Inference under stability of risk preferences

Levon Barseghyan
Department of Economics, Cornell University

Francesca Molinari
Department of Economics, Cornell University

Joshua C. Teitelbaum
Law Center, Georgetown University

We leverage the assumption that preferences are stable across contexts to partially identify and conduct inference on the parameters of a structural model of risky choice. Working with data on households' deductible choices across three lines of insurance coverage and a model that nests expected utility theory plus a range of non-expected utility models, we perform a revealed preference analysis that yields household-specific bounds on the model parameters. We then impose stability and other structural assumptions to tighten the bounds, and we explore what we can learn about households' risk preferences from the intervals defined by the bounds. We further utilize the intervals to (i) classify households into preference types and (ii) recover the single parameterization of the model that best fits the data. Our approach does not entail making distributional assumptions about unobserved heterogeneity in preferences.

Keywords. Inference, insurance, partial identification, revealed preference, risk preferences, stability.

JEL classification. D01, D12, D81, G22.

1. Introduction

Economists strive to develop models of decision making that can explain choices across multiple domains. At a minimum, we ask that a model's explanatory power extend across decision contexts that are essentially similar. Stated more formally, we require that a model satisfy a criterion of *stability*: a single agent-specific parameterization of the model should be consistent with the agent's choices in closely related domains.

In this paper, we demonstrate how one can exploit the stability criterion to conduct inference on the agent-specific parameters of a structural model of decision making under risk. We develop an approach that relies principally on the stability criterion and revealed preference arguments to bound the model parameters. Working with data on households' deductible choices across three lines of insurance coverage and a model that nests expected utility theory plus a broad range of non-expected utility models, we first show how one can infer household-specific bounds on the model parameters from a household's choices and then leverage the stability criterion and other structural assumptions to sharpen the inference—that is, tighten the bounds. We then show how one can utilize the intervals defined by the bounds to (i) classify households into preference types and (ii) recover the single parameterization of the model that best fits the data. Importantly, our approach does not entail making arbitrary assumptions about the distribution of unobserved heterogeneity in preferences. Rather, in line with the partial identification paradigm (e.g., Manski (2003)), it explores what we can learn about the structure of risk preferences without distributional assumptions that are motivated by statistical convenience. It thus yields more credible inferences than standard approaches to identification and estimation that rely on such assumptions (e.g., parametric maximum likelihood estimation (MLE) of a random utility model).

In Section 2, we describe our data. The data hail from a U.S. property and casualty insurance company that specializes in personal auto and home coverage. The full data set comprises annual information on a large sample of households who purchased auto or home policies from the company between 1998 and 2006. For reasons we explain, we restrict attention to a sample of 4170 households who purchased both auto and home policies in the same year, in either 2005 or 2006. For each household, we observe its deductible choices in three lines of coverage: auto collision, auto comprehensive, and home all perils. We also observe the coverage-specific pricing menus of premium–deductible combinations that each household faced when choosing its deductibles. In addition, we observe the households' claims histories and an array of demographics for each household. We utilize the data on claims and demographics to assign each household a claim probability in each line of coverage.

In Section 3, we outline the model. The model is a generalization of objective expected utility theory that allows for generic probability distortions through an unspecified function, $\Omega(\cdot)$. The probability distortions in the model are generic in the sense that they can capture in a reduced form way a wide range of different behaviors, including subjective beliefs, rank-dependent probability weighting (Quiggin (1982)), Kahneman–Tversky (KT) probability weighting (Kahneman and Tversky (1979)), Gul disappointment aversion (Gul (1991)), and Kőszegi–Rabin (KR) loss aversion (Kőszegi and Rabin (2006, 2007)).[1] Consequently, the model offers a parsimonious representation of a number of different classes of risk preferences.

In Section 4, we develop our approach. We show that a household's choice of deductible in a given line of coverage implies lower and upper bounds on its distorted

---

[1]In using the term probability distortions (or distorted probabilities), we do not mean to imply any negative connotation. Rather, we use the term probability distortions merely to refer to subjective beliefs or decision weights that differ from objective claim probabilities (as we estimate them).

probability of experiencing a claim in that coverage. Because the bounds are defined in terms of utility differences between options on its pricing menu, the household's choice effectively implies a relationship between its utility and probability distortion functions.[2] Because we observe three choices per household (one choice per coverage), we obtain three pairs of bounds—or *intervals*—per household (one interval per coverage). If a household had the same claim probability in each coverage, we could exploit the stability criterion to conduct inference in a relatively straightforward way: stability would require that the household's three intervals must intersect, and that its probability distortion function evaluated at this claim probability must be contained in the intersection. However, because a household's claim probabilities differ across coverages, each choice bounds its probability distortion function evaluated at a different point. Therefore, additional structure on the utility and probability distortion functions is necessary to give empirical content to the stability assumption.

We make two basic assumptions in addition to stability. The first is *constant absolute risk aversion* (CARA), our main restriction on the shape of the utility function.[3] Given CARA, a household's utility function is characterized by a single parameter, the Arrow–Pratt coefficient of absolute risk aversion, which we denote by r. The second basic assumption is *plausibility*: we require that there exist a single coefficient of absolute risk aversion and three distorted claim probabilities (one per coverage) that together can rationalize a household's choices. Altogether, 3629 households satisfy plausibility. Moving forward, we focus on this subsample of "rationalizable" households. This is out of necessity: by definition, no parameterization of the model, and thus none of the various behaviors and underlying models that it nests, can rationalize the choices of a household that violate plausibility.

In addition to CARA and plausibility, we consider five restrictions on the shape of the probability distortion function. The principal shape restriction is *monotonicity*, which requires that $\Omega(\cdot)$ is increasing. It ensures that the model obeys stochastic dominance in objective risk. It also places restrictions on subjective beliefs, depending on the underlying model. For instance, if the underlying model is subjective expected utility theory, monotonicity restricts subjective beliefs to be monotone transformations of objective risk. This is less restrictive, however, than the usual approach taken in the literature—assuming that subjective beliefs correspond to objective risk (see Barseghyan, Molinari, O'Donoghue, and Teitelbaum (2015b)). While we always consider monotonicity in the first instance (and generally view the results under monotonicity as our main results),

---

[2]For example, if a household chooses a deductible of $200 from a menu of $100, $200, and $250, then, loosely speaking, the lower bound is a function of the utility difference between the $200 and $250 options (more specifically, the difference in utility attributable to the difference in price between the $200 and $250 options and the difference in utility attributable to the $50 difference in coverage), and the upper bound is a function of the utility difference between the $100 and $200 options (more specifically, the difference in utility attributable to the difference in price between the $100 and $200 options and the difference in utility attributable to the $100 difference in coverage).

[3]In Section B of the Supplement, we show that our results are very similar if we instead assume constant relative risk aversion (CRRA) for reasonable levels of wealth. As we explain in Section 4.1, the utility differences among deductible options are nearly the same under CARA and CRRA, because the deductibles are small relative to wealth.

we often proceed to consider four additional shape restrictions on $\Omega(\cdot)$, adding them to the model sequentially in order of increasing strength. They are *quadraticity*, *linearity*, *unit slope*, and *zero intercept*. Together, these additional restrictions reduce the model to objective expected utility theory.

In Section 5, we use the rationalizable households' intervals to conduct inference on r and $\Omega(\cdot)$. First, we recover the distribution of the lower bound on r under each shape restriction on $\Omega(\cdot)$. We find, inter alia, that the distribution is skewed to the right under each shape restriction, and that the median is zero under each nondegenerate shape restriction. Next, we perform kernel regressions of the lower and upper bounds of the households' intervals as a function of their claim probabilities and use the results to draw inferences about the shape of $\Omega(\cdot)$. Under each nondegenerate shape restriction, the results evince a probability distortion function that substantially overweights small probabilities. Last, we use the intervals to analyze the benefits (in terms of gains in precision) and costs (in terms of loss of model fit) of imposing shape restrictions on $\Omega(\cdot)$.[4] We measure the benefit of a shape restriction by the average reduction in the size of the households' intervals due to the restriction. We measure the cost by the average perturbations to the households' intervals that would be required for every rationalizable household to satisfy the restriction. We conclude the section by drawing a connection between our cost statistic (which we label $\overline{Q}$), which measures the extent to which choice data violate expected utility maximization as generalized by the probability distortion model (with CARA utility and a given shape restriction on $\Omega(\cdot)$), and the efficiency index developed by Afriat (1967, 1972) and Varian (1990, 1993), which measures the extent to which choice data violate utility maximization (with a concave utility function).

In Section 6, we apply our approach to the problem of classifying households into preference types, where each type corresponds to a special case of the model. We find that four in five rationalizable households have intervals (i.e., make choices) that are consistent with a model with linear utility and monotone probability distortions,[5] whereas two in five have intervals that are consistent with a model with concave utility and no probability distortions (i.e., objective expected utility). Moreover, we find that nearly one in two rationalizable households *requires* monotone probability distortions to explain its intervals, whereas less than one in 20 requires concave utility. However, we also find that if we restrict the probability distortions to conform to either Gul disappointment aversion or KR loss aversion, then (i) the fraction of rationalizable households that have intervals that are consistent with the model (with either linear or concave utility) falls to two in five and (ii) the fraction that require probability distortions to explain their intervals falls to one in 30 (while the fraction that require concave utility rises to more than one in six), suggesting that other behaviors—namely, subjective beliefs or probability weighting—are playing an important role.[6] Indeed, when we restrict

---

[4]These benefits and costs are transparent and readily quantified under our approach. By contrast, they are difficult to isolate and measure under standard parametric approaches, because the impact of the shape restrictions is mediated in a complex way by the distributional assumptions.

[5]Insofar as the probability distortions reflect rank-dependent probability weighting, this model corresponds to Yaari's (1987) dual theory.

[6]To the extent that subjective beliefs obey monotonicity, they cannot be distinguished from probability weighting in our setting (Barseghyan, Molinari, O'Donoghue, and Teitelbaum (2013a)).

the model to have unit slope probability distortions (which we view as a parsimonious representation of KT probability weighting), we find that (i) three in five rationalizable households have intervals that are consistent with the model and (ii) nearly one in five require unit slope probability distortions to explain their intervals (while one in 10 require concave utility). At the end of the section, we explore the power of our revealed preference test, as measured by the success index proposed by Beatty and Crawford (2011). The results confirm that a model with monotone probability distortions is substantially more successful than a model with no probability distortions, and that unit slope distortions are more successful than those implied by Gul disappointment aversion or KR loss aversion.

From the results in Sections 5 and 6 we learn something about the extent and nature of preference heterogeneity among the rationalizable households. In many areas of research, however, economists study models that abstract from heterogeneity in preferences (e.g., representative agent models) and seek a single parameterization that best fits the data. In Section 7, we show how one can use the households' intervals to point estimate $\Omega(\cdot)$. Intuitively, we find the single probability distortion function that comes closest (in the Euclidean sense) to the monotone households' intervals. We prove that under mild conditions (satisfied in our data) the function is point identified, and we establish the consistency and asymptotic normality of our sample analog estimator. We then assess model fit given the minimum distance $\Omega(\cdot)$. For instance, we find that the model, when equipped with the minimum distance $\Omega(\cdot)$, can rationalize all three choices of nearly one in five monotone households. We also highlight the fact that, given the shape of $\Omega(\cdot)$, the residual deviation between the intervals and the minimum distance $\Omega(\cdot)$ gives us precisely the lower bound on the degree of heterogeneity in probability distortions among households.

In Sections 8 and 9, we wrap up our analysis by addressing two issues. First, we demonstrate a close connection between rank correlation of choices and stability of risk preferences under the probability distortion model. More specifically, we document that households' deductible choices are rank correlated across lines of coverage, echoing a similar finding by Einav, Finkelstein, Pascu, and Cullen (2012), and we show that it is the rationalizable households with monotone intervals who are driving these rank correlations. Second, we address concerns that the asymmetric information twins—moral hazard (unobserved action) and adverse selection (unobserved type)—may be biasing our results. With respect to moral hazard, we consider both ex ante and ex post moral hazard, and we conclude that neither is a significant issue in our data. With respect to adverse selection, we consider two possibilities—(i) there is heterogeneity in claim risk that is observed by the households but unobserved by the econometrician or (ii) there is heterogeneity in claim risk that is observed by the econometrician but unobserved by the households—and we show that our results and conclusions regarding probability distortions are robust to either possibility.

We offer concluding remarks in Section 10. Additional material is available in supplementary files on the journal website, http://qeconomics.org/supp/511/supplement.pdf and http://qeconomics.org/supp/511/code_and_data.zip.

1.1 *Related literature*

This paper builds on the literature on partial identification in econometrics (e.g., Manski (1989, 2003), Tamer (2010)). Our approach starts by asking what we can learn about the functionals characterizing risk preferences—in our model, the utility and probability distortions functions—when only stability and other minimal assumptions are imposed and revealed preference arguments are used to bound these functionals. We then sequentially add shape restrictions that increasingly constrain the model, so as to transparently show the role that each plays in sharpening the inference. To conduct statistical inference on the functional of primary interest—the probability distortion function—we apply recent techniques to build confidence sets for partially identified functionals (Imbens and Manski (2004), Beresteanu and Molinari (2008), Stoye (2009)). Next, we extend our approach to the problem of classification: we suppose that the data comprise a mixture of preference types and use our approach to bound the prevalence of each type. Last, we show how one can apply our approach to the problem of point estimation in a representative agent framework, and we develop a consistent estimator for the parameters of a linear predictor of the probability distortion function.

