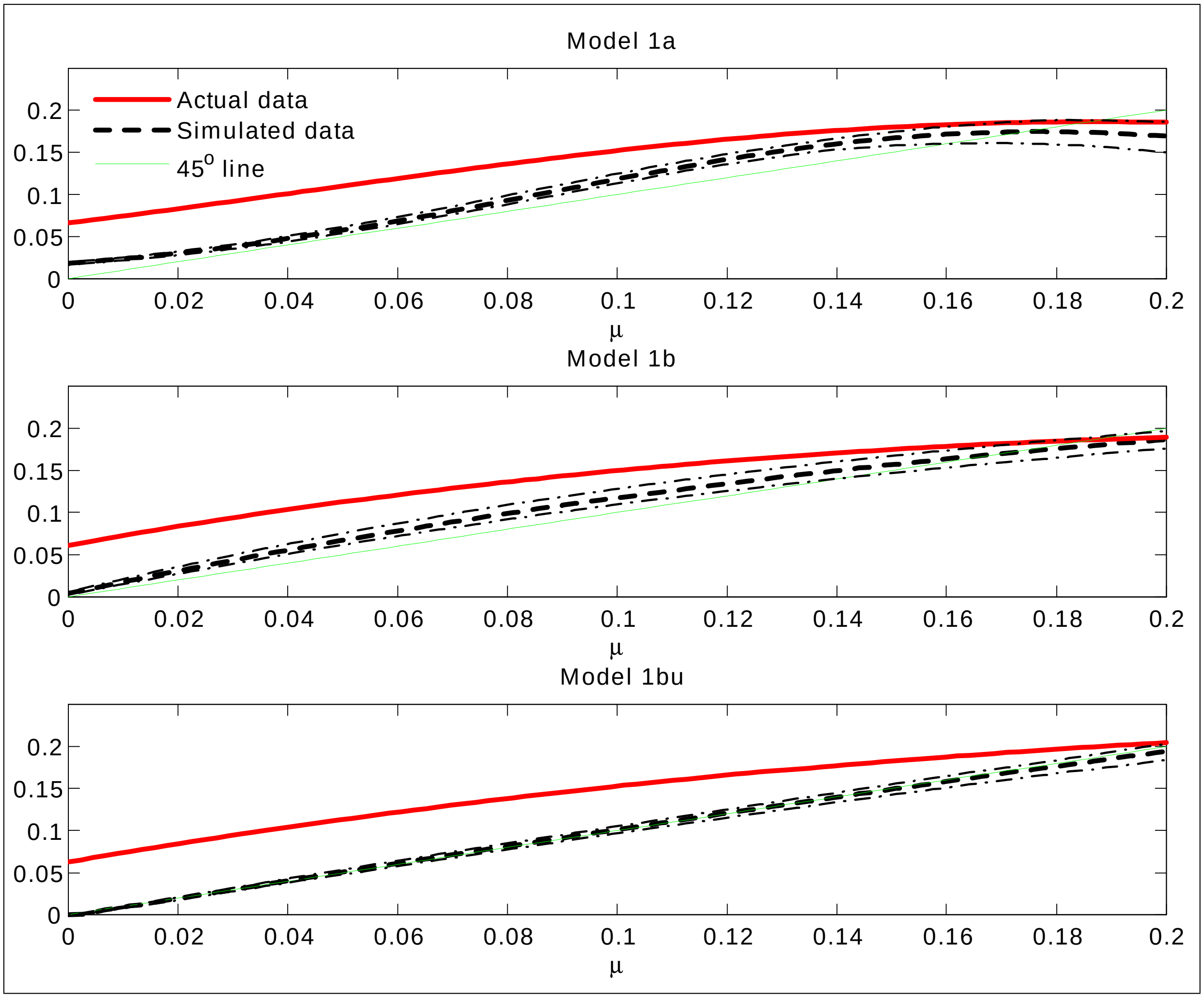

Figure 4: Mean Estimated $\Omega(\mu)$ – Unobserved Heterogeneity in r

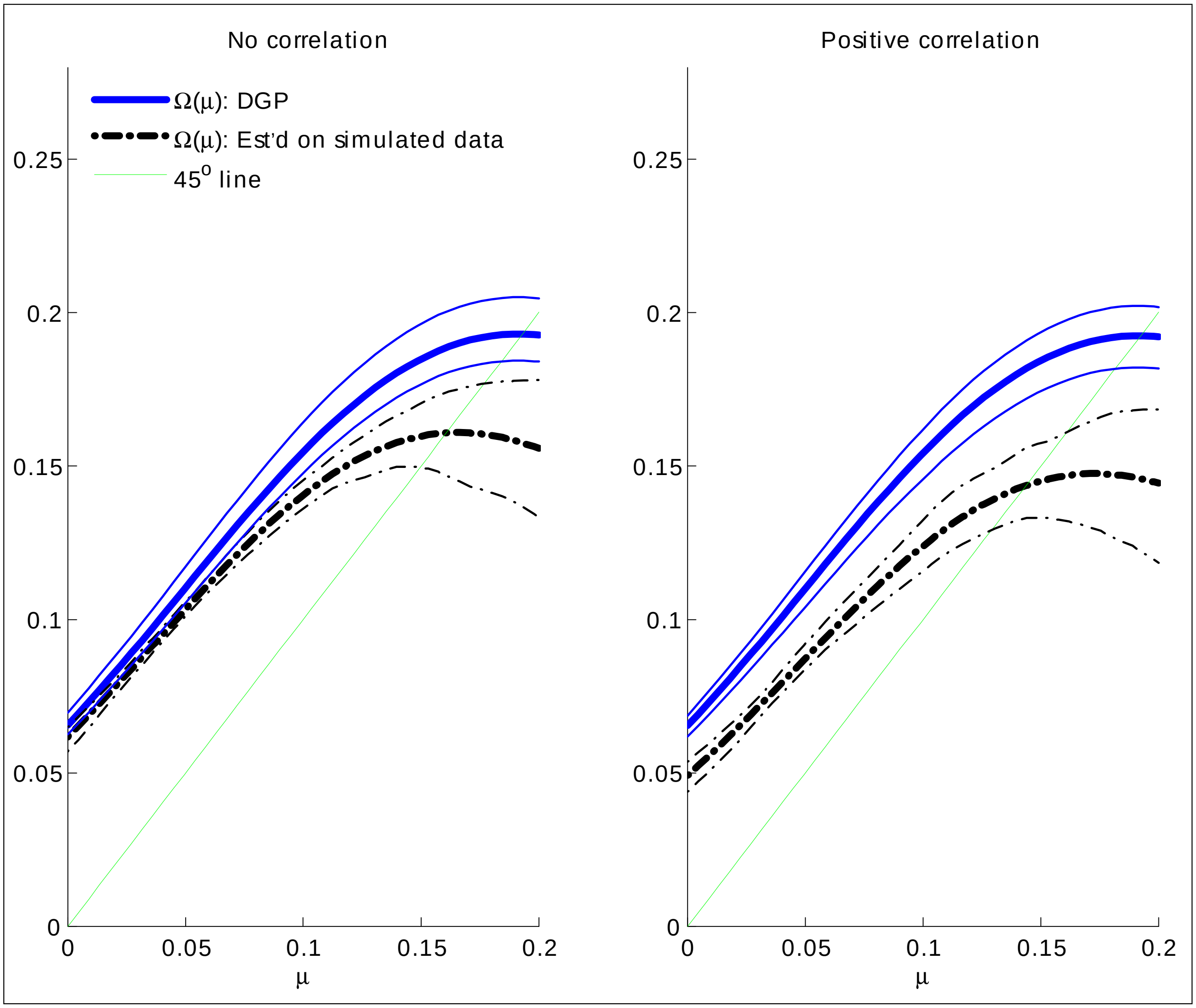

Figure 5: Unobserved Heterogeneity in r and $\Omega(\mu)$ – Model 1a

Appendix (Not for Publication)
to
The Nature of Risk Preferences: Evidence from Insurance Choices

Levon Barseghyan
Cornell University

Francesca Molinari
Cornell University

Ted O'Donoghue
Cornell University

Joshua C. Teitelbaum
Georgetown University

June 30, 2011

Contents

A Coverage Descriptions

Auto collision coverage pays for damage to the insured vehicle caused by a collision with another vehicle or object, without regard to fault. *Auto comprehensive* coverage pays for damage to the insured vehicle from all other causes (e.g., theft, fire, flood, windstorm, glass breakage, vandalism, hitting or being hit by an animal, or by falling or flying objects), without regard to fault. If the insured vehicle is stolen, auto comprehensive coverage also provides a certain amount per day for transportation expenses (e.g., rental car or public transportation). *Home all perils* coverage pays for damage to the insured home from all causes (e.g., fire, windstorm, hail, tornadoes, vandalism, or smoke damage), except those that are specifically excluded (e.g., flood, earthquake, or war). For simplicity, we often refer to home all perils simply as *home*.

B Kőszegi-Rabin Model

In this section, we describe the Kőszegi-Rabin (KR) model and derive the utility function in equation (4) from Section 3.3.1. In the KR model, the utility from choosing lottery $Y \equiv (y_n, q_n)_{n=1}^N$ given a reference lottery $\tilde{Y} \equiv (\tilde{y}_m, \tilde{q}_m)_{m=1}^M$ is

$$U(Y|\tilde{Y}) \equiv \sum_{n=1}^{N} \sum_{m=1}^{M} q_n \tilde{q}_m \left[u(y_n) + v(y_n|\tilde{y}_m) \right].$$

The function u represents standard "intrinsic" utility defined over final wealth states, just as in the expected utility model. The function v represents "gain-loss" utility that results from experiencing gains or losses relative to the reference point. For v, KR use

$$v(y|\tilde{y}) = \begin{cases} \eta \left[u(y) - u(\tilde{y}) \right] & \text{if } u(y) > u(\tilde{y}) \\ \eta\lambda \left[u(y) - u(\tilde{y}) \right] & \text{if } u(y) \leq u(\tilde{y}) \end{cases}.$$

In this formulation, the magnitude of gain-loss utility is determined by the intrinsic utility gain or loss relative to consuming the reference point. Moreover, gain-loss utility takes a two-part linear form, where $\eta \geq 0$ captures the importance of gain-loss utility relative to intrinsic utility and $\lambda \geq 1$ captures loss aversion. The model reduces to expected utility when $\eta = 0$ or $\lambda = 1$.

KR propose that the reference lottery equals recent expectations about outcomes—i.e., if a household expects to face lottery $\tilde{Y}$, then its reference lottery becomes $\tilde{Y}$. However, because situations vary in terms of when a household deliberates about its choices and when it commits to its choices, KR offer a number of solution concepts for the determination of the reference lottery. We assume that the reference lottery is determined according to what KR call a "choice-acclimating personal equilibrium" (CPE). Formally:

Definition (CPE)**.** *Given a choice set* $\mathcal{Y}$*, a lottery* $Y \in \mathcal{Y}$ *is a* choice-acclimating personal equilibrium *if for all* $Y' \in \mathcal{Y}$*,* $U(Y|Y) \geq U(Y'|Y')$*.*

In a CPE, a household's reference lottery corresponds to its choice. KR argue that CPE is appropriate in situations where the household commits to a choice well in advance of the resolution of uncertainty, and thus it knows that by the time the uncertainty is resolved and it experiences utility, it will have become accustomed to its choice and hence expect the lottery induced by its choice.[1] In particular, KR suggest that CPE is the appropriate solution concept for insurance applications.