Our application of the partial identification approach to estimate and conduct inference on the parameters of a non-expected utility model has no precedent in the empirical literature on risk preferences, including in particular the strand of the literature that relies on data on market choices.

A handful of prior studies pursue related approaches to infer bounds on a single risk aversion parameter (e.g., the Arrow–Pratt coefficient of absolute risk aversion) within an expected utility framework. In particular, Barseghyan, Prince, and Teitelbaum (2011) and Einav et al. (2012) use data on insurance choices across multiple domains of coverage to obtain agent–domain-specific intervals of risk aversion parameters, and then calculate the fraction of agents for whom a single risk aversion parameter can rationalize its choices across domains.[7] Meanwhile, Sydnor (2010) uses data on deductible choices in home insurance to generate household-specific bounds on risk aversion, and then argues that the implied levels of risk aversion are implausibly large.

However, the extant papers that study non-expected utility models take a different approach to identification and estimation: they specify a random utility model, make statistical assumptions about the distribution of unobserved heterogeneity in preferences to obtain point identification of a single parameterization of the model, and estimate the model by parametric or nonparametric methods. For instance, Cicchetti and Dubin (1994) use data on telephone wire insurance choices to estimate a rank-dependent expected utility model by parametric maximum likelihood; Jullien and Salanié (2000) use data on bets on U.K. horse races to estimate a rank-dependent expected utility model and a cumulative prospect theory model by parametric maximum

---

[7]Barseghyan, Prince, and Teitelbaum (2011) use data on choices in three insurance domains. Einav et al. (2012) use data on choices in five insurance domains. Their data also include choices in one investment domain. Einav et al. (2012) also pursue a model-free approach in which they rank by risk the options within each domain and examine the rank correlation of agents' choices across domains. Observe that Barseghyan, Prince, and Teitelbaum (2011) and Einav et al. (2012) treat stability as a testable hypothesis, whereas we treat stability as an identifying restriction.

likelihood; Kliger and Levy (2009) use data on call options on the S&P 500 index to estimate a rank-dependent expected utility model and a cumulative prospect theory model by nonlinear least squares; Chiappori, Gandhi, Salanié, and Salanié (2012) use data on bets on U.S. horse races to estimate a non-expected utility model by nonparametric regression using generalized additive models (GAMs);[8] and Andrikogiannopoulou and Papakonstantinou (2013) use data on bets in an online sports book to estimate a cumulative prospect theory model by parametric Markov chain Monte Carlo (MCMC).[9] Andrikogiannopoulou and Papakonstantinou (2013) also estimate a mixture model of cumulative prospect theory that classifies bettors into preference types; however, they again estimate the model by parametric MCMC.[10]

We build directly upon Barseghyan, Molinari, O'Donoghue, and Teitelbaum (2013b) (hereafter, BMOT), who use the same data and model that we use in this paper. Like the previous studies, however, BMOT assume random utility and make assumptions to obtain point identification of a single parameterization of the model. They then estimate the model by semi-nonparametric maximum likelihood, parametric maximum likelihood, and parametric MCMC. By contrast, we aim to leverage the stability criterion to characterize the set of household-specific model parameterizations that are consistent with their choices across domains.

There are, of course, related empirical studies that adopt a partial identification approach in other areas of economic research. Examples include, among many others, Manski (2014), who uses revealed preferences arguments and shape restrictions to partially identify preferences for income and leisure and study their consequences for the evaluation of income tax policy; Dominitz and Manski (2011), who analyze probabilistic expectations of equity returns measured at two points in time, and use the partial identification approach to obtain bounds on the prevalence of expectations types in their sample; Chetty (2012), who obtains bounds on price elasticities in the presence of frictions such as adjustment costs or inattention; Ciliberto and Tamer (2009), who estimate payoff functions in a static, complete information entry game in airline markets in the presence of multiple equilibria; Haile and Tamer (2003), who study an incomplete model of English auctions and derive bounds on the distribution function characterizing bidder demand, on the optimal reserve price, and on the effects of observable covariates on bidder valuations, and apply their methodology to U.S. Forest Service timber auctions to evaluate reserve price policy; and Manski and Pepper (2000), who derive sharp bounds in the presence of monotone instrumental variables, and apply them to a study of the returns to education.

---

[8]Gandhi and Serrano-Padial (2015) use data on bets on U.S. horse races to estimate a cumulative prospect theory model by parametric maximum likelihood.

[9]Cohen and Einav (2007) use data on auto deductible choices to estimate an expected utility model by parametric MCMC. Paravisini, Rappoport, and Ravina (2015) use data on portfolio choices on a person-to-person lending platform to estimate an expected utility model by ordinary least squares (OLS).

[10]Bruhin, Fehr-Duda, and Epper (2010) use experimental data (choices over binary money lotteries) to estimate a mixture model of cumulative prospect theory by parametric maximum likelihood. Conte, Hey, and Moffatt (2011) use similar experimental data to estimate a mixture model of rank-dependent expected utility theory by parametric maximum simulated likelihood.

    This paper also builds on two strands of the literature on revealed preference.[11] The first strand pursues nonparametric methods for testing whether choice data are consistent with utility maximization (e.g., Afriat (1967), Diewert (1973), Varian (1982), Blundell, Browning, and Crawford (2003, 2008)), including whether such data are consistent with various restrictions on the form of the utility function, such as homotheticity, additive separability, infinite differentiability, strict concavity, and quasilinearity (e.g., Varian (1983), Chiappori and Rochet (1987), Matzkin and Richter (1991), Brown and Calsamiglia (2007), Cherchye, Demuynck, and De Rock (2015)), and for estimating or otherwise recovering the set of utility functions that are consistent with choice data (e.g., Varian (1982), Knoblauch (1992), Blundell, Browning, and Crawford (2003, 2008)).[12] Within this strand, our work most closely relates to the papers that study expected utility maximization and non-expected utility models of decision making under risk or uncertainty (e.g., Varian (1983), Green and Srivastava (1986), Varian (1988), Green and Osband (1991), Kubler, Selden, and Wei (2014), Echenique and Saito (2015), Polisson, Quah, and Renou (2015)). In essence, our monotonicity test is a semiparametric test in the revealed preference tradition of the consistency of individual choice data with a broad class of models of risky choice.[13] An important difference between our work and these papers, however, is that we study risky choice in a setting with discrete choice sets while they study settings with continuous choice sets.[14] Consequently their tests typically rely on differentiable demand conditions while our approach does not.

    The second related strand of the revealed preference literature comprises papers that develop measures of goodness-of-fit of revealed preference tests, which assess the extent to which choice data violate the utility maximization hypothesis (e.g., Afriat (1972), Houtman and Maks (1985), Varian (1985), Swofford and Whitney (1987), Varian (1990, 1993), Famulari (1995), Gross (1995)) , as well as papers that develop measures of the power of revealed preference tests (e.g., Bronars (1987), Beatty and Crawford (2011), Andreoni, Gillen, and Harbaugh (2013)). As noted above, we develop a measure—our $\overline{Q}$ statistic—that assesses the fit of the probability distortion model given different shape restrictions on $\Omega(\cdot)$, and we draw a connection between our $\overline{Q}$ statistic and the prominent Afriat–Varian efficiency index. Moreover, we employ the Beatty–Crawford success measure to gauge the power of our tests of shape restrictions on $\Omega(\cdot)$.

---

[11]For reviews of this literature, see, for example, Varian (2005), Cherchye, Crawford, De Rock, and Vermeulen (2009), and Crawford and De Rock (2014).

[12]Much of the work in this strand contemplates individual, static choice, and linear budget sets. In related work, researchers pursue revealed preference tests for intertemporal choice (e.g., Browning (1989)), nonlinear budget sets (e.g., Matzkin and Richter (1991)), market data (e.g., Brown and Matzkin (1996)), and collective choice (e.g., Chiappori (1988), Cherchye, De Rock, and Vermeulen (2007)). This work is surveyed by Crawford and De Rock (2014).

[13]We take a nonparametric approach with respect to the probability distortion function, which is the object of our primary interest. At the same time, we assume a parametric form (namely, CARA) for the Bernoulli utility function. This assumption is not very restrictive in our setting, however, as the range of deductible options in each coverage is not very large and, therefore, assuming a differentiable Bernoulli utility function, local curvature is what matters.

[14]A handful of papers study riskless choice with discrete choice sets (e.g., Polisson and Quah (2013), Forges and Iehlé (2014), Cosaert and Demuynck (2015)).

2. Insurance data

We acquired the data from a large U.S. property and casualty insurance company. The company offers several lines of insurance, including auto and home. The full data set contains annual information on more than 400,000 households who held auto or home policies between 1998 and 2006. The data contain all the information in the company's records regarding the households and their policies.

We focus on three lines of coverage: auto collision coverage, auto comprehensive coverage, and home all perils coverage. 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, or vandalism), without regard to fault. 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 the sake of brevity, we often refer to home all perils simply as home.

In our analysis, we restrict attention to households who (i) purchased all coverages (auto collision, auto comprehensive, and home) and (ii) first purchased each coverage in the same year, in either 2005 or 2006. The former restriction maximizes the number of choices that we observe per household. The more choices we observe for a household, the more precise is the inference we can make about the household's risk preferences. The latter restriction avoids temporal issues, such as changes in the economic environment. For households who first purchased their auto and home policies in 2005 and renewed their policies in 2006, we observe their deductible choices at the time of first purchase and at the time of renewal. In our analysis, we consider only the deductible choices at the time of first purchase. This is meant to increase confidence that we are working with active choices; one might worry that households renew their policies without actively reassessing their deductible choices (Handel (2013)). Together, these restrictions yield a core sample of 4170 households.

2.1 *Deductible choices and pricing menus*

For each household in our sample, we observe its deductible choices in each coverage, as well as the premium paid by the household in each coverage. Moreover, we observe the coverage-specific menus of premium–deductible combinations that were available to each household at the time it made its deductible choices. According to conversations with the company and an independent agent who sells auto and home policies for the company, the choice environment is conducive to households making active and informed deductible choices: there are no default choices, the full pricing menu of premium–deductible combinations is available to a household at the time it makes a choice, and a household must choose a deductible separately for each coverage (the choice made in one coverage does not automatically become the default choice in another coverage).[15]

[15]Indeed, the choice set is not exactly the same across coverages (see Table S1 in the Supplement), and so it could not be the case that the choice made in one coverage automatically becomes the default choice

In each coverage, the company uses the same basic procedure to generate a household's pricing menu. The company first determines a household's base price $\bar{p}$ according to a coverage-specific rating function, which takes into account the household's coverage-relevant characteristics and any applicable discounts. Using the base price, the company then generates the household's pricing menu $\mathcal{M} = \{(p(d), d) : d \in \mathcal{D}\}$, 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, $p(d) = (g(d) \cdot \bar{p}) + \delta$, where $g(\cdot)$ is a decreasing positive function and $\delta > 0$. The multiplicative factors $\{g(d) : d \in \mathcal{D}\}$ are known as the deductible factors and δ is a small markup known as the expense fee. The deductible factors and the expense fee are coverage specific but household invariant.

Table S1 in the Supplement displays the deductible choices of the households in our core sample. In each coverage, the modal deductible choice is $500. Table S2 in the Supplement summarizes the pricing menus. For each coverage, it summarizes the (annual) premium associated 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 premium for coverage with a $500 deductible is $180 for auto collision, $115 for auto comprehensive, and $679 for home. The average cost of decreasing the deductible from $500 to $250 is $54 for auto collision, $30 for auto comprehensive, and $56 for home. The average savings from increasing the deductible from $500 to $1000 is $41 for auto collision, $23 for auto comprehensive, and $74 for home.[16]

2.2 *Claim probabilities*

For purposes of our analysis, we need to estimate each household's risk of experiencing a claim in 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 data set: 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 mean λ_{ijt}. In addition, we assume that deductible choices do not influence claim rates, that is, households do not suffer from moral hazard.[17] We treat the claim rates as latent random variables and assume that

$$\ln \lambda_{ijt} = \mathbf{X}'_{ijt}\boldsymbol{\beta}_j + \varepsilon_{ij},$$

where $\mathbf{X}_{ijt}$ is a vector of observables and $\exp(\varepsilon_{ij})$ follows a gamma distribution with unit mean and variance ϕ_j. We perform Poisson panel regressions with random effects to obtain maximum likelihood estimates of $\boldsymbol{\beta}_j$ and ϕ_j for each coverage j.[18]

---

in another coverage. That said, we cannot know what advice or guidance a selling agent may or may not provide to a household about its several deductible choices.

[16]Tables S1 and S2 also appear in BMOT. They are reproduced in the Supplement for the reader's convenience.

[17]We revisit this assumption in Section 9.

[18]The results of the regressions are reported in Tables A.4 and A.5 of the BMOT Online Appendix.

Next, we use the regression results to assign claim probabilities to the households in the core sample. For each household i, we use the regression estimates to calculate a fitted claim rate $\widehat{\lambda}_{ij}$ for each coverage j, conditional on the household's ex ante observables and ex post claims experience.[19] 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 a household experiences at most one claim.[20] Given this assumption, we transform $\widehat{\lambda}_{ij}$ into a claim probability μ_{ij} using

$$\mu_{ij} \equiv 1 - \exp(-\widehat{\lambda}_{ij}),$$

which follows from the Poisson probability mass function.