Under the KR model using CPE, the utility to the household from choosing deductible lottery $L_d = (-p_d, 1-\mu; -p_d - d, \mu)$ is

$$\begin{aligned} U(L_d|L_d) &= (1-\mu)u(w-p_d) + \mu u(w-p_d-d) \\ &\quad -\Lambda(1-\mu)\mu[u(w-p_d) - u(w-p_d-d)], \end{aligned} \tag{A.1}$$

where $\Lambda \equiv \eta(\lambda - 1)$ and w is the household's initial wealth. From equation (A.1), it is clear that we can not separately identify the parameters η and λ. Instead, we estimate the product $\eta(\lambda-1) \equiv \Lambda$.[2] Under the assumption

$$\frac{u(w+\Delta)}{u'(w)} - \frac{u(w)}{u'(w)} = \Delta - \frac{r}{2}\Delta^2$$

from Section 3.2, equation (A.1) becomes

$$\begin{aligned} U(L_d) &= -\left[p_d + \mu d\right] - \frac{r}{2}\left[(1-\mu)\left(p_d\right)^2 + \mu\left(p_d+d\right)^2\right] \\ &\quad -\Lambda\left(1-\mu\right)\mu\left[d + \frac{r}{2}\left[(p_d+d)^2 - (p_d)^2\right]\right], \end{aligned}$$

which is equation (4) from Section 3.3.1.

[1] The assumption that the household commits to its choice is important. Suppose instead that the household has the opportunity to revise its choice just before the uncertainty is resolved. Then even after "choosing" Y and coming to expect it, if $U(Y'|Y) > U(Y|Y)$ the household would want to revise its choice just before the uncertainty is resolved. KR propose alternative solution concepts that are more appropriate in such situations, where a household thinks about the problem in advance but does not commit to a choice until just before the uncertainty is resolved.

[2] The inability to separately identify η and λ applies to any application of CPE, and not just deductible lotteries, because for any lottery Y, η and λ appear in $U(Y|Y)$ only as the product $\eta(\lambda-1)$. For other solution concepts, η and λ become separately identified.

C Model Implications

In this section, we prove Properties 1, 2, 3, and 4 from Section 3.3.4. Take any three deductible options $a, b, c \in \mathcal{D}$, $a > b > c$. For a household with premium p_a for deductible a and claim probability μ, define $\widetilde{p}_b(p_a, \mu)$ as the premium for deductible b that makes the household indifferent between a and b, and define $\widetilde{p}_c(p_a, \mu)$ as the the premium for deductible c that makes the household indifferent between a and c. In other words, $\widetilde{p}_b(p_a, \mu) - p_a$ reflects the household's maximum willingness to pay (WTP) to reduce its deductible from a to b, and $\widetilde{p}_c(p_a, \mu) - p_a$ reflects the household's WTP to reduce its deductible from a to c. In what follows, we simplify notation by suppressing the explicit dependence of $\widetilde{p}_b$ and $\widetilde{p}_c$ on p_a, μ, r, and $\Omega(\mu)$. We also suppress the argument of Ω. In addition, let L_a denote the deductible lottery associated with deductible a at premium p_a.

Recall equation (7) from Section 3.3.3:

$$U(L_d) = -\left[p_d + \Omega d\right] - \frac{r}{2}\left[(1-\Omega)\left(p_d\right)^2 + \Omega\left(p_d + d\right)^2\right].$$

Define $p(x)$ as the premium for deductible x such that the household is indifferent between the resulting lottery and lottery L_a. Hence, $p(b) = \widetilde{p}_b$ and $p(c) = \widetilde{p}_c$. Applying equation (7), $p(x)$ is defined by each of the following equations (both of which we use below):[3]

$$-p(x) - \Omega x - \frac{r}{2}\left[(1-\Omega)(p(x))^2 + \Omega(p(x)+x)^2\right] = U(L_a) \tag{A.2}$$

$$(p(x) - p_a) - \Omega\left(a - x\right) + \frac{r}{2}(p(x)^2 - p_a^2) + \Omega\frac{r}{2}\left\{(x^2 - a^2) + 2(p(x)x - p_a a)\right\} = 0. \tag{A.3}$$

For the proofs below, it is useful to define W and V as

$$\begin{aligned} W(p, x, r, \Omega) &\equiv -p - \Omega x - \frac{r}{2}\left[(1-\Omega)(p)^2 + \Omega(p+x)^2\right] \\ V(p, x, r, \Omega) &\equiv (p - p_a) - \Omega\left(a - x\right) + \frac{r}{2}(p^2 - p_a^2) + \Omega\frac{r}{2}\left\{(x^2 - a^2) + 2(px - p_a a)\right\}. \end{aligned}$$

Our first lemma establishes that $p(x)$ is well behaved.

Lemma 1. *For any $r \geq 0$, $\Omega \in (0,1)$, $p_a > 0$, and $x \leq a$, $p(x)$ is a continuous and differentiable function with $dp/dx < 0$ (and thus $p(c) > p(b) > p_a$).*

Proof. It is straightforward to derive that W is twice differentiable and satisfies the conditions

[3] These equations are equivalent, where the latter merely expands $U(L_a)$ and rearranges terms.

of the implicit function theorem. Thus $p(x)$ is a continuous and differentiable function, and

$$\frac{dp}{dx} = \frac{-\frac{\partial W}{\partial x}}{\frac{\partial W}{\partial p}} = \frac{-\Omega\left[1 + r(p(x)+x)\right]}{1 + rp(x) + r\Omega x} < 0.$$

□

Our second lemma states that standard risk aversion implies that a household's WTP to reduce its deductible is strictly greater than the expected reduction in the deductible paid (evaluated at a claim probability of Ω).

Lemma 2. *For any $x' < x \leq a$, if $r = 0$ then $p(x') - p(x) = \Omega(x - x')$, and if $r > 0$ then $p(x') - p(x) > \Omega(x - x')$.*

Proof. The result for $r = 0$ is straightforward. For $r > 0$, define $\tilde{V}$ as

$$\tilde{V}(p, x', r, \Omega) \equiv [p - p(x)] - \Omega\,(x - x') + \frac{r}{2}(p^2 - p(x)^2) + \Omega\frac{r}{2}\left\{(x'^2 - x^2) + 2(px' - p(x)x)\right\},$$

in which case $p(x')$ is defined by $\tilde{V}(p(x'), x', r, \Omega) = 0$. Note that $p = \Omega(x - x') + p(x)$ implies

$$\begin{aligned}
\tilde{V}(p, x', r, \Omega) \;\; &\equiv \;\; [(\Omega(x - x') + p(x)) - p(x)] - \Omega\,(x - x') + \frac{r}{2}((\Omega(x - x') + p(x))^2 - (p(x))^2) \\
&\quad + \Omega\frac{r}{2}\left\{(x'^2 - x^2) + 2((\Omega(x - x') + p(x))x' - p(x)x)\right\} \\
&= \frac{r}{2}(\Omega^2(x - x')^2 + 2p(x)\Omega(x - x')) \\
&\quad + \Omega\frac{r}{2}\left\{(x'^2 - x^2) + 2(\Omega(x - x')x' + p(x)(x' - x))\right\} \\
&= \frac{r}{2}\Omega(x - x')[\Omega(x + x') - (x + x')] < 0.
\end{aligned}$$

Since $\partial\tilde{V}/\partial p = 1 + rp + \Omega r x' > 0$, it follows that $p(x') > \Omega(x - x') + p(x)$. □

Property 1 establishes the relationship between the magnitude of willingness to pay and risk preferences.

Property 1. *For any $x < a$, $p(x)$ is strictly increasing in r and Ω.*

Proof. By implicit function theorem:

$$\frac{\partial p(x)}{\partial r} = \frac{-\frac{\partial V}{\partial r}}{\frac{\partial V}{\partial p}} \qquad \text{and} \qquad \frac{\partial p(x)}{\partial \Omega} = \frac{-\frac{\partial V}{\partial \Omega}}{\frac{\partial V}{\partial p}}.$$

Note that

$$\frac{\partial V}{\partial r} = -\frac{1}{r}[(p(x) - p_a) - \Omega\,(a - x)] < 0,$$

where the equality uses equation (A.3) and the inequality follows from Lemma 2. Note further that

$$\frac{\partial V}{\partial \Omega} = -\frac{1}{\Omega}[(p(x) - p_a) + \frac{r}{2}(p(x)^2 - p_a^2)] < 0,$$

where the equality uses equation (A.3) and the inequality follows from Lemma 1. Finally, given that $\partial V / \partial p = 1 + rp + \Omega rx > 0$, it follows that $\frac{\partial p(x)}{\partial r} > 0$ and $\frac{\partial p(x)}{\partial \Omega} > 0$. □

We next establish Property 2, which shows that a risk averse household's WTP to avoid an incremental loss depends positively on the magnitude of the absolute loss.

Property 2. *If $r = 0$ then $\frac{p(b)-p_a}{p(c)-p(b)} = \frac{a-b}{b-c}$. If $r > 0$ then $\frac{p(b)-p_a}{p(c)-p(b)} > \frac{a-b}{b-c}$.*

Proof. The result for $r = 0$ is straightforward. From Lemma 1, $p(x)$ is continuous and differentiable, and thus

$$p(b) - p(a) = \int_b^a \left(-\frac{dp}{dx}\right) dx \qquad \text{and} \qquad p(c) - p(b) = \int_c^b \left(-\frac{dp}{dx}\right) dx.$$

From the proof of Lemma 1, $-\frac{dp}{dx} = \frac{\Omega[1+rp(x)+rx]}{1+rp(x)+r\Omega x} > 0$, and thus

$$\frac{d\left[-\frac{dp}{dx}\right]}{dx} = \Omega \frac{r(1-\Omega)(1+rp-rx\frac{dp}{dx})}{[1+rp(x)+r\Omega x]^2} > 0.$$

In words, $-\frac{dp}{dx}$ reflects the household's marginal willingness to pay to reduce its deductible, and $-\frac{dp}{dx} > 0$ reflects that a household is indeed willing to pay a higher premium to reduce its deductible. More importantly, $d\left[-\frac{dp}{dx}\right]/dx > 0$ reflects that the larger is its deductible, the larger is the household's marginal willingness to pay to reduce that deductible (or equivalently the smaller is its deductible, the smaller is the household's marginal willingness to pay to reduce that deductible). Finally, $d\left[-\frac{dp}{dx}\right]/dx > 0$ implies

$$\begin{aligned} p(b) - p(a) &= \int_b^a \left(-\frac{dp}{dx}\right) dx > (a-b)\left(-\left.\frac{dp}{dx}\right|_{x=b}\right) \\ p(c) - p(b) &= \int_c^b \left(-\frac{dp}{dx}\right) dx < (b-c)\left(-\left.\frac{dp}{dx}\right|_{x=b}\right), \end{aligned}$$

which together imply $\frac{p(b)-p(a)}{a-b} > \frac{p(c)-p(b)}{b-c}$, from which the result follows. □

Property 3 states that the implication of risk aversion in Property 2 is magnified as a household gets more risk averse.