Table S3 in the Supplement summarizes the claim probabilities in the core sample. The mean claim probabilities in auto collision, auto comprehensive, and home are 0.069, 0.021, and 0.084, respectively. In our analysis, we assume these estimated claim probabilities are correct in the sense that they correspond to the households' true claim probabilities. In Section 9, we revisit this assumption and address the concern that unobserved heterogeneity in households' claim risk may be biasing our results.

3. The model

Households have access to three lines of insurance coverage: auto collision (L), auto comprehensive (M), and home all perils (H). Policies in each line of coverage provide full insurance against covered losses in excess of a deductible chosen by the household. We assume that a household treats its deductible choices as independent decisions. This assumption is motivated by the literature on narrow bracketing (e.g., Read, Loewenstein, and Rabin (1999)).

In each coverage $j \in \{L, M, H\}$, household i faces a menu of premium–deductible pairs, $\mathcal{M}_{ij} = \{(p_{ij}(d), d) : d \in \mathcal{D}_j\}$, where $p_{ij}(d)$ is the premium associated with deductible d and $\mathcal{D}_j$ is the set of deductible options. We assume that the household experiences at most one claim during the policy period, and that the probability of experiencing a claim is μ_{ij}. We also assume that any claim exceeds the highest available deductible, payment of the deductible is the only cost associated with a claim, and the household's deductible choice does not influence its claim probability.[21]

---

[19]More specifically, $\widehat{\lambda}_{ij} = \exp(\mathbf{X}'_{ij}\widehat{\boldsymbol{\beta}}_j)E(\exp(\varepsilon_{ij})|\mathbf{Y}_{ij})$, where $\mathbf{Y}_{ij}$ records household i's claims experience under coverage j after purchasing the policy and $E(\exp(\varepsilon_{ij})|\mathbf{Y}_{ij})$ is calculated assuming $\exp(\varepsilon_{ij})$ follows a gamma distribution with unit mean and variance $\widehat{\phi}_j$.

[20]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. Given this assumption, we could use a binary choice model such as logit or probit instead of the Poisson model. However, this would lead to a loss of precision in estimation (see, generally, Cameron and Trivedi (1998, pp. 85–87)).

[21]We make the first 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) and BMOT. Including the $2500 and $5000 deductible options in the home menu would not materially change our results; see Section H of the Supplement.

Under the foregoing assumptions, the household's choice of deductible in each coverage involves a choice among lotteries of the form

$$L_{ij}(d) \equiv \big(-p_{ij}(d), 1 - \mu_{ij}; -p_{ij}(d) - d, \mu_{ij}\big).$$

To model households' preferences over deductibles, we adopt the probability distortion model considered by BMOT. The model is a generalization of objective expected utility theory that allows for probability distortions. According to the model, a household chooses deductible $d \in \mathcal{D}_j$ to maximize

$$V_{ij}\big(L_{ij}(d)\big) \equiv \big(1 - \Omega_{ij}(\mu_{ij})\big)u_{ij}\big(w_i - p_{ij}(d)\big) + \Omega_{ij}(\mu_{ij})u_{ij}\big(w_i - p_{ij}(d) - d\big),$$

where w_i is the household's wealth, $u_{ij}(\cdot)$ is its utility function, and $\Omega_{ij}(\cdot)$ is its probability distortion function.[22]

The probability distortion model has two principal virtues. The first is that it allows for the possibility that a household's aversion to risk is driven not only by the shape of its utility function, but also by the shape of its probability distortion function.[23] Stated another way, the model allows for the possibility that a household's demand for insurance is driven not only by the way in which it evaluates outcomes, but also by the way in which it evaluates risk. The second principal virtue of the model is that the probability distortion function can capture a wide range of different behaviors, including the following.

- Subjective beliefs: when $\Omega(\mu) \neq \mu$.

- Rank-dependent probability weighting: when $\Omega(\mu)$ is a probability weighting function (PWF), that is, an increasing function that maps $[0, 1]$ onto $[0, 1]$.

- KT probability weighting: when $\Omega(\mu)$ is a PWF that satisfies overweighting and subadditivity for small probabilities, as well as subcertainty and subproportionality.[24]

- Gul disappointment aversion: when $\Omega(\mu) = \mu(1 + \beta)/(1 + \beta\mu)$, $\beta \geq 0$.

- KR loss aversion: when $\Omega(\mu) = \mu + \Lambda(1 - \mu)\mu$, $\Lambda \geq 0$.[25]

---

[22]Because of the narrow bracketing assumptions, one could argue that the model is not a strict generalization of expected utility theory, on the grounds that expected utility theory requires broad bracketing (integration of all risky choices into a single, joint decision). However, one could also argue that expected utility theory is consistent with narrow bracketing and does not necessarily require broad bracketing, particularly in light of the "small worlds" discussion of Savage (1954). For a thoughtful discussion of this issue, see Read (2009).

[23]Under expected utility theory (objective or subjective), by contrast, aversion to risk is driven solely by the shape of the utility function, which arguably is problematic (Rabin (2000)).

[24]A PWF $\Omega(\mu)$ satisfies overweighting if $\Omega(\mu) > \mu$; satisfies subadditivity if $\Omega(\nu\mu) > \nu\Omega(\mu)$ for $0 < \nu < 1$; satisfies subcertainty if $\Omega(\mu) + \Omega(1 - \mu) < 1$ for $0 < \mu < 1$; and satisfies subproportionality if $\Omega(\mu\nu)/\Omega(\mu) \leq \Omega(\mu\nu\tau)/\Omega(\mu\tau)$ for $0 < \mu, \nu, \tau \leq 1$. See Kahneman and Tversky (1979).

[25]In Gul's model, β captures the degree of disappointment aversion. In KR's model, Λ effectively captures the degree of loss aversion. (To be clear, we refer to the version of KR's model in which the solution concept is a choice-acclimating personal equilibrium.) For details, see BMOT.

As a result, the model nests a number of underlying models, including the following.

- Objective expected utility theory: when $\Omega(\mu) = \mu$.

- Subjective expected utility theory: when $\Omega(\mu) \neq \mu$.

- Yaari's dual theory: when u is linear and $\Omega(\mu)$ is a PWF.

- Rank-dependent expected utility theory: when $\Omega(\mu)$ is a PWF.[26]

- Gul's disappointment aversion model: when $\Omega(\mu) = \mu(1 + \beta)/(1 + \beta\mu)$, $\beta \geq 0$.

- KR's loss aversion model: when $\Omega(\mu) = \mu + \Lambda(1 - \mu)\mu$, $\Lambda \geq 0$.[27]

4. STABILITY AND Ω-INTERVALS

The model as presented in Section 3 allows preferences to be context dependent, that is, $V_{ij}(L_{ij}(d)) \neq V_{ik}(L_{ik}(d))$ for $j \neq k$. Economists, however, desire models of decision making that obey context invariance, or stability, both because they seek a theory of decision that can explain choices across multiple domains and because they view stability as an essential aspect of rationality (Kahneman (2003)).[28] Stability requires that $V_{ij}(L_{ij}(d)) = V_i(L_{ij}(d))$ for every coverage j. In particular, stability requires that a household's utility and probability distortion functions are context invariant.

A0. STABILITY. *Both $u_{ij}(\cdot) = u_i(\cdot)$ and $\Omega_{ij}(\cdot) = \Omega_i(\cdot)$ for all j.*

Under stability, the principle of revealed preference implies that $V_i(L_{ij}(d^*)) \geq V_i(L_{ij}(d))$ for every deductible $d \in \mathcal{D}_j$ when household i chooses deductible $d^* \in \mathcal{D}_j$ under coverage j. It follows that a household's choice of deductible implies bounds on its distorted probability $\Omega_i(\mu_{ij})$, which bounds are defined in terms of utility differences among deductible options:

$$\text{LB}_{ij} \leq \Omega_i(\mu_{ij}) \leq \text{UB}_{ij},$$

where

$$\text{LB}_{ij} \equiv \max\left\{0, \max_{d > d^*} \Delta_{ij}\right\} \quad \text{and} \quad \text{UB}_{ij} \equiv \min\left\{1, \min_{d < d^*} \Delta_{ij}\right\}$$

---

[26]Because all deductible lotteries are in the loss domain, the model also nests cumulative prospect theory in our setting (Tversky and Kahneman (1992)).

[27]Observe that the probability distortion function that corresponds to Gul disappointment aversion is a PWF, but that the probability distortion function that corresponds to KR loss aversion is not a PWF (because $\Omega(\mu) = \mu + \Lambda(1 - \mu)\mu$ can lie outside [0, 1] for some $\mu \in (0, 1)$ if Λ is large).

[28]Of course, nonstable subjective beliefs do not violate rationality (in the sense of Savage). But if subjective beliefs are wholly nonstable—that is, if they are entirely context dependent and lack any domain-general component—then we have an extreme "small worlds" problem (again, in the sense of Savage). We cannot hope to develop any model of decision making that can explain choices across multiple domains, even domains that are closely related or essentially similar. Instead, we can only hope to develop ad hoc models, each capable of explaining choices within a specific domain.

and

$$\Delta_{ij} \equiv \big(u_i\big(w_i - p_{ij}(d)\big) - u_i\big(w_i - p_{ij}(d^*)\big)\big)$$
$$/\big\{\big[u_i\big(w_i - p_{ij}(d)\big) - u_i\big(w_i - p_{ij}(d) - d\big)\big]$$
$$- \big[u_i\big(w_i - p_{ij}(d^*)\big) - u_i\big(w_i - p_{ij}(d^*) - d^*\big)\big]\big\}.$$

Let $\mathcal{I}_{ij} \equiv [\text{LB}_{ij}, \text{UB}_{ij}]$. We refer to $\mathcal{I}_{ij}$ as the household's Ω-*interval* for coverage j.

The model under stability has empirical content, but it is limited. Provided that $\mathcal{I}_{ij}$ is nonempty (i.e., $\text{LB}_{ij} \leq \text{UB}_{ij}$), the model is rejected for a household only if (i) it has identical claim probabilities in two lines of coverage j and k and (ii) its Ω-intervals for coverages j and k do not intersect. In general, however, a household's pricing menus and claim probabilities differ across coverages. When this is the case, the model cannot be rejected for the household if its Ω-intervals do not intersect. To increase the model's empirical content, it is necessary to impose additional structure on the household's utility and probability distortion functions. With this additional structure, we can use the households' Ω-intervals to conduct inference on $u_i(\cdot)$ and $\Omega_i(\cdot)$ and draw conclusions about the various behaviors and underlying models that are encompassed by the model.

4.1 *CARA and plausibility*

In addition to stability, we make two basic assumptions. The first is *constant absolute risk aversion* (CARA).

A1. CARA. *The ratio $u_i''(w_i)/u_i'(w_i)$ is a constant function of w_i.*

This is the principal shape restriction on the utility function.[29] Assuming CARA has two key virtues. First, $u_i(\cdot)$ is fully characterized by a single household-specific parameter—the coefficient of absolute risk aversion, $r_i \equiv -u_i''(w_i)/u_i'(w_i)$. Second, the bounds of the Ω-intervals do not depend on wealth w_i, which is unobserved:[30]

$$\Delta_{ij} = \big(\exp\big(r_i p_{ij}(d)\big) - \exp\big(r_i p_{ij}(d^*)\big)\big)$$
$$/\big\{\big[\exp\big(r_i p_{ij}(d)\big) - \exp\big(r_i\big(p_{ij}(d) + d\big)\big)\big]$$
$$- \big[\exp\big(r_i p_{ij}(d^*)\big) - \exp\big(r_i\big(p_{ij}(d^*) + d^*\big)\big)\big]\big\}.$$

CARA is a common assumption in economics. However, the lack of a wealth effect is troubling to some, particularly those who believe that decreasing absolute risk aversion is more plausible. This leads some to assume constant relative risk aversion (CRRA). It is easy to show that for reasonable levels of wealth, the utility differences among deductible options under CRRA—which assumes that the coefficient of relative

---

[29]Observe that CARA implicitly relies on two presumptions: (i) $u_i(\cdot)$ is twice differentiable and (ii) $u_i'(w) \neq 0$ for all w. Of course, the latter presumption follows from the fact that $u_i(\cdot)$ is a utility function, which implies that it is increasing.

[30]If $r = 0$, then $\Delta_{ij} = [p_{ij}(d) - p_{ij}(d^*)]/[d^* - d]$.

risk aversion, $\rho_i \equiv w_i \times r_i$, is a constant function of wealth—are very similar to those under CARA. This is because the deductibles are small relative to wealth,[31] and thus what matters is the local curvature of the utility function around initial wealth.[32] Consequently, assuming CRRA instead of CARA yields very similar results, as we show in Section B of the Supplement.[33] We also show that another class of utility functions used in the literature—namely, those with a negligible third derivative (NTD) (Cohen and Einav (2007), Barseghyan, Prince, and Teitelbaum (2011))—yields very similar results as well.

The second basic assumption is *plausibility*.

A2. PLAUSIBILITY. *There exists $r_i \geq 0$ such that* $\mathrm{LB}_{ij} \leq \mathrm{UB}_{ij}$ *for all j.*

Plausibility requires that there exists a positive coefficient of absolute risk aversion such that the household's Ω-intervals are nonempty. Stated another way, it requires that there exist a concave utility function and three distorted claim probabilities (one for each coverage) that together can rationalize the household's choices.[34] The requirement that there exist some r_i such that $\mathrm{LB}_{ij} \leq \mathrm{UB}_{ij}$ is a prerequisite for making any inferences about $u_i(\cdot)$ and $\Omega_i(\cdot)$. Restricting $r_i \geq 0$ is motivated by the law of diminishing marginal utility.

In what follows, we also restrict $r_i \leq 0.0108$. Placing an upper bound on r_i is necessary to make checking plausibility computationally feasible. We set the upper bound at 0.0108 for reasons we explain below in Section 5. Increasing the upper bound on r_i would not substantially change our results; see Section C of the Supplement.

Altogether, 541 households (13.0 percent) violate plausibility. Of these households, virtually every one chose an auto collision deductible of $200. Given the pricing menu for auto collision coverage, this is an implausible choice for nearly every household. The intuition is best illustrated in the case of linear utility ($r = 0$). For auto collision coverage, the pricing rule, $p(d) = (g(d) \cdot \bar{p}) + \delta$, is such that $g(100) = 1.15$, $g(200) = 1.00$, and $g(250) = 0.85$. For any base price $\bar{p}$, therefore, the choice of $200 implies

$$\mathrm{LB} = \frac{p(200) - p(250)}{250 - 200} = \frac{0.15}{50}\bar{p} \quad \text{and} \quad \mathrm{UB} = \frac{p(100) - p(200)}{200 - 100} = \frac{0.15}{100}\bar{p}.$$

Hence, the lower bound exceeds the upper bound, whatever the base price. The intuition is straightforward: if a household's distorted claim probability is high enough that it prefers a deductible of $200 over a deductible of $250, then it also should prefer $100 over $200. Conversely, if the household's distorted claim probability is low enough that it prefers $200 over $100, then it also should prefer $250 over $200. Allowing $r > 0$ disrupts

---

[31]Each of the households in our sample owns a home and at least one auto.

[32]Note that under CRRA, the household's absolute risk aversion, $r_i = \rho_i/w_i$, can be driven by its wealth or its relative risk aversion. Hence, the model implicitly allows for heterogeneity both in wealth and in utility curvature.

[33]The CRRA results are robust to substantial changes in wealth. It follows that our results are also robust to a broader class of risk preferences, hyperbolic absolute risk aversion (HARA), provided that the absolute value of the additive term is not too large relative to wealth.

[34]Observe that plausibility embeds a restriction contained in the definitions of LB_{ij} and UB_{ij}, namely that the household's distorted claim probabilities lie between 0 and 1.

this logic only for absurd levels of absolute risk aversion; see Section C of the Supplement.

We note that violations of plausibility are not an artifact of assuming CARA. We find very similar violation rates under CRRA and NTD utility. Given this and the fact that the model nests expected utility theory and several of the leading alternative models, we treat the households that violate plausibility as nonrationalizable and drop them moving forward. We refer to the remaining subsample of 3629 households that satisfy plausibility as the *rationalizable* households.[35]

4.2 *Shape restrictions on Ω*

We complete the model with shape restrictions on the probability distortion function.

The principal restriction is *monotonicity*, which by itself does not impose parametric restrictions on $\Omega_i(\cdot)$.

A3. MONOTONICITY. *If $\mu_{ij} \leq \widetilde{\mu}_{ij}$, then $\Omega_i(\mu_{ij}) \leq \Omega_i(\widetilde{\mu}_{ij})$.*

Monotonicity requires that the probability distortion function is increasing. It is a standard assumption in prospect theory and other models that feature probability transformations. Kahneman and Tversky (1979, p. 280) go so far as to say that it is a "natural" assumption. By definition, a probability distortion function that satisfies monotonicity is a PWF. In the case of our model, monotonicity ensures that the model obeys stochastic dominance in objective risk (e.g., Ingersoll (2008)). It also places restrictions on subjective beliefs, depending on the underlying model. For instance, if the underlying model is subjective expected utility theory, monotonicity restricts subjective beliefs to be monotone transformations of objective risk. This is less restrictive, however, than the usual approach taken in the literature—assuming that subjective beliefs correspond to objective risk (see Barseghyan et al. (2015b)).

While we always impose monotonicity in the first instance, we often proceed to consider four additional shape restrictions on $\Omega_i(\cdot)$, adding them to the model sequentially in order of increasing strength.

The first two are *quadraticity* and *linearity*.

A4. QUADRATICITY. $\Omega_i(\mu_{ij}) = a + b\mu_{ij} + c(\mu_{ij})^2$, *where $b \geq 0$ and $c \geq -b/2$.*

A5. LINEARITY. $\Omega_i(\mu_{ij}) = a + b\mu_{ij}$, *where $b \geq 0$.*

---

[35]This raises an intriguing question for future research: What model could rationalize the choices of the households that violate plausibility? There are, of course, numerous other models that one could study. However, their potential to rationalize the choices of the households in our sample is unclear. Take, for example, an expected utility model with state-dependent utility (Karni (1985)). While a state-dependent utility model may be apposite in a number of insurance settings (e.g., flight insurance, life insurance, catastrophic health insurance, and disability insurance), we believe it is not well suited to ours. In our setting, households are not insuring against death or disability, where it seems reasonable that the utility of money would depend on the state of the world, but rather against damage to replaceable property, where it does not.

Quadraticity and linearity require that the probability distortion function be quadratic and linear, respectively. The parameter restrictions on b and c in A4 and on b in A5 follow from monotonicity. A quadratic specification parsimoniously allows for nonlinear distortions; importantly, it is sufficiently flexible to capture the left side of the classic inverse-S shape. The left side is what is relevant for our data: 98.1 percent of the claim probabilities in the rationalizable subsample lie between zero and 0.16, and 99.8 percent lie between zero and 0.25. A linear specification is more restrictive, permitting only linear distortions. It turns out, however, that nearly every household that satisfies quadraticity also satisfies linearity.[36]

The final two restrictions are *unit slope* and *zero intercept*.

A6. UNIT SLOPE. $\Omega_i(\mu_{ij}) = a + \mu_{ij}$.

A7. ZERO INTERCEPT. $\Omega_i(\mu_{ij}) = \mu_{ij}$.

Unit slope requires that the probability distortion function is linear and perfectly sensitive to changes in probability. Zero intercept requires unit slope and $\Omega_i(0) = 0$. Observe that imposing zero intercept effectively disallows probability distortions and reduces the model to objective expected utility theory.

5. INFERENCE ON r AND Ω

Table 1, column (a) reports the percentage of rationalizable households that satisfy each shape restriction on $\Omega_i(\cdot)$. To be clear, a household satisfies a shape restriction if there exists some probability distortion function that both satisfies the restriction and is consistent with the household's Ω-intervals for some $r_i \in [0, 0.0108]$. Roughly, five in six ra-

TABLE 1. Shape restrictions on Ω. Rationalizable subsample (3629 households).

Shape Restriction	(a) Percent of Households Satisfying Restriction	(b) Average Reduction in Size of Ω-Intervals		(c) Ratio of Average Size of Ω-Intervals to Average Size of BMOT 95 Percent Confidence Band	(d) Average Value of Q (PP)
		PP	Percent		
Monotonicity	84.8	4.4	24.1	3.17	0.6
Quadraticity	82.0	6.2	28.4	2.59	1.9
Linearity	80.4	8.2	41.7	2.06	2.0
Unit slope	61.6	14.0	69.2	0.98	3.6
Zero intercept	39.6	15.6	100.0	0.00	11.9

Note: PP stands for percentage points. In column (b), we calculate the average reduction in the size of the Ω-intervals due to a shape restriction as follows: (i) we restrict attention to the subsample of rationalizable households that satisfy the shape restriction and (ii) we compare, for each household in this subsample, the size of its Ω-intervals before imposing any shape restrictions (calculated at the minimum r_i for which the household satisfies plausibility) to the size of its Ω-intervals after imposing the shape restriction (calculated at the minimum plausible r_i for which the household satisfies the shape restriction). In column (c), we calculate the size of the Ω-intervals under a shape restriction in the same manner.

[36]See Table 1, column (a).

tionalizable households satisfy monotonicity; four in five satisfy quadraticity and linearity; three in five satisfy unit slope; and two in five satisfy zero intercept.[37]

We emphasize that the relatively low success rate of the zero intercept model (i.e., the objective expected utility model) is not an artifact of restricting $r_i \leq 0.0108$. In fact, this upper bound on r_i was chosen to maximize the success rate of the zero intercept model. What is more, increasing the upper bound on r_i would not substantially change our results. For further discussion, see Section C of the Supplement.

5.1 *Minimum plausible r*

When a rationalizable household satisfies a shape restriction on $\Omega_i(\cdot)$, it satisfies the restriction for more than one plausible value of r_i. We focus on the *minimum plausible r_i*—that is, the minimum plausible value of r_i for which the household satisfies the restriction. We focus on the minimum plausible r_i for two reasons. The first is the Rabin (2000) critique, which implies that relying on large values of r_i to rationalize aversion to modest-stakes risk is problematic, because large values of r_i imply absurd levels of aversion to large-stakes risk. The second reason is that once we restrict attention to the small values of r_i for which the household satisfies the restriction, focusing on the smallest value is effectively without loss of generality; the household's Ω-intervals are effectively the same whether we fix r_i at the minimum plausible value or at some other small value. Moving forward, we always pin down a household's Ω-intervals by fixing r_i at its minimum plausible value.

Figure 1 plots, for each shape restriction on $\Omega_i(\cdot)$, the distribution of the minimum plausible r_i among the rationalizable households that satisfy the restriction. In general, the figure evinces that there is heterogeneity in the minimum plausible r_i across households (cf. Cohen and Einav (2007)). More specifically, it reveals two important facts. First, the distribution is skewed to the right under each restriction, and indeed the median is zero under each restriction, save only zero intercept. This implies that once we allow for probability distortions (even linear, unit slope distortions), a majority of rationalizable households do not require concave utility to rationalize their choices. Second, the mean strictly decreases as we relax the shape restrictions on $\Omega_i(\cdot)$ (i.e., move from zero intercept to monotonicity). This implies that as we allow for more flexible probability distortions, the rationalizable households on average require less utility curvature to rationalize their choices.

Figure 2 displays the percentage of rationalizable households that satisfy each shape restriction as we increase the upper bound on r_i from zero to 0.0108. It reveals two additional key facts about the distribution of the minimum plausible r_i. First, between 70 and 80 percent of rationalizable households do not require a positive r_i to satisfy monotonicity, quadraticity, or linearity, and nearly 50 percent do not require a positive r_i to satisfy unit slope. Put differently, even if we impose linear utility, the model can rationalize the choices of the vast majority of rationalizable households if we allow for monotone,

---

[37]Each shape restriction can be represented by a system of linear inequalities. The statistics reported in the table are constructed by checking for each household whether the implied system of inequalities has a nonempty solution.
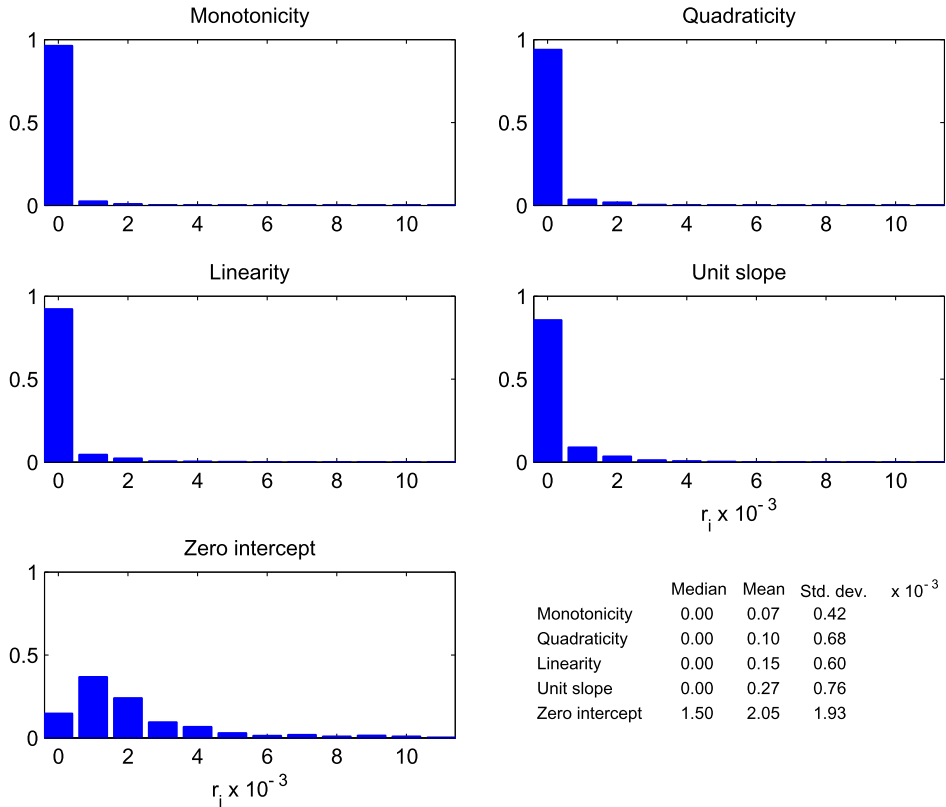
FIGURE 1. Distribution of minimum plausble r_i. Each graph is a histogram of the minimum plausible r_i for a given shape restriction on $\Omega_i(\cdot)$.

quadratic, or linear probability distortions, and it can rationalize the choices of nearly a majority of rationalizable households if we allow for unit slope probability distortions. By contrast, if we allow for concave utility but do not allow for probability distortions, the model can rationalize the choices of less than 40 percent of rationalizable households. Second, as we increase the upper bound on r_i above zero, the percentage of rationalizable households that satisfy monotonicity, quadraticity, linearity, and unit slope increases by 5–10 percentage points and then levels off once the upper bound on r_i surpasses about 0.003.[38] On the one hand, this confirms that there is important heterogeneity in the minimum plausible r_i. On the other hand, however, it suggests that once we allow for monotone, quadratic, linear, or even unit slope probability distortions, we gain relatively little by admitting this heterogeneity, and almost nothing by allowing for large values of r_i.