Property 3. *Holding Ω fixed, the ratio $\frac{p(b)-p_a}{p(c)-p(b)}$ is strictly increasing in r.*

Proof. Note that $d\left(\frac{p(b)-p_a}{p(c)-p(b)}\right)/dr > 0$ if and only if $\frac{1}{p(b)-p_a}\frac{\partial p(b)}{\partial r} > \frac{1}{p(c)-p_a}\frac{\partial p(c)}{\partial r}$. Applying $\frac{\partial p(x)}{\partial r}$ from the proof of Property 1,

$$\begin{aligned}
\frac{1}{p(b)-p_a}\frac{\partial p(b)}{\partial r} &= \frac{1}{p(b)-p_a}\frac{\frac{1}{r}[(p(b)-p_a)-\Omega\,(a-b)]}{1+rp(b)+\Omega rb}, \\
\frac{1}{p(c)-p_a}\frac{\partial p(c)}{\partial r} &= \frac{1}{p(c)-p_a}\frac{\frac{1}{r}[(p(c)-p_a)-\Omega\,(a-c)]}{1+rp(c)+\Omega rc}.
\end{aligned}$$

We have $\frac{1}{p(b)-p_a}\frac{1}{r}[(p(b)-p_a)-\Omega\,(a-b)] > \frac{1}{p(c)-p_a}\frac{1}{r}[(p(c)-p_a)-\Omega\,(a-c)]$, because

$$\begin{aligned}
(p(c)-p_a)\,[(p(b)-p_a)-\Omega\,(a-b)] &> (p(b)-p_a)\,[(p(c)-p_a)-\Omega\,(a-c)] \\
\Leftrightarrow (p(b)-p_a)\,(a-c) &> (p(c)-p_a)\,(a-b)\,,
\end{aligned}$$

where the last inequality follows from Property 2—specifically, because $\frac{p(c)-p_a}{a-c}$ is a convex combination of $\frac{p(b)-p_a}{a-b}$ and $\frac{p(c)-p(b)}{b-c}$, Property 2 implies $\frac{p(c)-p_a}{a-c} < \frac{p(b)-p_a}{a-b}$. Finally, we have $1+rp(b)+\Omega rb < 1+rp(c)+\Omega rc$, because

$$\begin{aligned}
rp(b)+\Omega rb &< rp(c)+\Omega rc \\
\Leftrightarrow \Omega\,(b-c) &< p(c)-p(b),
\end{aligned}$$

where the last inequality follows from Lemma 2. The result follows. □

We conclude with the key property for identification, which establishes that different pairs of r and Ω have different implications for willingness to pay (provided there are at least three deductible options on the menu).

Property 4. *If $p(c)$ is the same for $(r,\Omega(\mu))$ and a different $(r',\Omega(\mu)')$, then $p(b)$ is different for $(r,\Omega(\mu))$ and $(r',\Omega(\mu)')$. In particular, if $r > r'$ (in which case $\Omega(\mu) < \Omega(\mu)'$), then $p(b)$ is larger for $(r,\Omega(\mu))$.*

Proof. For a fixed $p(c)$, define $\Omega^c(r)$ by $V\left(p(c),c,r,\Omega^c(r)\right)=0$, so that any pair $(r,\Omega^c(r))$ yields the same $p(c)$. Then

$$\frac{d\Omega^c}{dr} = \frac{-\frac{\partial V}{\partial r}}{\frac{\partial V}{\partial \Omega}} = -\frac{\frac{1}{r}[(p(c)-p_a)-\Omega^c(r)\,(a-c)]}{\frac{1}{\Omega^c(r)}[(p(c)-p_a)+\frac{r}{2}(p(c)^2-p_a^2)]}.$$

Next, define $\check{p}_b(r)$ by $V(\check{p}_b(r),b,r,\Omega^c(r)) = 0$, so that $\check{p}_b(r)$ is the $p(b)$ associated with pair $(r,\Omega^c(r))$. The goal is to show that $d\check{p}_b(r)/dr > 0$, from which the result follows.

Differentiating $V(\check{p}_b(r), b, r, \Omega^c(r)) = 0$ yields

$$\frac{d\left[V(\check{p}_b(r), b, r, \Omega^c(r))\right]}{dr} = \frac{\partial V}{\partial p}\frac{d\check{p}_b(r)}{dr} + \frac{\partial V}{\partial r} + \frac{\partial V}{\partial \Omega}\frac{d\Omega^c}{dr} = 0.$$

Note that

$$\begin{aligned} \frac{\partial V}{\partial r} + \frac{\partial V}{\partial \Omega}\frac{d\Omega^c}{dr} &= -\frac{1}{r}[(\check{p}_b(r) - p_a) - \Omega^c(r)\,(a-b)] \\ &\quad + \frac{1}{\Omega^c(r)}\left[(\check{p}_b(r) - p_a) + \frac{r}{2}(\check{p}_b(r)^2 - p_a^2)\right] \frac{\frac{1}{r}[(p(c) - p_a) - \Omega^c(r)\,(a-c)]}{\frac{1}{\Omega^c(r)}[(p(c) - p_a) + \frac{r}{2}(p(c)^2 - p_a^2)]}. \end{aligned}$$

We have

$$(p(c) - p_a)\frac{1}{r}[(\check{p}_b(r) - p_a) - \Omega^c(r)\,(a-b)] > (\check{p}_b(r) - p_a)\frac{1}{r}[(p(c) - p_a) - \Omega^c(r)\,(a-c)]$$

as in the proof of Property 3. In addition,

$$(p(c) - p_a)[(\check{p}_b(r) - p_a) + \frac{r}{2}(\check{p}_b(r)^2 - p_a^2)] < (\check{p}_b(r) - p_a)[(p(c) - p_a) + \frac{r}{2}(p(c)^2 - p_a^2)],$$

because

$$\begin{aligned} (p(c) - p_a)(\check{p}_b(r)^2 - p_a^2) &< (\check{p}_b(r) - p_a)(p(c)^2 - p_a^2) \\ \Leftrightarrow (p(c) - p_a)(\check{p}_b(r) - p_a)(\check{p}_b(r) + p_a) &< (\check{p}_b(r) - p_a)(p(c) - p_a)(p(c) + p_a) \\ \Leftrightarrow \check{p}_b(r) &< p(c), \end{aligned}$$

where the last inequality follows from Lemma 1. Together, these imply $\frac{\partial V}{\partial r} + \frac{\partial V}{\partial \Omega}\frac{d\Omega^c}{dr} < 0$, and therefore $\frac{\partial V}{\partial p}\frac{d\check{p}_b(r)}{dr} > 0$. Hence, $\frac{\partial V}{\partial p} > 0$ implies $\frac{d\check{p}_b(r)}{dr} > 0$. The result follows. □

D Robustness to Unobserved Heterogeneity in Preferences

In this section, we describe in detail the robustness exercises reported in Section 7. In each exercise, we generate simulated deductible choices using an alternative model and then estimate our model on the "simulated data"—i.e., the actual data but with the simulated deductible choices substituted for the actual deductible choices.

At the simulation stage of each exercise, we allow for both observed and unobserved heterogeneity in preferences and risk. However, because of severe computational burden, at the estimation stage we use Models 1a and 1b, which assume homogenous preferences and do not allow for unobserved heterogeneity in risk (although in one case in which we find some small bias in favor of probability distortions, we also estimate Model 1b allowing for unobserved heterogeneity in risk).[4]

D.1 Unobserved Heterogeneity in Standard Risk Aversion

We first consider whether unobserved heterogeneity in standard risk aversion could lead us to find large probability distortions when in fact none exist.

Alternative Model We use a variant of the Cohen-Einav model to generate the simulated deductible choices. For each coverage j, each household i has a subjective claim probability μ_{ij} and coefficient of absolute risk aversion r_{ij}.[5] Given its menu of premium-deductible pairs $\{(p_d, d) : d \in \mathcal{D}\}$, the household chooses the deductible that maximizes its expected utility (i.e., maximizes equation (1)).

Data Generating Process (DGP) We generate simulated deductible choices for each household i and coverage j, as follows. First, we generate the subjective claim probability $\mu_{ij} = 1-\exp(-\lambda_{ij})$ using the claim rate $\lambda_{ij} = \widehat{\lambda}_{ij} \exp(\varepsilon_{ij}^{\lambda})$, where $\widehat{\lambda}_{ij}$ is household i's predicted claim rate for coverage j and $\varepsilon_{ij}^{\lambda}$ is a draw from a gamma distribution with unit mean and variance $\widehat{\phi}_{ij}$ (the estimated variance from the claim rate regression for coverage j updated for household i's (ex post) claims experience). Second, we generate standard risk aversion $r_{ij} = \exp(\widehat{\beta}_r Z_i + \varepsilon_{ij}^{r})$, where $\widehat{\beta}_r$ is the vector of coefficient estimates assuming only standard risk aversion and ε_{ij}^{r} is a draw from a normal distribution with variance α. We consider several values of α and several correlations between $\varepsilon_{ij}^{\lambda}$ and ε_{ij}^{r} defined as $\rho \equiv corr(\varepsilon_{ij}^{\lambda}, \varepsilon_{ij}^{r})$. We assume that $\varepsilon_{ij}^{\lambda}$ and ε_{ij}^{r} are not correlated across coverages. Finally, we generate the simulated deductible choice $\widetilde{D}_{ij}$ by applying the expected utility model (equation (1)) when the household faces menu P_{ij}. In the end, the simulated data comprise $\{\widetilde{D}_{ij}, \widehat{\mu}_{ij}, P_{ij}\}$, where

[4] In each exercise, we generate 100 independent simulated datasets of 4170 households making the three deductible choices. We then estimate our model on each dataset, and report the mean estimated $\Omega(\mu)$ along with the 2.5th and 97.5th percentiles.

[5] This formulation permits household i's risk aversion r_{ij} to be coverage specific. Thus, it is as if each household-coverage observation is coming from a distinct household. This approach permits us to assume a strong correlation between unobserved heterogeneity in risk and unobserved heterogeneity in risk aversion, while still assuming that unobserved heterogeneity in risk is uncorrelated across coverages. An alternative approach is to instead assume that household i's risk aversion r_i is the same across coverages, but then to further assume that unobserved heterogeneity in risk is perfectly correlated across coverages so that we can permit a strong correlation between unobserved heterogeneity in risk and unobserved heterogeneity in risk aversion (see Section D.3 below). Under that approach, the results do not change in any noticeable way.

$\widetilde{D}_{ij}$ is household i's simulated deductible choice for coverage j and $\widehat{\mu}_{ij}$ and P_{ij} come from the actual data.

Parameter Values In the DGP, the predicted claim rate $\widehat{\lambda}_{ij}$, the distribution of $\varepsilon_{ij}^{\lambda}$, and the vector of coefficient estimates $\widehat{\beta}_r$ are taken from our main analysis. In contrast, α and ρ are new parameters that emerge when we introduce unobserved heterogeneity in risk aversion, and because our main analysis does not permit such heterogeneity, it provides no guidance on their magnitudes.

We consider three values of ρ. First, as a benchmark, we consider $\rho = 0$. Second, we consider $\rho = 0.84$, which is the correlation estimated by Cohen and Einav (2007). Finally, for symmetry, we also consider $\rho = -0.84$.

In choosing α, we face a tradeoff. Increasing α increases the potential for bias, but at the same time it increases the frequency with which the alternative model predicts extreme deductible choices (because it increases the number of households with a very low or very high r_{ij}). In our simulations, we found that, even at $\alpha = 1$, the DGP already implies a substantially higher frequency of extreme deductible choices than we observe in the actual data. Hence, we focus on two values, $\alpha = 1$ and $\alpha = 3.15$, where the latter is the value estimated by Cohen and Einav (2007).