---

[38]In Section D of the Supplement, we discuss the intuition for why the percentage of rationalizable households that satisfy monotonicity increases as we increase the upper bound on r_i.
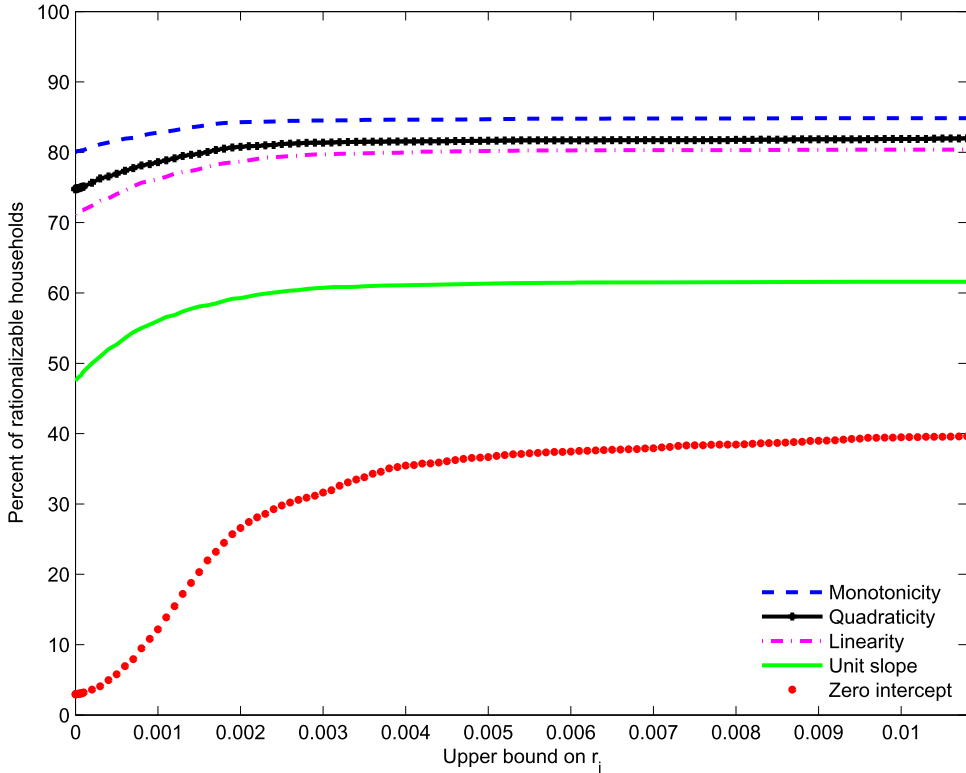
FIGURE 2. Percentage of rationalizable households that satisfy each shape restriction as we increase the upper bound on r_i.

5.2 Ω-Intervals

Figure 3 depicts the average bounds on $\Omega(\mu)$ under each nondegenerate shape restriction (i.e., A3–A6). In particular, each frame displays, for a given restriction, kernel regressions of the lower and upper bounds of the Ω-intervals as a function of μ for the subsample of rationalizable households that satisfy the restriction.[39] We can draw several conclusions from the Ω-intervals depicted in Figure 3.

First, the Ω-intervals evidence large probability distortions. Under each nondegenerate shape restriction, the households' Ω-intervals are consistent with a probability distortion function that substantially overweights small probabilities. Under monotonicity, for instance, the midpoints of the Ω-intervals imply $\Omega(0.02) = 0.11$, $\Omega(0.05) = 0.17$, and

---

[39]To lessen the first-order bias term typical of kernel regression, we use a fourth-order Gaussian kernel. The bandwidth used in the regressions is chosen via cross-validation. Specifically, we obtain via cross-validation an optimal bandwidth for the lower points of the Ω-intervals and an optimal bandwidth for the upper points of the Ω-intervals. We then use the average of the two bandwidths, which leads to undersmoothing of the lower bound and some oversmoothing of the upper bound. For the lower bound, undersmoothing obtains that the asymptotic distribution of the estimator is centered at zero (see, e.g., Jones (1995), Horowitz (2009)). We report confidence bands that (pointwise in μ) cover the estimated regression intervals with asymptotic probability 95 percent, using the nonparametric bootstrap procedure detailed in Beresteanu and Molinari (2008, Section 3).
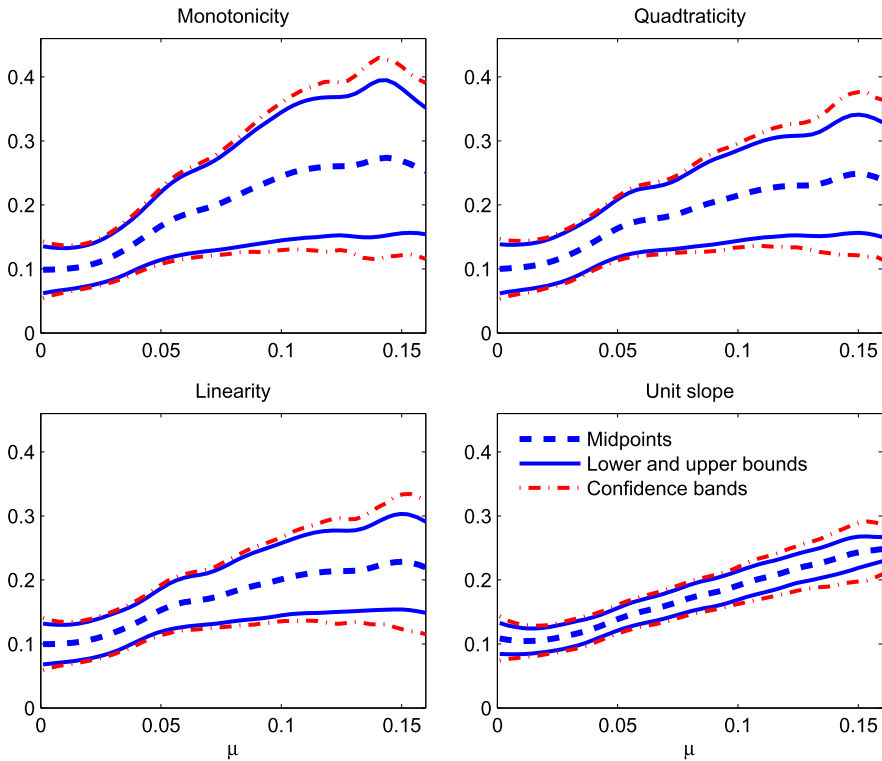
FIGURE 3. Average bounds on $\Omega(\mu)$. Each frame displays, for a nondegenerate shape restriction, kernel regressions of the lower and upper bounds of the Ω-intervals as a function of μ for the subsample of rationalizable households that satisfy the restriction. We use a fourth-order Gaussian kernel. The bandwidth is chosen via cross-validation. Specifically, we obtain via cross–validation an optimal bandwidth for the lower points of the Ω-intervals (0.016) and an optimal bandwidth for the upper points of the Ω-intervals (0.003). We then use the average of the two bandwidths. We report confidence bands that (pointwise in μ) cover the estimated Ω-intervals with asymptotic probability 95 percent, using the nonparametric bootstrap procedure detailed in Beresteanu and Molinari (2008, Section 3).

$\Omega(0.10) = 0.25$, and indeed even the lower bounds of the Ω-intervals imply $\Omega(0.02) = 0.07$, $\Omega(0.05) = 0.11$, and $\Omega(0.10) = 0.15$.[40]

Second, the Ω-intervals suggest a probability distortion function that bears a striking resemblance to the probability weighting function originally posited by Kahneman and Tversky (1979), in the range of our data. In particular, the Ω-intervals are consistent with a function that exhibits overweighting and subadditivity for small probabilities, exhibits mild insensitivity to changes in probabilities, and trends toward a positive intercept as μ approaches zero (though we have relatively little data for $\mu < 0.005$).

---

[40]The results are very similar under quadraticity, linearity, and unit slope. Under quadraticity, for example, the midpoints imply $\Omega(0.02) = 0.11$, $\Omega(0.05) = 0.16$, and $\Omega(0.10) = 0.22$, and the lower bounds imply $\Omega(0.02) = 0.07$, $\Omega(0.05) = 0.12$, and $\Omega(0.10) = 0.14$.

(A) Tversky-Kahneman PWF

(B) Prelec PWF

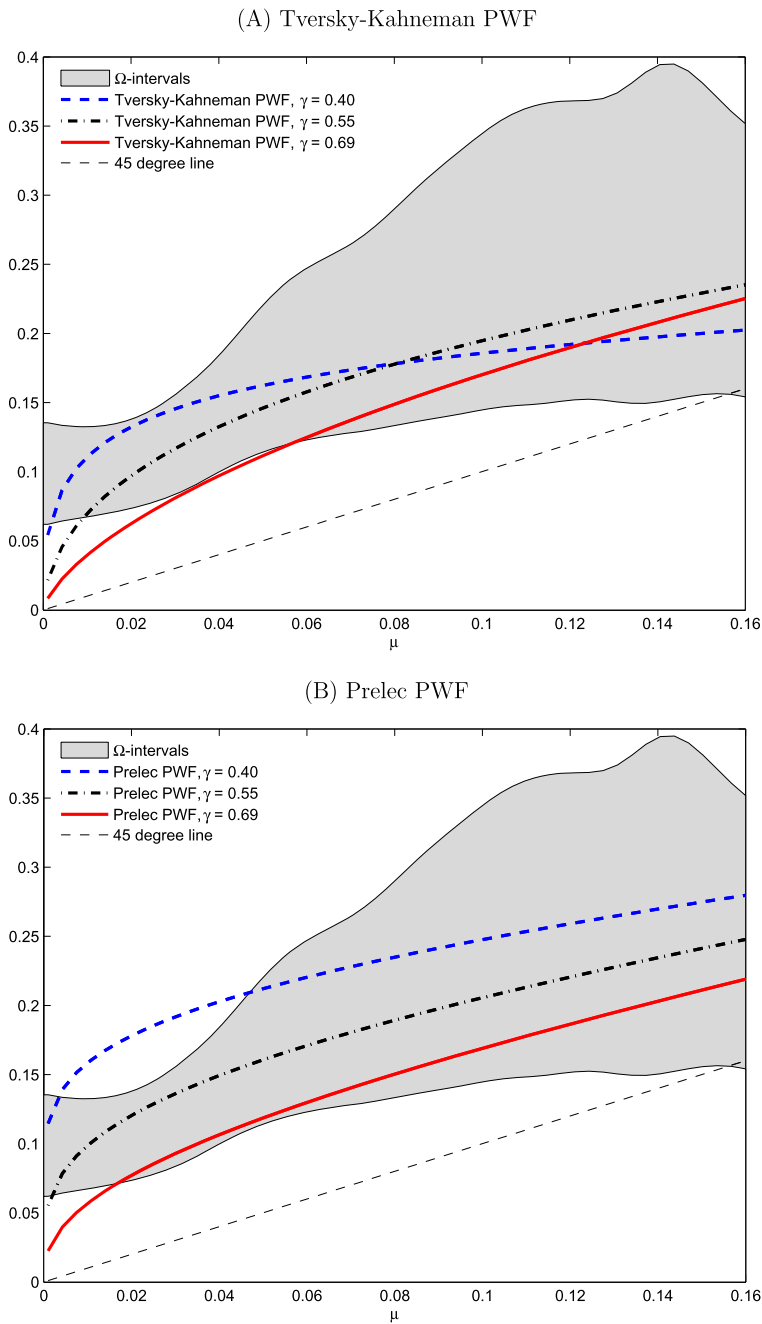FIGURE 4. Inference on Tversky–Kahneman and Prelec PWFs.

    Third, if we assume that $\Omega_i(\cdot)$ has a specific parametric form, we can utilize the Ω-intervals to conduct inference on the shape parameters. For example, Figure 4 super-imposes the one-parameter probability weighting functions suggested by Tversky and Kahneman (1992) (panel A) and Prelec (1998) (panel B), in each case for three parameter

values $\gamma \in \{0.40, 0.55, 0.69\}$, over the monotone Ω-intervals.[41] In the case of the Tversky–Kahneman PWF, the Ω-intervals favor $\gamma = 0.40$. In the case of the Prelec PWF, the Ω-intervals favor $\gamma = 0.55$.

5.3 Benefits and costs of shape restrictions on Ω

A key advantage of our approach is that it makes transparent the benefits and costs of restricting the shape of $\Omega_i(\cdot)$. The benefit of adding shape restrictions on $\Omega_i(\cdot)$ is a gain in precision—that is, they shrink the Ω-intervals. Table 1, column (b) reports for each shape restriction the average reduction in the size of the Ω-intervals due to the restriction. It reveals that each shape restriction yields large gains in precision. Monotonicity alone shrinks the Ω-intervals roughly by a quarter. Moving to linearity shrinks them roughly by two-fifths, and assuming unit slope shrinks them roughly by two-thirds. Of course, imposing zero intercept collapses the Ω-intervals to points. To put these gains in precision into perspective, we compare the size of the Ω-intervals under each shape restriction to the size of the 95 percent confidence bands on $\Omega(\cdot)$ that result from semi-nonparametric MLE of the model, as reported in BMOT. Table 1, column (c) reports, under each restriction, the ratio of the average size of the Ω-intervals to the average size of the BMOT 95 percent confidence bands on $\Omega(\cdot)$. Roughly, the ratio is three to one under monotonicity, two and a half to one under quadraticity, two to one under linearity, and one to one under unit slope.

    The cost of adding shape restrictions is a loss of model fit, that is, the model can rationalize the choices of fewer households. The percentage of rationalizable households that satisfy a shape restriction is a telling measure of model fit. However, it does not take into account the extent to which violating households fail to satisfy the restriction. For this reason, we introduce a second measure, tailored to our approach, that accounts for the extent of the violations. It is constructed as follows. For a given assumption A, we first assign each rationalizable household a number $Q_i \geq 0$. For households that satisfy A for some $r_i \in [0, 0.0108]$, we set $Q_i = 0$. For households that violate A, we fix $r_i = 0$ and set Q_i equal to the solution to the problem

$$\text{minimize} \quad \omega_{iL} + \omega_{iM} + \omega_{iH}$$

$$\text{such that} \quad i \text{ satisfies } A \text{ with } \mathcal{I}_{ij} = [\text{LB}_{ij} - \underline{\omega}_{ij}, \text{UB}_{ij} + \overline{\omega}_{ij}] \text{ for all } j \in \{L, M, H\},$$

$$\text{where } \omega_{ij} = \underline{\omega}_{ij} + \overline{\omega}_{ij} \text{ and } \underline{\omega}_{ij}, \overline{\omega}_{ij} \geq 0 \text{ for all } j \in \{L, M, H\}.$$

Intuitively, we take the household's Ω-intervals at $r_i = 0$, imagine perturbing them such that the household satisfies the assumption, and set Q_i equal to the minimum required perturbation.[42] We then take the average value of Q_i among the rationalizable households: $\overline{Q} = \sum_i Q_i / 3629$. Table 1, column (d) reports $\overline{Q}$ for each shape restriction. The

---

[41]The Tversky–Kahneman PWF is $\Omega(\mu) = \mu^\gamma / [\mu^\gamma + (1 - \mu)^\gamma]^{1/\gamma}$ with $0 < \gamma \leq 1$. The Prelec PWF is $\Omega(\mu) = \exp[-(-\ln \mu)^\gamma]$ with $\gamma > 0$.

[42]We note that fixing $r_i = 0$ is conservative: it forces $\Omega_i(\cdot)$ to do all the work of rationalizing the household's choices. Of course, one could fix r_i at any plausible value. For example, one could fix r_i at the relevant estimate reported by BMOT (0.00049). This would yield very similar results.

minimum perturbations that would be required for every rationalizable household to satisfy monotonicity, quadraticity, linearity, and unit slope are relatively small—ranging from less than one percentage point on average in the case of monotonicity to less than four percentage points on average in the case of unit slope. By comparison, the minimum perturbations that would be required for every rationalizable household to satisfy zero intercept are relatively large—roughly 12 percentage points on average.

Although the foregoing cost–benefit accounting does not readily lend itself to a marginal analysis that selects the optimal shape restriction on $\Omega_i(\cdot)$, it unequivocally evidences the importance of probability distortions in general. On the one hand, if we do not allow for probability distortions, the model not only fails to explain the choices of three in five rationalizable households, it fails badly according to the $\overline{Q}$ statistic. On the other hand, if we allow for probability distortions, we greatly improve the model's fit (even permitting only linear, unit slope distortions) and still achieve fairly tight bounds (even imposing only monotonicity).

5.4 *More on $\overline{Q}$*

There is a noteworthy kinship between the $\overline{Q}$ statistic and the efficiency index developed by Afriat (1967, 1972) and Varian (1990, 1993). The Afriat–Varian efficiency index measures the extent to which choice data violate the Generalized Axiom of Revealed Preference (GARP); in other words, it measures the extent to which choice data are inconsistent with utility maximization (with a concave utility function). The $\overline{Q}$ statistic measures the extent to which choice data are inconsistent with expected utility maximization as generalized by the probability distortion model (with a CARA utility function and a given shape restriction on $\Omega_i(\cdot)$). Of course, the $\overline{Q}$ statistic and the Afriat–Varian efficiency index have different units of measurement: the $\overline{Q}$ statistic is denominated in probability units whereas the Afriat–Varian efficiency index is denominated in budget units (i.e., wealth). However, we can readily translate the $\overline{Q}$ statistic into wealth. For example, suppose that we want to compute the expected wealth loss associated with the deviations from monotonicity among the rationalizable households in our sample. Equipped with $\mathcal{I}_{ij}$ and ω_{ij} (as defined in Section 5.3) for all i and j, we can compute, for each household and coverage, the smallest difference between the expected value of the household's actual choice and the expected value of the household's predicted choice under the model with a monotone Ω function.[43] Doing this, we find that among the rationalizable households that violate monotonicity, the average expected wealth loss is $23 (or a 15.1 percent average expected utility loss).

Furthermore, we can apply a similar logic to measure the extent of the plausibility violations among the nonrationalizable households in our sample. As noted above, virtually every nonrationalizable household chose an auto collision deductible of $200, which is an implausible choice given the menu of prices. Per our model, we can ask, "What is the smallest deviation from rationality that can explain this choice?" The $200 deductible would be the smallest deviation possible if the household's optimal deductible choice was either $100 or $250. Thus, the $\Omega(\mu)$ at which the expected utility cost of the

---

[43]For instance, suppose for a given household i that $\underline{\omega}_{iL} > 0$, $\underline{\omega}_{ij} = 0$ for $j = M, H$, and $\overline{\omega}_{ij} = 0$ for $j = L, M, H$. The household's wealth loss is given by $(-p(d_{iL}^*) - \mathrm{LB}_{iL} d_{iL}^*) - \max_{d_{iL}} \{-p(d_{iL}^*) - (\mathrm{LB}_{iL} - \underline{w}_{iL}) d_{iL}\}$.

deviation is the smallest is the $\Omega(\mu)$ at which the household is indifferent between these deductibles, that is, $\Omega(\mu) = (p(250) - p(100))/(250 - 100)$.[44] Equipped with this $\Omega(\mu)$, we can measure the differences in expected utility between choosing a deductible of \$100 or \$250, on the one hand, and \$200, on the other.[45] Doing this, we find that the average expected utility loss is \$9.58 (or 4.7 percent).

In sum, one can use the $\overline{Q}$ statistic to measure the wealth/utility cost of deviating from a given model of risky choice. Moreover, for a researcher estimating a model of risky choice using parametric methods, the $\overline{Q}$ statistic also provides information that may be useful in guiding the econometric specification. Under the probability distortion model, for example, violations of plausibility suggest the need for choice noise (i.e., disturbances to the von Neumann–Morgenstern (vNM) utility function), while violations of monotonicity suggest the need for choice noise or preference noise (e.g., disturbances to the Ω function or its parameters), and in either case the $\overline{Q}$ statistic is informative about the presence and amount of such noise.

6. Classification

In the previous section, we utilize the rationalizable households' Ω-intervals—which come out of stability and revealed preference—to begin to draw conclusions about the nature of their risk preferences. In this section, we use the Ω-intervals to classify the rationalizable households into preference types, each of which corresponds to a special case of the model. In addition to providing evidence about the distribution of preference types, the classification results reinforce the importance of probability distortions. The results also reveal, however, that certain forms of probability distortions have more purchase within our application than others. In particular, the results strongly favor the unit slope form, which we view as a parsimonious representation of KT probability weighting, over the forms that correspond to Gul disappointment aversion and KR loss aversion.

6.1 *Concave utility and probability distortions*

As a first step, we consider the following set of preference types.

- Expected value theory: when $r = 0$ and $\Omega(\mu) = \mu$.
- Objective expected utility theory: when $r \in [0, 0.0108]$ and $\Omega(\mu) = \mu$.
- Yaari's dual theory: when $r = 0$ and $\Omega(\mu)$ is a PWF.
- Rank-dependent expected utility theory: when $r \in [0, 0.0108]$ and $\Omega(\mu)$ is a PWF.

Expected value theory allows for neither concave utility nor probability distortions. Objective expected utility theory allows for concave utility but not probability distortions. Yaari's dual theory allows for monotone probability distortions but not concave utility. Rank-dependent expected utility theory allows for concave utility and monotone probability distortions.

---

[44]Here we assume that $r = 0$. The utility cost is the smallest under linear utility.

[45]It is straightforward to see that with linear utility and this $\Omega(\mu)$, this utility difference is the same for both the \$100 versus \$200 comparison and the \$250 versus \$200 comparison.

TABLE 2. Classification. Rationalizable subsample (3629 households).

Preference Type	Lower Limit	Upper Limit	Imbens–Manski/Stoye 95 Percent Confidence Interval	
Panel A				
Expected value theory [$r = 0$ & $\Omega(\mu) = \mu$]	0.00	2.95	0.00	2.96
Objective expected utility theory [$0 \leq r \leq 0.0108$ & $\Omega(\mu) = \mu$]	1.02	39.63	1.01	39.65
Yaari's dual theory [$r = 0$ & $\Omega(\mu)$ is a PWF]	41.44	80.05	41.42	80.07
Rank-dependent EUT [$0 \leq r \leq 0.0108$ & $\Omega(\mu)$ is a PWF]	3.78	84.84	3.77	84.86
Panel B				
Expected value theory [$r = 0$ & $\Omega(\mu) = \mu$]	0.00	2.95	0.00	2.96
Objective expected utility theory [$0 \leq r \leq 0.0108$ & $\Omega(\mu) = \mu$]	13.39	39.63	13.38	39.65
Gul's disappointment aversion model with $r = 0$	2.92	29.15	2.91	29.17
[$\Omega(\mu) = \mu(1 + \beta)/(1 + \beta\mu)$]				
Gul's disappointment aversion model	0.44	42.99	0.44	43.01
[$0 \leq r \leq 0.0108$ & $\Omega(\mu) = \mu(1 + \beta)/(1 + \beta\mu)$]				
Panel C				
Expected value theory [$r = 0$ & $\Omega(\mu) = \mu$]	0.00	2.95	0.00	2.96
Objective expected utility theory [$0 \leq r \leq 0.0108$ & $\Omega(\mu) = \mu$]	15.98	39.63	15.97	39.65
KR's loss aversion model with $r = 0$ [$\Omega(\mu) = \mu + \Lambda(1 - \mu)\mu$]	1.74	25.38	1.73	25.40
KR's loss aversion model	0.85	42.22	0.85	42.24
[$0 \leq r \leq 0.0108$ & $\Omega(\mu) = \mu + \Lambda(1 - \mu)\mu$]				
Panel D				
Expected value theory [$r = 0$ & $\Omega(\mu) = \mu$]	0.00	2.95	0.00	2.96
Objective expected utility theory [$0 \leq r \leq 0.0108$ & $\Omega(\mu) = \mu$]	10.50	39.63	10.48	39.65
Unit slope distortions model with $r = 0$ [$\Omega(\mu) = a + \mu$]	18.32	47.45	18.31	47.47
Unit slope distortions model [$0 \leq r \leq 0.0108$ & $\Omega(\mu) = a + \mu$]	3.64	61.59	3.63	61.61
Panel E				
Objective expected utility theory [$0 \leq r \leq 0.0108$ & $\Omega(\mu) = \mu$]	4.60	39.63	4.59	39.65
Gul's disappointment aversion model with $r = 0$	0.17	29.15	0.16	29.17
[$\Omega(\mu) = \mu(1 + \beta)/(1 + \beta\mu)$]				
KR's loss aversion model with $r = 0$ [$\Omega(\mu) = \mu + \Lambda(1 - \mu)\mu$]	0.11	25.38	0.11	25.40
Unit slope distortions model with $r = 0$ [$\Omega(\mu) = a + \mu$]	15.90	47.45	15.88	47.47
Percentage consistent with all four models	18.57			
Percentage consistent with at least one model	58.58			

Note: For each preference type, the upper limit is the percentage of rationalizable households with Ω-intevals that are consistent with that preference type. For a core preference type (italicized), the lower limit is the percentage of rationalizable households with Ω-intervals that are consistent with that preference type but are inconsistent with the other core preference type. For a noncore preference type, the lower limit is the percentage of rationalizable households with Ω-intervals that are consistent with that preference type but are inconsistent with both core preference types. Imbens–Manski/Stoye confidence intervals uniformly cover each element of the bound (i.e., [lower limit, upper limit]) on each preference type, with asymptotic probability 95 percent.

Table 2, panel A reports for each preference type in the set the percentage of rationalizable households with Ω-intervals that are consistent with that preference type. We refer to this percentage as the *upper limit* for the preference type, because some rationalizable households have Ω-intervals that are consistent with more than one preference type. This occurs for two reasons. First, there are overlaps among the preference types.