Estimating our Model using the Simulated Data After choosing α and ρ and generating the simulated data, we estimate Models 1a and 1b—i.e., homogenous preferences with a quadratic Chebyshev polynomial expansion of $\ln \Omega(\mu)$ (Model 1a) or $\Omega(\mu)$ (Model 1b). Figures A.1 and A.2 depict the estimated probability distortion function $\Omega(\mu)$ when we estimate Models 1a and 1b, respectively. Each panel presents, for a specific combination of α and ρ, the mean estimated $\Omega(\mu)$ using the simulated data (along with the 2.5th and 97.5th percentiles). For comparison, it also depicts the mean estimated $\Omega(\mu)$ using the actual data. For each model, the message is the same. When $\varepsilon_{ij}^{\lambda}$ and ε_{ij}^{r} are uncorrelated ($\rho = 0$), there is some bias in favor of finding probability distortions (which is roughly the same for both variances), but it is nothing close to the estimated $\Omega(\mu)$ using the actual data. When $\varepsilon_{ij}^{\lambda}$ and ε_{ij}^{r} are strongly positively correlated, the bias is somewhat larger (and more pronounced for the higher variance), but again nothing close to the estimated $\Omega(\mu)$ using the actual data. When $\varepsilon_{ij}^{\lambda}$ and ε_{ij}^{r} are strongly negatively correlated, the bias is small when the variance is small, and the bias disappears when the variance is high.

In sum, the probability distortions we find using the simulated data are rather small and significantly less than the large probability distortions we find using the actual data. In order to investigate the source of the bias towards finding probability distortions, we also estimate

Model 1b when we allow for unobserved heterogeneity in risk (Model 1bu). As Figure A.3 illustrates, when we estimate Model 1bu on the simulated data, we now find no bias in favor of finding probability distortions. By contrast, when we estimate Model 1bu on the actual data, the probability distortions do not go away—indeed, they are very similar to the probability distortions we find when we estimate Model 1b on the actual data. Moreover, recall that in our analysis in Section 5 (where we allow for observed heterogeneity in preferences) the estimated probability distortions also persist when we allow for unobserved heterogeneity in risk (Models 2u, 4u, and 5u). These results lead us to conclude that unobserved heterogeneity in standard risk aversion cannot explain the large probability distortions we find in the data.

FIGURES A.1, A.2 & A.3

D.2 Unobserved Heterogeneity in Standard Risk Aversion and Probability Distortions

We next consider whether unobserved heterogeneity in both r and $\Omega(\mu)$ could bias our results in favor of finding probability distortions.

Alternative Model Here we use equation (7) to generate the simulated deductible choices. For each coverage j, each household i has a subjective claim probability μ_{ij}, a coefficient of absolute risk aversion r_{ij}, and a probability distortion function $\Omega_{ij}(\mu)$. Given its menu of premium-deductible pairs $\{(p_d, d) : d \in \mathcal{D}\}$, the household chooses the deductible that maximizes equation (7).

Data Generating Process (DGP) We generate simulated deductible choices for each household i and coverage j, as follows. First, we generate the subjective claim probability μ_{ij} exactly as in Section D.1. Second, we generate standard risk aversion $r_{ij} = \exp(\widehat{\beta}_r Z_i + \varepsilon_{ij}^r)$ as in Section D.1, except that now $\widehat{\beta}_r$ is the vector of coefficient estimates from Model 2. Third, we generate probability distortions $\Omega_{ij}(\mu) = \exp(\widehat{\beta}_{\Omega,1} Z_i + \widehat{\beta}_{\Omega,2} Z_i \mu + \widehat{\beta}_{\Omega,3} Z_i \mu^2 + \varepsilon_{ij}^{\Omega})$, where $\widehat{\beta}_{\Omega,1}$, $\widehat{\beta}_{\Omega,2}$, and $\widehat{\beta}_{\Omega,3}$ are the vectors of coefficient estimates from Model 2 and $\varepsilon_{ij}^{\Omega}$ is a draw from a normal distribution with variance α (same as ε_{ij}^r). As before, we consider several values of α and several correlation structures among $\varepsilon_{ij}^{\lambda}$, ε_{ij}^r, and $\varepsilon_{ij}^{\Omega}$, and we assume that $\varepsilon_{ij}^{\lambda}$, ε_{ij}^r, and $\varepsilon_{ij}^{\Omega}$ are not correlated across coverages. Finally, we generate the simulated deductible choice $\widetilde{D}_{ij}$ by applying our model (equation (7)) when the household faces menu P_{ij}. In the end, the simulated data comprise $\{\widetilde{D}_{ij}, \widehat{\mu}_{ij}, P_{ij}\}$, where $\widetilde{D}_{ij}$ is household i's simulated deductible choice for coverage j and $\widehat{\mu}_{ij}$ and P_{ij} come from the actual data.

Parameter Values In the DGP, the predicted claim rate $\widehat{\lambda}_{ij}$, the distribution of $\varepsilon^{\lambda}_{ij}$, and the coefficient vectors $\widehat{\beta}_r$ and $\widehat{\beta}_\Omega$ are taken from our main analysis. The new parameters (relative to our model) are α and the correlations among $\varepsilon^{\lambda}_{ij}$, ε^{r}_{ij}, and $\varepsilon^{\Omega}_{ij}$. For α, we restrict attention to $\alpha = 1$.[6] For the correlations, we consider the case where there is no correlation among $\varepsilon^{\lambda}_{ij}$, ε^{r}_{ij}, and $\varepsilon^{\Omega}_{ij}$ as well as three extreme cases of possible pairwise correlations: $corr(\varepsilon^{r}_{ij}, \varepsilon^{\Omega}_{ij}) = 1$ and $corr(\varepsilon^{r}_{ij}, \varepsilon^{\lambda}_{ij}) = 0.84$ for each j; $corr(\varepsilon^{r}_{ij}, \varepsilon^{\Omega}_{ij}) = -1$ and $corr(\varepsilon^{r}_{ij}, \varepsilon^{\lambda}_{ij}) = 0.84$ for each j; and $corr(\varepsilon^{r}_{ij}, \varepsilon^{\Omega}_{ij}) = -1$ and $corr(\varepsilon^{\Omega}_{ij}, \varepsilon^{\lambda}_{ij}) = 0.84$ for each j.

Estimating our Model using the Simulated Data We proceed exactly as in the previous exercise, except that we estimate only Model 1a because the DGP assumes the log form for $\Omega_{ij}(\mu)$. Figure A.4 depicts (i) the mean $\Omega(\mu)$ used to generate the simulated data (along with the 2.5th and 97.5th percentiles) and (ii) the mean estimated $\Omega(\mu)$ using the simulated data (along with the 2.5th and 97.5th percentiles) for the cases (A) where $\varepsilon^{\lambda}_{ij}$, ε^{r}_{ij}, and $\varepsilon^{\Omega}_{ij}$ are uncorrelated, (B) where $\varepsilon^{\lambda}_{ij}$, ε^{r}_{ij}, and $\varepsilon^{\Omega}_{ij}$ are strongly positively correlated, (C) where $\varepsilon^{\lambda}_{ij}$ and ε^{r}_{ij} are strongly positively correlated and ε^{r}_{ij} and $\varepsilon^{\Omega}_{ij}$ are perfectly negatively correlated, and (D) where $\varepsilon^{\lambda}_{ij}$ and $\varepsilon^{\Omega}_{ij}$ are strongly positively correlated and ε^{r}_{ij} and $\varepsilon^{\Omega}_{ij}$ are perfectly negatively correlated. In each case, disregarding unobserved heterogeneity in r and $\Omega(\mu)$ either leads to essentially no bias or leads to a bias *against* finding probability distortions. If anything, therefore, this exercise indicates that disregarding unobserved heterogeneity in both standard risk aversion and probability distortions might lead us to *underestimate* the magnitude of probability distortions.

FIGURE A.4

D.3 Correlated Unobserved Heterogeneity in Risk

The exercises reported in Sections D.1 and D.2 assume that unobserved heterogeneity in risk is not correlated across coverages (i.e., $\varepsilon^{\lambda}_{iH}$, $\varepsilon^{\lambda}_{iL}$, and $\varepsilon^{\lambda}_{iM}$ are uncorrelated). In fact, the results of these exercises are robust to various correlation structures. For instance, whether we assume $\varepsilon^{\lambda}_{iH}$, $\varepsilon^{\lambda}_{iL}$, and $\varepsilon^{\lambda}_{iM}$ are perfectly correlated or have pairwise correlations of -0.33 (the strongest possible negative pairwise correlations given three pairs), our conclusions in Sections D.1 and D.2 do not change in any noticeable way. It is worth noting that, in some instances, allowing for correlation among $\varepsilon^{\lambda}_{iH}$, $\varepsilon^{\lambda}_{iL}$, and $\varepsilon^{\lambda}_{iM}$ causes unobserved heterogeneity in preferences to be correlated across coverages. In particular, we considered: (1) the case in which standard risk aversion is the same across coverages and unobserved heterogeneity in

[6] For larger values of α (e.g., $\alpha = 3.15$, which we consider in Section D.1 above), this alternative model is not well behaved, because too many households are assigned an $\Omega_{ij}(\mu) \approx 0$.

risk is perfectly correlated across coverages and perfectly positively (negatively) correlated with ε_i^r; and (2) the case in which both standard risk aversion and probability distortions are the same across coverages and unobserved heterogeneity in risk is perfectly correlated across coverages, perfectly positively (negatively) correlated with ε_i^r, and perfectly positively (negatively) correlated with ε_i^Ω. Finally, we note that if we assume the same correlation structures for unobserved heterogeneity in risk, but without permitting unobserved heterogeneity in preferences (i.e., setting $\alpha = 0$), we again find essentially no bias in favor of finding probability distortions.

E Appendix Tables

On the ensuing pages, we report Tables A.1 through A.20.

Table A.1: Summary of Premium Menus - Auto Collision

Core Sample (4170 Households)

	Deductible Choice				
	$100	**$200**	**$250**	**$500**	**$1000**
Mean annual premium for coverage with $500 deductible	110	129	146	189	255
Standard deviation	54	54	66	96	168
Mean cost of decreasing deductible from $500 to $250	33	38	44	57	77
Standard deviation	17	17	20	29	52
Mean savings from increasing deductible from $500 to $1000	24	29	33	43	58
Standard deviation	12	12	15	22	39
Number of households	42	559	467	2822	280

Note: All values in dollars, except number of households.

Table A.2: Summary of Premium Menus - Auto Comprehensive

Core Sample (4170 Households)

	Deductible Choice					
	$50	**$100**	**$200**	**$250**	**$500**	**$1000**
Mean annual premium for coverage with $500 deductible	61	70	92	98	136	258
Standard deviation	27	33	43	41	71	247
Mean cost of decreasing deductible from $500 to $250	16	18	24	26	36	68
Standard deviation	7	9	11	11	19	66
Mean savings from increasing deductible from $500 to $1000	12	14	18	19	27	51
Standard deviation	5	7	9	8	14	49
Number of households	216	171	1397	440	1795	151

Note: All values in dollars, except number of households.