In particular, expected value theory is nested by each of the others; objective expected utility theory and Yaari's dual theory are nested by rank-dependent expected utility and share a degenerate special case (expected value theory); and rank-dependent expected utility theory nests each of the others. Second, some rationalizable households have Ω-intervals that are consistent with multiple nonnested preference types, even excluding common special cases. Specifically, some rationalizable households are consistent with both objective expected utility theory and Yaari's dual theory (the only nonnested pair in the set of preference types), even excluding those that are consistent with expected value theory.

Looking at the upper limits, we find that 3.0 percent of rationalizable households have Ω-intervals that are consistent with expected value theory. These households require neither concave utility nor probability distortions to explain their choices. Furthermore, we find that 39.6 percent of rationalizable households have Ω-intervals that are consistent with objective expected utility theory, 80.1 percent have Ω-intervals that are consistent with Yaari's dual theory, and 84.8 percent are consistent with rank-dependent expected utility.[46] In other words, roughly two in five rationalizable households make choices that are consistent with a model with concave utility and nondistorted probabilities, four in five make choices that are consistent with a model with linear utility and monotone probability distortions, and five in six are consistent with a model with concave utility and monotone probability distortions. Stated another way, concave utility alone is sufficient to explain the choices of roughly two in five rationalizable households; monotone probability distortions alone are sufficient to explain the choices of four in five rationalizable households; and concave utility and monotone probability distortions together can explain the choices of five in six rationalizable households.

Table 2, panel A also reports the *lower limit* for each preference type in the set. To understand the lower limits, it is helpful to distinguish between (i) the "core" nonnested preference types in the set (objective expected utility theory and Yaari's dual theory) and (ii) the "noncore" preference types in the set, which are either a degenerate special case of both core types (expected value theory) or a generalization of both core types (rank-dependent expected utility theory). For a core preference type, the lower limit is the percentage of rationalizable households with Ω-intervals that are consistent with that preference type but are inconsistent with the other core preference type. For a noncore preference type, the lower limit is the percentage of rationalizable households with Ω-intervals that are consistent with that preference type but are inconsistent with both core preference types.

Turning to the lower limits, we find that 1.0 percent of rationalizable households have Ω-intervals that are consistent with objective expected utility theory but not with Yaari's dual theory. In other words, we find that for one in 100 rationalizable households, (i) they require concave utility to explain their choices and (ii) their choices cannot be explained solely by monotone probability distortions. In addition, we find that 41.4 percent of rationalizable households have Ω-intervals that are consistent with Yaari's dual

---

[46]Note that the rationalizable households with Ω-intervals that are consistent with rank-dependent expected utility theory are the same households that satisfy monotonicity in Section 5.

theory but not with objective expected utility theory. That is, we find that for more than
two in five rationalizable households, (i) they require monotone probability distortions
to explain their choices and (ii) their choices cannot be explained solely by concave util-
ity. Together with the lower limit on rank-dependent expected utility theory (3.8 per-
cent),[47] which implies that less than one in 25 rationalizable households requires both
concave utility and monotone probability distortions to explain their choices, these
lower limits imply that concave utility is necessary to explain the choices of less than
one in 20 rationalizable households (4.8 percent), whereas monotone probability distor-
tions are necessary to explain the choices of nearly one in two rationalizable households
(45.2 percent).[48]

    The results clearly evince the importance of probability distortions. The upper limits
imply that probability distortions alone can explain the choices of four in five rational-
izable households, twice as many as can be explained by concave utility alone. At the
same time, the lower limits imply that probability distortions are required to explain
the choices of nearly half of the rationalizable households, almost 10 times as many as
require concave utility to explain their choices. In short, the results indicate that the
marginal contribution of probability distortions is high. Even allowing for concave util-
ity, the gain in explanatory power from allowing for probability distortions is large. By
contrast, the marginal contribution of concave utility is low. Once we allow for probabil-
ity distortions, the gain in explanatory power from allowing for concave utility is small.

6.2 *Gul disappointment aversion, KR loss aversion, and unit slope distortions*

As noted above, the probability distortion function can capture a wide range of differ-
ent behaviors, depending on the underlying model. We focus on three models, each of
which implies a specific restriction on the probability distortion function: Gul disap-
pointment aversion, which implies $\Omega(\mu) = \mu(1+\beta)/(1+\beta\mu)$; KR loss aversion, which
implies $\Omega(\mu) = \mu + \Lambda(1-\mu)\mu$; and unit slope distortions, which imply $\Omega(\mu) = a + \mu$, and

---

[47] The lower limit for expected value theory is zero by definition, because expected value theory is a de-
generate special case of both core preference types.

[48] Table 5 also reports Imbens–Manski–Stoye confidence intervals that uniformly cover each element of
the bound on each preference type, with asymptotic probability $1 - \alpha$ (Imbens and Manski (2004), Stoye
(2009)). Formally, let the bound on each type be $[\widehat{\theta}_l, \widehat{\theta}_u]$, where $\widehat{\theta}_l$ is the lower limit and $\widehat{\theta}_u$ is the upper
limit. Observe that each is a frequency estimator, so the estimator of each of their variances is $\widehat{\sigma}_j^2 = \widehat{\theta}_j(1 - \widehat{\theta}_j)/n$, $j = l, u$. Then a confidence interval that is asymptotically valid uniformly over a large class of data
generating processes is $[\widehat{\theta}_l - (c_\alpha \widehat{\sigma}_l/\sqrt{n}), \widehat{\theta}_u + (c_\alpha \widehat{\sigma}_u/\sqrt{n})]$, where c_α solves

$$\Phi\left(c_\alpha + \frac{\sqrt{n}(\widehat{\theta}_u - \widehat{\theta}_l)}{\max\{\widehat{\sigma}_u, \widehat{\sigma}_l\}}\right) - \Phi(-c_\alpha) = 1 - \alpha,$$

and where Φ is the standard normal cumulative distribution function. Validity follows because Imbens–
Manski–Stoye assumptions are satisfied in our context, in particular because $P(\widehat{\theta}_u \geq \widehat{\theta}_l) = 1$ by construction
and hence uniformly, and therefore Lemma 3 in Stoye (2009) applies. In the case of expected value theory,
in which the lower limit is zero by construction, our confidence intervals are obtained using a one-sided
confidence interval for the upper limit.

which we view as a parsimonious representation of KT probability weighting.[49] Note that all three models are parameterized with a single parameter.

As a second step, we reclassify households over three additional sets of preference types. The first set comprises objective expected utility theory and Gul's disappointment aversion model with $r = 0$, as the core preference types, plus expected value theory and Gul's model with $r \in [0, 0.0108]$, as the noncore preference types. The second set comprises objective expected utility theory and KR's loss aversion model with $r = 0$, as the core preference types, plus expected value theory and KR's model with $r \in [0, 0.0108]$, as the noncore preference types. The third set comprises objective expected utility theory and the unit slope distortions model with $r = 0$, as the core preference types, plus expected value theory and the unit slope model with $r \in [0, 0.0108]$, as the noncore preference types.

The results are reported in Table 2, panels B, C, and D, respectively. The results in panels B and C imply that restricting $\Omega(\cdot)$ to conform to Gul's disappointment aversion model or KR's loss aversion model has two related effects: it substantially reduces the model's explanatory potential and it nearly eliminates the marginal contribution of probability distortions. As before, the upper limit for expected value theory is 3.0 percent, and allowing for concave utility (i.e., moving to the objective expected utility model) raises the upper limit to 39.6 percent. However, allowing for Gul disappointment aversion or KR loss aversion, as the case may be, raises the upper limit only to 29.2 percent or 25.4 percent, respectively. Indeed, even if we allow for (i) concave utility and Gul disappointment aversion or (ii) concave utility and KR loss aversion, as the case may be, it raises the upper limit only to 43.0 percent or 42.2 percent, respectively. What is more, the lower limits imply that only 3.4 percent of rationalizable households require Gul disappointment aversion to explain their choices, and only 2.6 percent require KR loss aversion. In short, the results imply that the marginal contribution of these behaviors/models is small.

By contrast, the results in panel D imply that unit slope distortions enhance the model's explanatory potential and make a substantial marginal contribution. The upper limits imply that unit slope distortions alone can explain the choices of nearly half (47.5 percent) of rationalizable households, 20 percent more than can be explained by concave utility alone (39.6 percent), and that unit slope distortions in combination with concave utility can explain the choices of more than three in five (61.6 percent) rationalizable households. At the same time, the lower limits imply that unit slope distortions are necessary to explain the choices of nearly a fourth (22.0 percent) of rationalizable households, almost 60 percent more than require concave utility to explain their choices (14.1 percent).

As a final step, we reclassify households over a set that includes objective expected utility theory and each of the Gul, KR, and unit slope models with $r = 0$, as the core preference types, but includes no noncore preference types. The results are reported in Table 2, panel E. In addition to reporting the upper and lower limits, panel E also reports

---

[49]It is straightforward to show that the unit slope form with a positive intercept satisfies overweighting and subadditivity, as well as subproportionality. However, it does not satisfy subcertainty.

the percentage of rationalizable households that are consistent with all four models and the percentage that are consistent with at least one model. Again, the results point to unit slope probability distortions as the behavior/model with the greatest explanatory potential and marginal contribution. First, while the four models collectively can explain the choices of nearly six in 10 rationalizable households, the unit slope model alone can explain the choices of nearly five in 10. Second, unit slope distortions are necessary to explain the choices of nearly one in six rationalizable households, while fewer than one in 20 require concave utility to explain their choices, and fewer than one in 100 require Gul disappointment aversion or KR loss aversion to explain their choices.

6.3 *Power of revealed preference test*

A question that arises in interpreting the results of this section (as well as the previous section) concerns the power of our revealed preference test of shape restrictions on $\Omega_i(\cdot)$. Of course, the key challenge in measuring the power of any revealed preference test lies in selecting the alternative model of choice. After all, the power of a test is a function of the alternative, and there always exist alternatives against which the test will have low power (Blundell, Browning, and Crawford (2003)). In our setting, for example, a simple rule of always selecting the minimum deductible option is indistinguishable from expected utility maximization (with or without probability distortions) with extreme risk aversion, and hence our test of zero intercept (or any other shape restriction) would have no power against this alternative.

In influential work on the topic of the power of revealed preference tests, Bronars (1987) proposes uniform random choice as a general alternative to a null of optimizing behavior.[50] Adopting Bronars' alternative, we perform Monte Carlo simulations to estimate the probability of satisfying various shape restrictions $\Omega_i(\cdot)$. More specifically, we generate 200 simulated data sets, each comprising 3629 observations of three deductibles choices (one for each coverage), where each choice is drawn randomly from a uniform distribution over the coverage-specific options. We then compute the mean pass rate of our test across the simulated data sets under five shape restrictions: monotonicity, unit slope, KR loss aversion, Gul disappointment aversion, and zero intercept. The results are reported in Table 3, column (b).

Computing these pass rates, however, provides only a first step toward the end goal of assessing the success of the underlying model of choice. To that end, Beatty and Crawford (2011) combine Bronars' approach with Selten's (1991) measure of predictive success of area theories (i.e., theories that predict a subset of all possible outcomes) to fashion a success measure for models of choice. Essentially, they measure success by the difference between the test's pass rate under the null and the pass rate under Bronars' alternative. The intuition is that "a model should be counted as more successful in situations in which we observe both good pass rates and demanding restrictions" (Beatty and Crawford (2011, p. 2785)). After all, if the pass rate under Bronar's alternative is high,

---

[50]Bronars credits Becker (1962) for the basic idea. For a recent application of Bronars' approach, see, for example, Choi, Fisman, Gale, and Kariv (2007). For a thorough discussion of the topic, including a review of Bronars' approach and suggestions for alternative approaches, see Andreoni, Gillen, and Harbaugh (2013).

TABLE 3. Power of revealed preference test. Rationalizable subsample (3629 households).

Shape Restriction	(a) Percentage of Households Satisfying Restriction	(b) Percentage of Households Satisfying Restriction	(c)		(d)
	Actual	Uniform Random Choice	95 Percent Confidence Interval		Beatty–Crawford Success Measure
Monotonicity	84.8	38.8	37.2	40.2	46.1
Unit slope	61.6	13.8	12.9	14.9	47.8
KR loss aversion	42.2	11.3	10.3	12.1	30.9
Gul disappointment aversion	43.0	11.4	10.6	12.3	31.6
Zero intercept	39.6	6.3	5.6	7.1	33.4

Note: Column (a) reports results for the actual data. Column (b) reports means across 200 simulated data sets, each comprising 3629 observations of three deductible choices (one for each coverage), where each choice is drawn randomly from a uniform distribution over the coverage-specific options. Column (c) reports 95 percent confidence intervals for the means reported in column (b). The Beatty–Crawford success measure is the difference between columns (a) and (b).

then the underlying model is not very demanding and consequently the revealed preference test reveals very little.[51]

Table 3, column (d) reports the Beatty–Crawford measure of the success of the model under monotonicity, unit slope, KR loss aversion, Gul disappointment aversion, and zero intercept. The results indicate that a model with monotone probability distortions is substantially more successful than a model with no probability distortions (zero intercept). What is more, the results strongly favor unit slope distortions over those implied by KR loss aversion or Gul disappointment aversion. Indeed, the unit slope model slightly outperforms even the general monotone model, while the KR and Gul models slightly underperform the zero intercept model.[52]

7. POINT ESTIMATION

The classification results provide evidence about the extent and nature of preference heterogeneity. In many areas of research, however, economists study models that abstract from heterogeneity in preferences (e.g., representative agent models) and seek a single parameterization that best fits the data. In this section, we show how one can utilize the Ω-intervals to point estimate the probability distortion function.