Table A.3: Summary of Premium Menus - Home

Core Sample (4170 Households)

	Deductible Choice					
	$100	**$250**	**$500**	**$1000**	**$2500**	**$5000**
Mean annual premium for coverage with $500 deductible	366	520	631	972	2218	3366
Standard deviation	113	218	308	593	2289	1808
Mean cost of decreasing deductible from $500 to $250	31	42	52	80	183	275
Standard deviation	6	18	26	48	201	140
Mean savings from increasing deductible from $500 to $1000	41	57	69	107	244	368
Standard deviation	8	23	34	64	268	188
Number of households	36	1239	2166	664	50	15

Note: All values in dollars, except number of households.

Table A.4: Claim Rate Regressions - Auto

Poisson Panel Regression Model with Random Effects

Full Data Set (1,348,020 Household-Year Records)

	Collision		Comprehensive	
	Coef	**Std Err**	**Coef**	**Std Err**
Constant	-6.7646 **	0.0616	-7.9277 **	0.1057
Driver 2 Indicator	-0.0485	0.0593	-0.3542 **	0.1022
Driver 3+ Indicator	0.3215 **	0.0733	-0.1261	0.1201
Vehicle 2 Indicator	0.5991 **	0.0466	0.6502 **	0.0782
Vehicle 3+ Indicator	0.7312 **	0.0596	0.8766 **	0.0937
Young Driver	-0.0058	0.0296	0.0895 **	0.0453
Driver 1 Age	-0.0210 **	0.0015	0.0113 **	0.0029
Driver 1 Age Squared	0.0002 **	0.0000	-0.0002 **	0.0000
Driver 1 Female	0.1040 **	0.0093	-0.0672 **	0.0168
Driver 1 Married	0.0630 **	0.0111	0.0640 **	0.0201
Driver 1 Divorced	0.0186	0.0141	0.0914 **	0.0247
Driver 1 Separated	0.0392	0.0256	0.0791	0.0428
Driver 1 Single	.	.	.	.
Driver 1 Widowed	0.0031	0.0160	-0.0170	0.0335
Vehicle 1 Age	-0.0354 **	0.0019	-0.0286 **	0.0030
Vehicle 1 Age Squared	-0.0006 **	0.0001	0.0000	0.0002
Vehicle 1 Business	.	.	.	.
Vehicle 1 Farm	-0.2575 **	0.0872	0.0206	0.1194
Vehicle 1 Pleasure	-0.1094 **	0.0306	-0.1118 **	0.0526
Vehicle 1 Work	-0.0831 **	0.0304	-0.0620	0.0523
Vehicle 1 Passive Restraint	-0.1087 **	0.0239	-0.0858 **	0.0352
Vehicle 1 Anti-Theft	0.0754 **	0.0078	0.0735 **	0.0136
Vehicle 1 Anti-Lock	0.0581 **	0.0080	0.0729 **	0.0139
Driver 2 Age	0.0115 **	0.0024	0.0181 **	0.0042
Driver 2 Age Squared	-0.0001 **	0.0000	-0.0001 **	0.0000
Driver 2 Female	0.1204 **	0.0151	-0.0376	0.0257
Driver 2 Married	-0.0835 **	0.0191	-0.0408	0.0326
Driver 2 Divorced	-0.1579	0.1027	-0.1347	0.1636
Driver 2 Separated	0.0254	0.2130	0.1796	0.3226
Driver 2 Single	.	.	.	.
Driver 2 Widowed	-0.0802	0.1383	-1.1835 **	0.3864
Vehicle 2 Age	-0.0332 **	0.0016	-0.0229 **	0.0027
Vehicle 2 Age Squared	0.0004 **	0.0001	0.0002 **	0.0001
Vehicle 2 Business	.	.	.	.
Vehicle 2 Farm	-0.1703	0.1056	-0.1345	0.1500
Vehicle 2 Pleasure	-0.1805 **	0.0380	-0.0563	0.0663
Vehicle 2 Work	-0.1670 **	0.0381	0.0119	0.0664
Vehicle 2 Passive Restraint	-0.0428 **	0.0201	-0.0875 **	0.0294
Vehicle 2 Anti-Theft	0.0547 **	0.0103	0.0385 **	0.0171
Vehicle 2 Anti-Lock	0.0317 **	0.0105	0.0199	0.0170
Driver 1 Credit Score	-0.0017 **	0.0000	-0.0013 **	0.0001
Driver 1 Previous Accident	0.0913 **	0.0156	0.0756 **	0.0277
Driver 1 Previous Convictions	0.1476	0.0888	0.0648	0.1670
Driver 1 Previous Reinstated	0.0170	0.0558	0.0003	0.0996
Driver 1 Previous Revocation	-0.0218	0.1456	0.3156	0.1967
Driver 1 Previous Suspension	0.0463	0.0564	0.0125	0.1026
Driver 1 Previous Violation	0.0827 **	0.0093	0.0577 **	0.0161
Year Dummies	Yes		Yes	
Territory Codes	Yes		Yes	
Variance (ϕ)	0.2242 **	0.0065	0.5661	0.0198
Loglikelihood	-399,318		-169,817	

Note: Territory codes indicate rating territories, which are based on actuarial risk factors, such as traffic and weather patterns, population demographics, wildlife density, and the cost of goods and services.

** Significant at 5 percent level.

Table A.5: Claim Rate Regression - Home

Poisson Panel Regression Model with Random Effects

Full Data Set (1,265,229 Household-Year Records)

	Coef		Std Err
Constant	-7.3642	**	0.0978
Dwelling Value	0.0000	**	0.0000
Home Age	0.0016	**	0.0006
Home Age Squared	0.0000	**	0.0000
Number of Families	-0.0021		0.0023
Distance to Hydrant	0.0000		0.0000
Alarm Discount	0.2463	**	0.0195
Protection Devices	-0.1852	**	0.0239
Farm/Business	0.1044	**	0.0242
Primary Home	0.4832	**	0.0819
Owner Occupied	0.2674	**	0.0419
Construction: Fire Resistant	0.1525		0.1342
Construction: Masonry	0.0751	**	0.0172
Construction: Masonry/Veneer	0.0755	**	0.0252
Construction: Frame	.		.
Credit Score	-0.0026	**	0.0000
Year Dummies		Yes	
Protection Classes		Yes	
Territory Codes		Yes	
Variance (ϕ)	0.4514	**	0.0086
Loglikelihood		-347,278	

Note: Territory codes indicate rating territories, which are based on actuarial risk factors, such as traffic and weather patterns, population demographics, wildlife density, and the cost of goods and services. Protection classes guage the effectiveness of local fire protection and building codes.

** Significant at 5 percent level.

Table A.6: Model 3 Estimates

Core Sample (4170 Households)

	r		**Log Ω(μ): constant**		**Log Ω(μ): linear**		**Log Ω(μ): quadratic**	
	Coef	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**
Constant	-6.88 **	0.08	-2.83 **	0.03	12.55 **	0.50	-37.61 **	2.98
Driver 1 age	-0.81 **	0.23	0.06	0.11	3.16	2.05	-6.05	11.97
Driver 1 age squared	0.49 **	0.25	0.12	0.11	-5.63 **	2.21	17.59	12.97
Driver 1 female	0.06	0.04	-0.03	0.02	1.33 **	0.53	-10.24 **	3.03
Driver 1 single	-0.04	0.05	0.02	0.03	0.45	0.58	-3.71	3.00
Driver 1 married	0.02	0.07	-0.01	0.04	1.10	1.01	-7.61	5.01
Driver 1 credit score	-0.07	0.04	-0.02	0.02	2.16 **	0.48	-10.75 **	2.59
Driver 2 indicator	-0.01	0.06	0.01	0.03	-1.55 **	0.71	7.93 **	3.77
Home value	-0.70 **	0.05	0.03 **	0.01	0.40 *	0.23	0.13	0.75
Parameter mean	0.00126		-2.83		12.55		-37.61	
Parameter median	0.00110		-2.85		12.66		-37.29	
	Collision		**Comprehensive**		**Home**		**Loglikelihood value**	
σ	31.83 **	1.09	20.19 **	0.62	70.52 **	2.79	-14819.51	

Note: Each variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at 5 percent level.

* Significant at 10 percent level.

Table A.7: Model 4 Estimates

Core Sample (4170 Households)

	r		**Log Ω(μ): constant**		**Log Ω(μ): linear**		**Log Ω(μ): quadratic**	
	Coef	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**
Constant	3.70 **	0.14	-2.65 **	0.02	12.22 **	0.54	-36.50 **	3.47
Driver 1 age	-0.62	0.40	0.23 **	0.09	2.17	1.40	-1.23	7.19
Driver 1 age squared	-0.52	0.48	-0.04	0.09	-4.08 **	1.45	9.79	7.79
Driver 1 female	0.03	0.06	-0.03	0.02	1.38 **	0.50	-10.88 **	2.81
Driver 1 single	-0.11	0.07	0.01	0.02	0.56	0.47	-4.46 *	2.66
Driver 1 married	-0.03	0.10	0.00	0.03	0.90 *	0.52	-6.88 **	3.05
Driver 1 credit score	-0.08	0.06	-0.03	0.02	2.07 **	0.48	-10.21 **	2.70
Driver 2 indicator	0.10	0.09	-0.01	0.02	-1.31 **	0.54	6.14 *	3.22
Home value	0.20 **	0.03	-0.01	0.01	0.67 **	0.13	-0.34	0.34
Parameter mean	0.00050		-2.65		12.22		-36.50	
Parameter median	0.00028		-2.65		12.24		-35.90	
	Collision		**Comprehensive**		**Home**		**Loglikelihood value**	
σ	25.23 **	0.63	16.94 **	0.50	53.76 **	2.05	-14985.31	

Note: Each variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at 5 percent level.

* Significant at 10 percent level.

Table A.8: Model 5 Estimates

Core Sample (4170 Households)

	r		**Log Ω(μ): constant**		**Log Ω(μ): linear**		**Log Ω(μ): quadratic**	
	Coef	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**
Constant	2.62 **	0.05	-2.75 **	0.02	12.00 **	0.57	-35.52 **	3.43
Driver 1 age	-1.17 **	0.32	0.31 **	0.12	6.84 **	1.06	-14.43 **	2.43
Driver 1 age squared	0.69	0.37	-0.10	0.11	-8.71 **	1.10	23.78 **	2.49
Driver 1 female	-0.01	0.04	0.00	0.02	0.80 **	0.20	-8.07 **	1.96
Driver 1 single	-0.10 **	0.04	0.02	0.02	0.67 **	0.32	-4.00 **	2.00
Driver 1 married	-0.01	0.05	0.01	0.02	0.84 **	0.29	-6.07 **	2.24
Driver 1 credit score	-0.10 **	0.03	-0.02	0.02	2.33 **	0.35	-10.71 **	1.46
Driver 2 indicator	0.01	0.05	0.01	0.02	-1.72 **	0.23	8.54 **	1.54
Home value	0.45 **	0.01	-0.05 **	0.01	0.95 **	0.07	-0.71 **	0.15
Parameter mean	0.00056		-2.75		12.00		-35.52	
Parameter median	0.00044		-2.75		12.01		-35.42	
	Collision		**Comprehensive**		**Home**		**Loglikelihood value**	
σ	0.45 **	0.02	0.30 **	0.02	1.08 **	0.07	-14942.46	

Note: Each variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at 5 percent level.

Table A.9: Model 2u Estimates

Core Sample (4170 Households)

	r		**Log Ω(μ): constant**		**Log Ω(μ): linear**		**Log Ω(μ): quadratic**	
	Coef	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**
Constant	-7.11 **	0.08	-2.73 **	0.02	11.04 **	0.41	-21.78 **	1.95
Driver 1 age	-1.22 **	0.19	0.15 **	0.07	3.31 **	0.57	-3.86 **	1.66
Driver 1 age squared	0.75 **	0.20	0.02	0.06	-4.92 **	0.61	7.64 **	1.10
Driver 1 female	0.05	0.05	-0.01	0.02	1.24 **	0.29	-9.90 **	1.32
Driver 1 single	0.03	0.07	0.00	0.02	0.56 **	0.21	-3.25 **	1.20
Driver 1 married	0.02	0.12	-0.01	0.05	1.24 **	0.46	-6.87 **	1.00
Driver 1 credit score	-0.12 **	0.05	-0.01	0.03	1.49 **	0.62	-3.64	2.73
Driver 2 indicator	-0.09	0.08	0.01	0.04	-1.40 *	0.75	5.27 **	1.99
Parameter mean	0.00097		-2.73		11.04		-21.78	
Parameter median	0.00076		-2.74		11.05		-19.49	
	Collision		**Comprehensive**		**Home**		**Loglikelihood value**	
σ	25.93 **	0.80	18.02 **	0.56	66.01 **	2.78	-14952.75	

Note: Each variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at 5 percent level.

* Significant at 10 percent level.

Table A.10: Model 4u Estimates

Core Sample (4170 Households)

	r		**Log Ω(μ): constant**		**Log Ω(μ): linear**		**Log Ω(μ): quadratic**	
	Coef	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**
Constant	4.33 **	0.12	-2.69 **	0.02	11.68 **	0.40	-26.98 **	1.88
Driver 1 age	0.02	0.21	0.19 **	0.06	0.12	0.37	8.20 **	2.39
Driver 1 age squared	-1.04 **	0.26	0.05	0.06	-3.14 **	0.43	5.86 **	2.45
Driver 1 female	-0.01	0.05	0.03 *	0.02	-0.66	0.44	2.35	1.89
Driver 1 single	-0.12 **	0.06	0.02	0.03	0.52	0.78	-3.74	3.70
Driver 1 married	0.02	0.08	0.00	0.02	0.77 **	0.20	-5.87 **	1.15
Driver 1 credit score	-0.08	0.05	-0.02	0.02	1.97 **	0.42	-7.92 **	1.87
Driver 2 indicator	-0.04	0.08	0.03	0.03	-1.98 **	0.57	8.82 **	2.70
Home value	0.27 **	0.04	-0.04 **	0.01	1.85 **	0.13	-1.92 **	0.18
Parameter mean	0.00081		-2.69		11.68		-26.98	
Parameter median	0.00055		-2.71		11.65		-27.08	
	Collision		**Comprehensive**		**Home**		**Loglikelihood value**	
σ	24.66 **	0.68	17.34 **	0.53	53.30 **	2.29	-14976.45	

Note: Each variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at 5 percent level.

* Significant at 10 percent level.

Table A.11: Model 5u Estimates

Core Sample (4170 Households)

	r		**Log Ω(μ): constant**		**Log Ω(μ): linear**		**Log Ω(μ): quadratic**	
	Coef	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**
Constant	3.16 **	0.04	-2.96 **	0.02	13.06 **	0.40	-33.78 **	1.71
Driver 1 age	-0.60 **	0.11	0.50 **	0.12	6.69 **	1.47	-19.57 **	6.07
Driver 1 age squared	0.09	0.14	-0.17	0.11	-8.80 **	1.57	28.18 **	6.73
Driver 1 female	-0.01	0.02	0.08 **	0.02	-1.46 **	0.33	4.73 **	1.16
Driver 1 single	-0.09 **	0.02	0.03	0.02	0.69 **	0.32	-3.90 **	0.77
Driver 1 married	0.00	0.03	0.07 **	0.03	-1.58 **	0.52	7.56 **	1.78
Driver 1 credit score	-0.07 **	0.02	0.00	0.02	1.75 **	0.29	-4.48 **	0.95
Driver 2 indicator	-0.02	0.02	0.02	0.03	-1.25 **	0.51	3.95 **	1.83
Home value	0.68 **	0.01	-0.14 **	0.02	1.78 **	0.21	1.75 **	0.64
Parameter mean	0.00093		-2.96		13.06		-33.78	
Parameter median	0.00082		-2.96		12.81		-31.62	
	Collision		**Comprehensive**		**Home**		**Loglikelihood value**	
σ	0.87 **	0.04	0.60 **	0.02	1.67 **	0.09	-14718.29	

Note: Each variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at 5 percent level.

Table A.12: Model 2 with CARA Utility

Core Sample (4170 Households)

	r		**Log Ω(μ): constant**		**Log Ω(μ): linear**		**Log Ω(μ): quadratic**	
	Coef	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**
Constant	-6.83 **	0.05	-3.01 **	0.03	11.94 **	0.65	-34.82 **	3.94
Driver 1 age	-0.75 **	0.34	0.33 **	0.15	8.43 **	1.33	-18.69 **	2.73
Driver 1 age squared	0.53	0.39	-0.08	0.16	-10.42 **	1.49	28.08 **	3.37
Driver 1 female	0.04	0.02	-0.02	0.03	0.90	0.75	-9.00 **	4.25
Driver 1 single	0.02	0.02	-0.01	0.02	0.34	0.23	-3.13 *	1.90
Driver 1 married	-0.02	0.03	0.02	0.04	0.67	0.89	-4.54 *	2.74
Driver 1 credit score	-0.04	0.03	-0.02	0.03	2.30 **	0.78	-10.59 **	3.83
Driver 2 indicator	-0.05	0.03	0.01	0.03	-0.95 **	0.46	6.21 **	1.57
Parameter mean	0.00113		-3.01		11.94		-34.82	
Parameter median	0.00103		-3.01		12.06		-34.67	
	Collision		**Comprehensive**		**Home**		**Loglikelihood value**	
σ	33.31 **	0.98	21.30 **	0.64	104.30 **	5.29	-14713.74	

Note: Each variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at 5 percent level.

* Significant at 10 percent level.

Table A.13: Model 2 with Alternative Sample 1

Alternative Sample 1 (20,662 Households)

	r		**Log Ω(μ): constant**		**Log Ω(μ): linear**		**Log Ω(μ): quadratic**	
	Coef	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**
Constant	-6.83 **	0.04	-2.88 **	0.01	14.01 **	0.34	-43.10 **	2.78
Driver 1 age	1.14 **	0.34	-0.31 **	0.09	-1.14 **	0.54	-7.11 **	2.75
Driver 1 age squared	-1.28 **	0.36	0.41 **	0.08	-2.79 **	0.34	26.03 **	1.11
Driver 1 female	0.16 **	0.02	-0.01	0.01	-0.89 **	0.22	2.40 **	0.86
Driver 1 single	-0.08 **	0.04	0.03 **	0.02	0.14	0.21	-1.65	1.27
Driver 1 married	0.01	0.03	0.03 **	0.01	-0.97 **	0.14	2.68 **	0.32
Driver 1 credit score	-0.07 **	0.02	0.03 **	0.01	0.86 **	0.25	-4.51 **	1.21
Driver 2 indicator	0.10 **	0.03	0.01	0.01	-2.23 **	0.30	10.63 **	2.04
Parameter mean	0.00113		-2.88		14.01		-43.10	
Parameter median	0.00112		-2.92		14.26		-45.60	
	Collision		**Comprehensive**		**Home**		**Loglikelihood value**	
σ	31.17 **	0.48	17.91 **	0.43	18.93 **	0.27	-49570.46	

Note: Each variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at 5 percent level.

Table A.14: Model 2 with Alternative Sample 2

Alternative Sample 2 (6824 Households)

	r		**Log Ω(μ): constant**		**Log Ω(μ): linear**		**Log Ω(μ): quadratic**	
	Coef	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**
Constant	-7.57 **	0.08	-2.63 **	0.02	12.91 **	0.43	-42.12 **	2.97
Driver 1 age	-1.31 **	0.21	0.06	0.09	4.52 *	2.46	-11.77	15.32
Driver 1 age squared	0.91 **	0.22	0.08	0.09	-6.73 **	2.53	22.96	16.35
Driver 1 female	0.15 **	0.04	-0.03	0.02	0.94 **	0.29	-7.90 **	1.65
Driver 1 single	0.05	0.05	0.02	0.02	-0.10	0.33	-0.18	2.01
Driver 1 married	-0.13	0.08	0.04 **	0.02	0.42	0.45	-4.30 *	2.58
Driver 1 credit score	-0.08 **	0.04	-0.01	0.02	2.19 **	0.38	-12.41 **	2.31
Driver 2 indicator	0.13	0.07	-0.03	0.02	-1.32 **	0.41	8.32 **	2.86
Parameter mean	0.00060		-2.63		12.91		-42.12	
Parameter median	0.00048		-2.65		12.99		-42.40	
	Collision		**Comprehensive**		**Home**		**Loglikelihood value**	
σ	28.61 **	0.71	18.78 **	0.47	58.74 **	2.00	-24690.75	

Note: Each variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at 5 percent level.

* Significant at 10 percent level.

Table A.15: Model 2 with Restricted Menu I

Core Sample (4170 Households)

	r		**Log Ω(μ): constant**		**Log Ω(μ): linear**		**Log Ω(μ): quadratic**	
	Coef	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**
Constant	-8.79 **	0.27	-2.81 **	0.04	12.88 **	0.65	-30.29 **	3.12
Driver 1 age	-4.22 **	0.85	0.03	0.17	6.47 **	2.37	-8.09	8.02
Driver 1 age squared	3.92 **	0.94	0.38 **	0.17	-10.94 **	2.28	23.24 **	7.44
Driver 1 female	0.14	0.16	0.03	0.04	0.91	1.00	-9.94 **	5.00
Driver 1 single	0.18	0.19	-0.01	0.04	0.96	0.99	-6.98	5.07
Driver 1 married	0.20	0.27	-0.02	0.04	1.31 *	0.76	-8.92 **	3.09
Driver 1 credit score	-0.54 **	0.17	-0.04	0.03	2.59 **	0.66	-10.26 **	2.93
Driver 2 indicator	-0.07	0.21	0.00	0.04	-1.19	0.79	5.52	4.05
Parameter mean	0.00029		-2.81		12.88		-30.29	
Parameter median	0.00013		-2.88		13.46		-30.35	
	Collision		**Comprehensive**		**Home**		**Loglikelihood value**	
σ	26.26 **	0.84	19.83 **	0.89	76.56 **	4.49	-11517.11	

Note: Each variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at 5 percent level.

* Significant at 10 percent level.

Table A.16: Model 2 with Restricted Menu II

Core Sample (4170 Households)

	r		**Log Ω(μ): constant**		**Log Ω(μ): linear**		**Log Ω(μ): quadratic**	
	Coef	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**
Constant	-7.52 **	0.11	-2.34 **	0.03	8.43 **	0.55	-23.31 **	3.26
Driver 1 age	-1.85 **	0.44	0.29 **	0.13	1.97	2.39	0.66	3.08
Driver 1 age squared	1.49 **	0.52	-0.06	0.11	-4.39 **	1.57	9.50 **	4.03
Driver 1 female	0.17 **	0.06	-0.05 *	0.02	1.63 **	0.63	-12.81 **	4.15
Driver 1 single	0.17	0.09	-0.04	0.02	0.68 **	0.32	-5.30 **	1.39
Driver 1 married	0.16	0.12	-0.03	0.03	1.16 **	0.56	-8.66 **	2.92
Driver 1 credit score	-0.17 **	0.07	0.01	0.03	1.51 **	0.69	-7.41 *	3.92
Driver 2 indicator	-0.02	0.10	-0.01	0.04	-0.91	0.78	4.67	4.19
Parameter mean	0.00066		-2.34		8.43		-23.31	
Parameter median	0.00048		-2.34		8.48		-21.94	
	Collision		**Comprehensive**		**Home**		**Loglikelihood value**	
σ	33.21 **	1.28	19.70 **	0.74	73.79 **	4.37	-10299.92	

Note: Each variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at 5 percent level.

* Significant at 10 percent level.

Table A.17: Model 2 with Alternative Error Structure A and Restricted Menu I

Core Sample (4170 Households)

	r		Log Ω(μ): constant		Log Ω(μ): linear		Log Ω(μ): quadratic	
	Coef	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**
Constant	-9.22 **	0.20	-2.81 **	0.05	13.48 **	0.80	-32.52 **	3.46
Driver 1 age	-4.89 **	0.97	0.04	0.20	6.23 **	1.69	-6.01	5.40
Driver 1 age squared	4.61 **	1.22	0.35	0.24	-10.31 **	2.68	19.36 **	2.34
Driver 1 female	0.09	0.10	0.03	0.02	0.91 **	0.27	-9.67 **	2.15
Driver 1 single	0.30 **	0.13	-0.01	0.02	0.84 **	0.37	-5.96 **	1.87
Driver 1 married	0.44 **	0.13	-0.02	0.03	1.22 **	0.32	-8.34 **	1.26
Driver 1 credit score	-0.61 **	0.13	-0.04	0.03	2.57 **	0.76	-10.16 **	3.30
Driver 2 indicator	-0.07	0.18	0.00	0.03	-1.18 **	0.47	5.43 **	2.14
Parameter mean	0.00022		-2.81		13.48		-32.52	
Parameter median	0.00008		-2.87		13.96		-32.32	
	Collision		**Comprehensive**		**Home**		**Loglikelihood value**	
σ	51.63 **	1.34	36.49 **	1.46	130.57 **	6.96	-11571.67	

Note: Each variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at 5 percent level.

Table A.18: Model 2 with Alternative Error Structure A and Restricted Menu II

Core Sample (4170 Households)

	r		Log Ω(μ): constant		Log Ω(μ): linear		Log Ω(μ): quadratic	
	Coef	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**	**Coef**	**Std Err**
Constant	-9.28 **	0.37	-2.23 **	0.02	8.13 **	0.41	-22.78 **	1.88
Driver 1 age	0.41	0.50	0.34 **	0.11	2.76	1.79	-5.28	8.67
Driver 1 age squared	-2.42 **	0.80	-0.06	0.11	-5.23 **	1.76	16.15 *	8.63
Driver 1 female	0.14 **	0.06	-0.03	0.02	1.68 **	0.50	-13.36 **	2.83
Driver 1 single	0.10	0.08	-0.02	0.02	0.62	0.41	-4.63 *	2.45
Driver 1 married	0.09	0.12	-0.02	0.03	1.04 **	0.42	-7.85 **	2.53
Driver 1 credit score	-0.01	0.06	-0.02	0.02	1.27 **	0.38	-6.43 **	1.89
Driver 2 indicator	-0.04	0.10	-0.02	0.03	-0.53	0.50	2.62	2.71
Parameter mean	0.00035		-2.23		8.13		-22.78	
Parameter median	0.00015		-2.23		8.27		-21.79	
	Collision		**Comprehensive**		**Home**		**Loglikelihood value**	
σ	60.77 **	2.12	37.66 **	1.42	114.41 **	5.55	-10337.61	

Note: Each variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at 5 percent level.

* Significant at 10 percent level.

Table A.19: Model 2 with Alternative Error Structure B and Restricted Menu I

Core Sample (4170 Households)

	r			**Log Ω(μ): constant**			**Log Ω(μ): linear**			**Log Ω(μ): quadratic**		
	Coef		**Std Err**	**Coef**		**Std Err**	**Coef**		**Std Err**	**Coef**		**Std Err**
Constant	-7.04	**	0.09	-2.04	**	0.03	3.21	**	0.64	-7.30	**	3.34
Driver 1 age	-1.43	**	0.18	0.42	**	0.11	2.89		1.90	-7.31		5.79
Driver 1 age squared	1.16	**	0.19	-0.28	**	0.12	-3.60		2.27	10.37		8.06
Driver 1 female	0.00		0.03	0.03		0.02	0.08		0.50	-2.96		1.98
Driver 1 single	-0.05		0.04	-0.09	**	0.04	2.03	*	1.11	-7.82		5.07
Driver 1 married	-0.11	**	0.06	0.01		0.05	1.32		1.01	-7.73	**	3.23
Driver 1 credit score	-0.26	**	0.03	-0.01		0.01	2.25	**	0.29	-8.80	**	1.26
Driver 2 indicator	-0.07		0.05	-0.02		0.02	0.14		0.22	2.25		2.11
Parameter mean	0.00101			-2.04			3.21			-7.30		
Parameter median	0.00081			-2.01			3.36			-7.54		
	Collision			**Comprehensive**			**Home**			**Loglikelihood value**		
σ	62.75	**	2.22	53.33	**	3.23	65.80	**	2.94	-11702.47		

Note: Each variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at 5 percent level.

* Significant at 10 percent level.

Table A.20: Model 2 with Alternative Error Structure B and Restricted Menu II

Core Sample (4170 Households)

	r			**Log Ω(μ): constant**			**Log Ω(μ): linear**			**Log Ω(μ): quadratic**		
	Coef		**Std Err**	**Coef**		**Std Err**	**Coef**		**Std Err**	**Coef**		**Std Err**
Constant	-7.06	**	0.07	-2.26	**	0.03	7.22	**	0.55	-21.34	**	2.58
Driver 1 age	-1.07	**	0.37	0.15	**	0.07	3.29	**	1.33	-6.44		13.08
Driver 1 age squared	0.83	**	0.37	0.04		0.06	-5.43	**	1.45	14.89		14.72
Driver 1 female	0.09	**	0.03	-0.03		0.02	0.85	**	0.28	-6.66	**	1.92
Driver 1 single	-0.01		0.06	0.00		0.02	0.51		0.62	-2.99		4.11
Driver 1 married	-0.04		0.06	0.00		0.05	0.89		1.01	-5.66		3.89
Driver 1 credit score	-0.15	**	0.05	0.02		0.02	1.62	**	0.24	-8.35	**	1.30
Driver 2 indicator	-0.05		0.05	0.00		0.03	-0.52	**	0.25	4.09		3.24
Parameter mean	0.00093			-2.26			7.22			-21.34		
Parameter median	0.00082			-2.27			7.39			-21.42		
	Collision			**Comprehensive**			**Home**			**Loglikelihood value**		
σ	53.09	**	1.58	38.37	**	1.24	66.17	**	2.65	-10996.25		

Note: Each variable *z* is normalized as (*z* -mean(*z*))/stdev(*z*).

** Significant at 5 percent level.

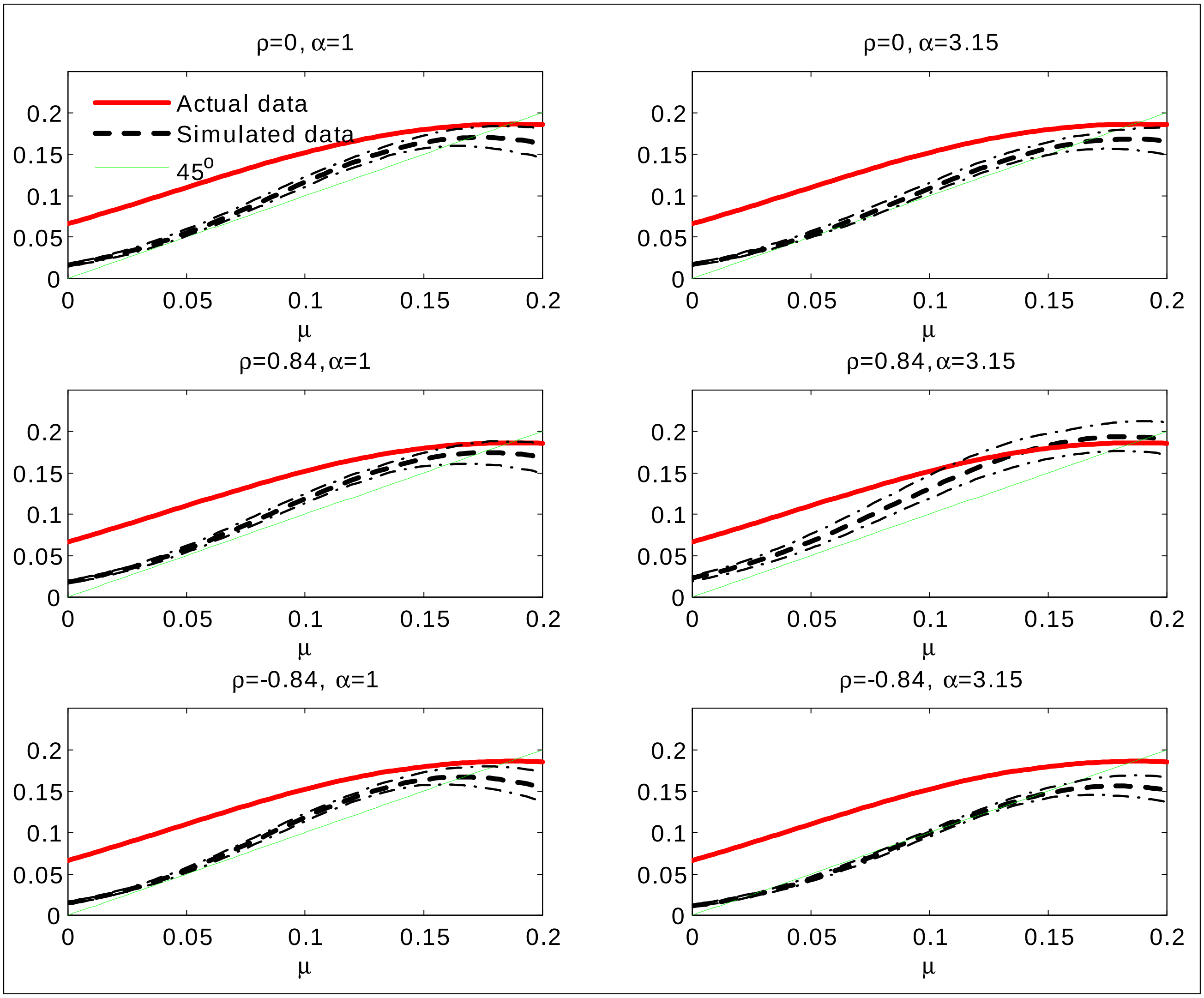

Figure A.1: Mean Estimated $\Omega(\mu)$ – Unobserved Heterogeneity in r – Model 1a

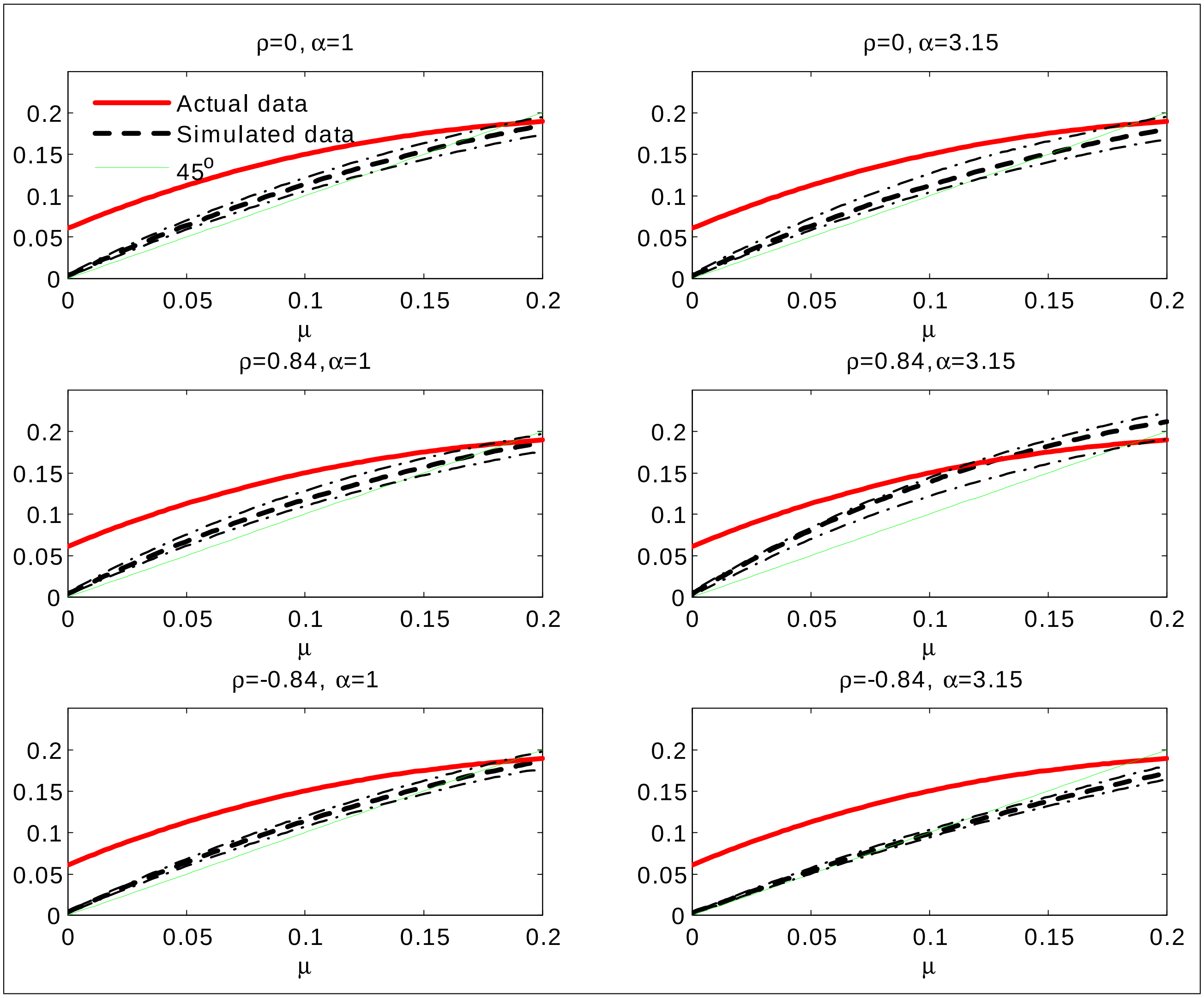

Figure A.2: Mean Estimated $\Omega(\mu)$ – Unobserved Heterogeneity in r – Model 1b

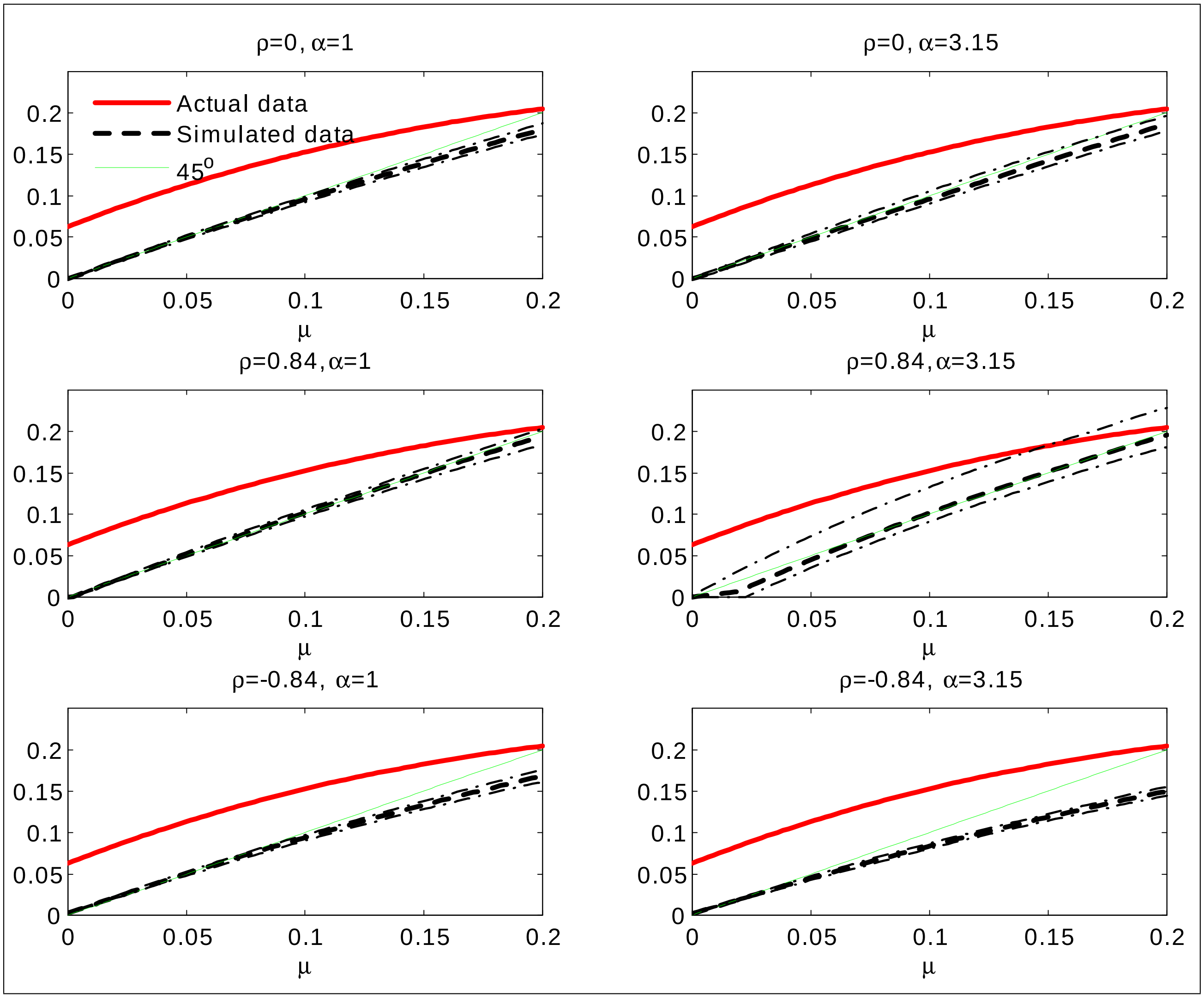

Figure A.3: Mean Estimated $\Omega(\mu)$ – Unobserved Heterogeneity in r – Model 1bu

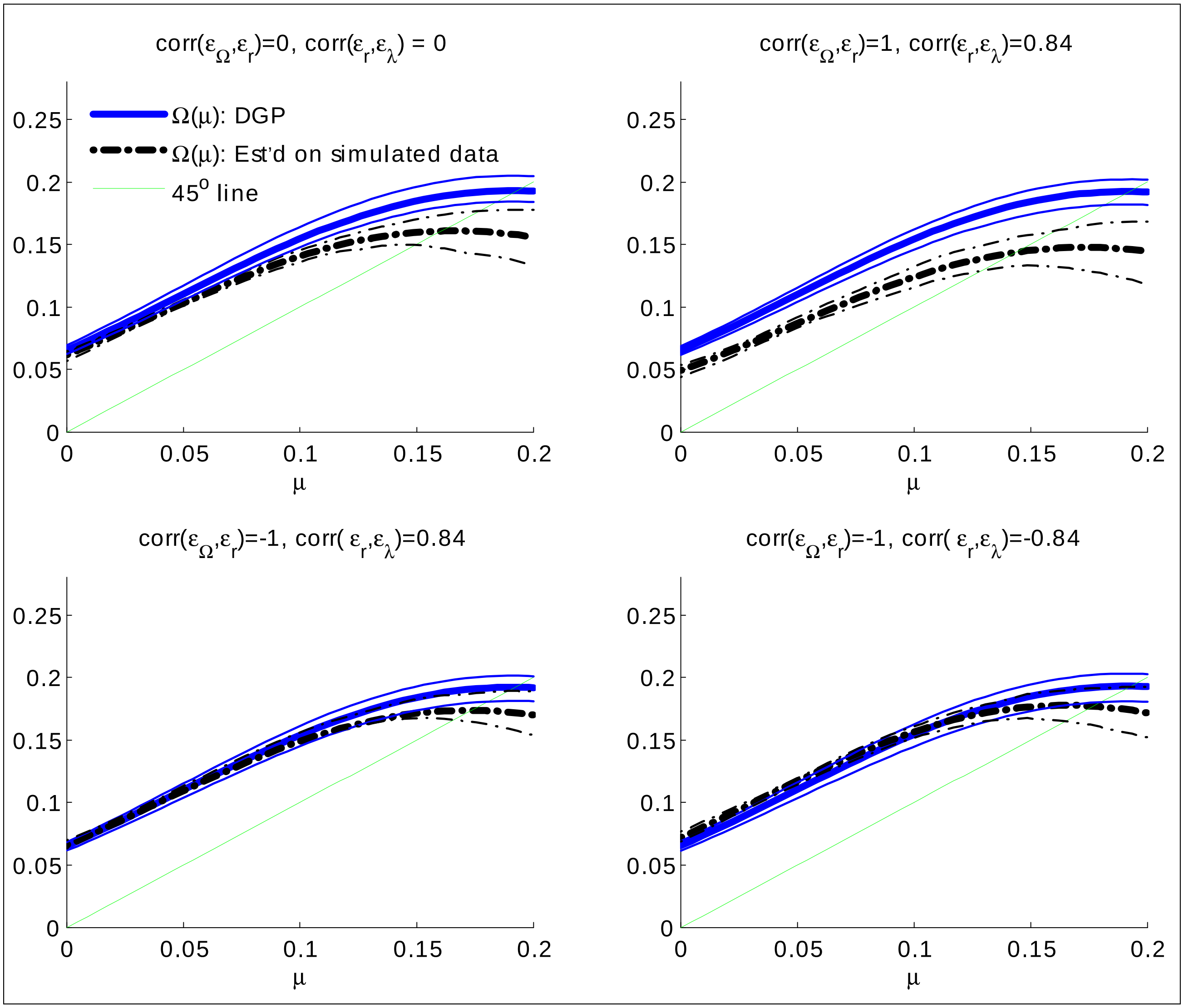

Figure A.4: Unobserved Heterogeneity in r and $\Omega(\mu)$ – Model 1a