Intuitively, we find the Ω function that comes closest (in a sense we make precise) to the monotone households' Ω-intervals. More specifically, we find the single Ω function

---

[51] In the extreme, if the pass rate under Bronar's alternative is 1, then the underlying model has no empirical content, as any choice is consistent with the model, and the revealed preference test reveals nothing.

[52] Dean and Martin (forthcoming) propose a modification of Beatty and Crawford's approach that, in our application, calls for replacing Bronars' alternative of uniform random choice in each coverage with an alternative of random choice according to the marginal empirical distribution of choices in each coverage. (Andreoni, Gillen, and Harbaugh (2013) propose a similar approach.) Naturally, under Dean and Martin's alternative, which is closer to the null, the pass rates are higher and the Beatty–Crawford statistics are lower for each shape restriction. Nevertheless, the relative success of the shape restrictions are roughly the same. See Table S4 in the Supplement.

TABLE 4. Minimum distance Ω. Monotone subsample (3079 households).

Parameter	Estimate	Std. Err.
a	0.055*	0.003
b	2.22*	0.14
c	−14.88*	1.95
d	43.58*	8.65
e	−54.45*	18.59
f	25.73	25.49

Note: 9237 observations (3079 × 3). *Significant at the 5 percent level.

that minimizes the average distance between these Ω-intervals and the function. We then assess how well this minimum distance Ω fits the data. As before, we consider two notions of fit. The first is the percentage of monotone households that the model can rationalize when equipped with the minimum distance Ω. The second is the average distance between the minimum distance Ω and their Ω-intervals.

7.1 *Minimum distance Ω*

We estimate the best linear predictor,[53]

$$\widetilde{\Omega}(\mu_{ij}) = a + b\mu_{ij} + c(\mu_{ij})^2 + d(\mu_{ij})^3 + e(\mu_{ij})^4 + f(\mu_{ij})^5,$$

by finding the value of $\boldsymbol{\theta} \equiv (a, b, c, d, e, f)$ that minimizes, over every choice j of every monotone household i, the average Euclidean distance between the point $\widetilde{\Omega}(\mu_{ij})$ and the Ω-interval $\mathcal{I}_{ij}$.[54] This distance is zero if $\widetilde{\Omega}(\mu_{ij})$ is contained in $\mathcal{I}_{ij}$; otherwise, it equals the Euclidean distance to the nearest bound of $\mathcal{I}_{ij}$. As before, we pin down $\mathcal{I}_{ij}$ by fixing the household's coefficient of absolute risk aversion r_i at its minimum plausible value (here, the minimum plausible value under which the household satisfies monotonicity). In Section F of the Supplement, we prove that under mild conditions (satisfied in our data) the parameter vector $\boldsymbol{\theta}$ is point identified, and we establish the consistency and asymptotic normality of our sample analog estimator. We also demonstrate that its critical values can be consistently approximated by nonparametric bootstrap.

The results are reported in Table 4 and depicted in Figure 5. The minimum distance Ω, which is monotone on the relevant range,[55] exhibits substantial overweighting of small probabilities. For example, it implies that claim probabilities of 2 percent, 5 per-

---

[53]We specify a fifth-degree polynomial because it is flexible enough to closely approximate most monotone functions. However, we do not restrict the parameters of the function to impose monotonicity. As it turns out, the estimated function is monotone on the relevant range (see below).

[54]For a given point $t \in \mathbb{R}$ and interval $T = [\tau_L, \tau_U]$, the Euclidean distance between t and T is given by $d(t, T) = \inf_{\tau \in T} |t - \tau| = \max\{(\tau_L - t)_+, (t - \tau_U)_+\}$, where $(z)_+ = \max(0, z)$.

[55]Specifically, it is increasing between 0 and 0.16, wherein lie 98.1 percent of the claim probabilities in the rationalizable subsample.
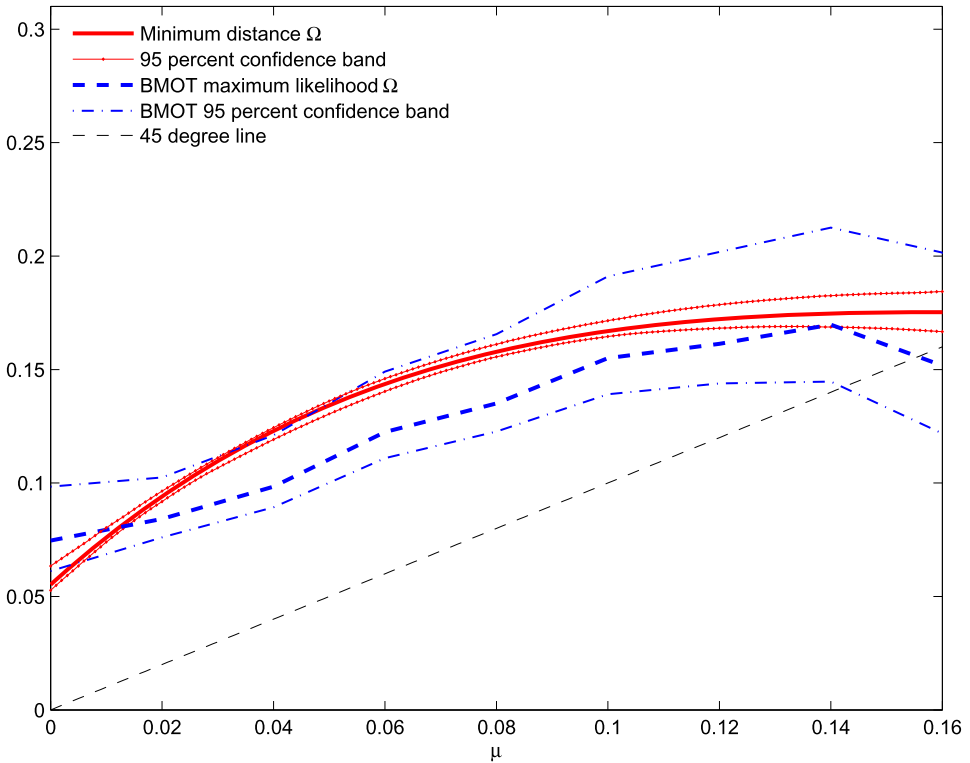
FIGURE 5. Minimum distance Ω.

cent, and 10 percent are distorted (overweighted) to 9 percent, 14 percent, and 17 percent, respectively. We note that the results are very similar if we instead specify a lower order polynomial, including even a first-degree function; see Section F of the Supplement.

We emphasize that the minimum distance Ω is obtained without making parametric assumptions about the distribution of unobserved heterogeneity in preferences. Instead, it relies only on the economic model, including the shape restrictions on the utility and probability distortion functions, and the Ω-intervals, which come out of stability and revealed preference, to recover the probability distortion function that best fits the data, that is, best describes the probability distortions of the average (representative) monotone household.

It is useful to contrast this minimum distance Ω with the "maximum likelihood" Ω estimated by BMOT, which was obtained by sieve MLE of a semi-nonparametric econometric model that assumes random utility with additively separable, independent type 1 extreme value distributed choice noise. As Figure 5 shows, the minimum distance Ω and the maximum likelihood Ω are remarkably similar. Because they are obtained by two very different methods, they act as mutual robustness checks and serve to reinforce each other and increase confidence in their common result.
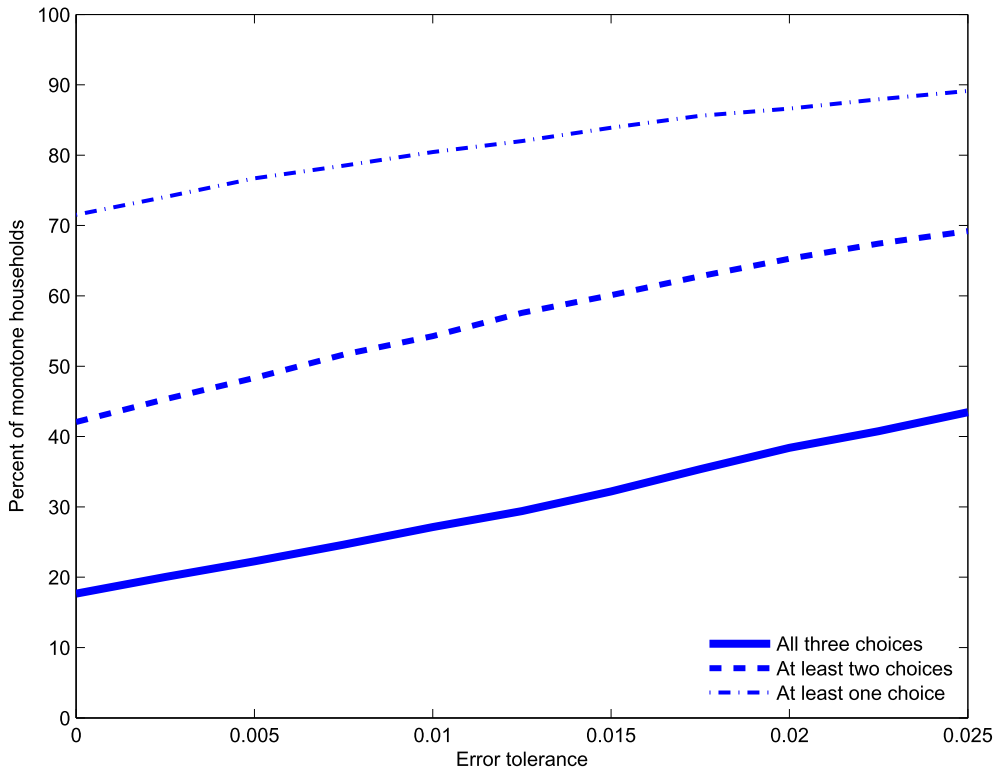
FIGURE 6. Percentage of monotone households that the model can rationalize when equipped with the minimum distance Ω.

7.2 *Model fit*

Next, we assess model fit given the minimum distance Ω. First, we consider the percentage of monotone households that the model can rationalize when equipped with the minimum distance Ω. Figure 6 presents the results. With zero tolerance for error—that is, if we require zero distance between $\widetilde{\Omega}$ and the Ω-interval—we find that the model can rationalize all three choices of 18 percent of monotone households, at least two choices of 42 percent of monotone households, and at least one choice of 72 percent of monotone households. In other words, a single probability distortion function can rationalize all three choices of nearly one in five monotone households, at least two choices of more than two in five monotone households, and at least one choice of more than seven in 10 monotone households. If we tolerate some error, the percentages increase quite rapidly. With a tolerance of 2.0 percentage points, for instance, the model can rationalize all three choices of nearly two in five monotone households, at least two choices of more than three in five monotone households, and at least one choice of nearly nine in 10 monotone households.

Second, we consider the average distance between the minimum distance Ω and the monotone households' Ω-intervals. This is akin to the $\overline{Q}$ statistic introduced in Section 5. At the minimizer $\boldsymbol{\theta^*}$, the average distance between $\widetilde{\Omega}(\mu_{ij})$ and $\mathcal{I}_{ij}$ is 2.70 percent-

age points. To gauge the magnitude of this distance, we compare it to the average distance between μ_{ij} and $\mathcal{I}_{ij}$, which is 5.82 percentage points. That is to say, on average, there is a 5.82 percentage point "gap" that is unexplained if we restrict $\Omega(\mu_{ij}) = \mu_{ij}$. The minimum distance Ω can explain 46 percent of this gap. The remainder is attributed to heterogeneity in probability distortions.

The last point is worth emphasizing. Suppose we want to measure the heterogeneity in probability distortions among the monotone households (or any other subset of the rationalizable households). Given the specification of $\widetilde{\Omega}(\mu_{ij})$, the distances between $\widetilde{\Omega}(\mu_{ij})$ and $\mathcal{I}_{ij}$ give us precisely the lower bound on the degree of heterogeneity, and it is obtained without making assumptions about the nature of the heterogeneity. By contrast, if we estimate the model by maximum likelihood or other parametric methods, the residuals between the model and the data depend, inextricably and opaquely, on the parametric assumptions, and consequently the exact nature of the relationship between these residuals and the degree of heterogeneity is obscured.

8. Rank correlation of choices

In this section, we shift gears to address a puzzle in the recent literature on the stability of risk preferences. On the one hand, Barseghyan, Prince, and Teitelbaum (2011) (using data on choices in three insurance domains) and Einav et al. (2012) (using data on choices in five insurance domains and one investment domain) provide evidence that, within an expected utility framework, people do not exhibit a stable degree of risk aversion across contexts. At the same time, Einav et al. (2012) provide evidence that people's risky choices are rank correlated across contexts, implying that there exists an important domain-general component of risk preferences.[56] We show that one can resolve this puzzle with stable probability distortions.

More specifically, we demonstrate that, in our data, there is a close connection between rank correlation of choices and stability of risk preferences under the probability distortion model. First, we document that the rationalizable households' deductible choices are rank correlated across lines of coverage. Table 5, column (a) reports, for the full subsample of 3629 rationalizable households, the pairwise Spearman rank correlations of the households' deductible choices in auto collision, auto comprehensive, and home. The rank correlations are positive and range from 0.285 to 0.490, and each is statistically significant at the 1 percent level. Notably, the results are remarkably similar to those of Einav et al. (2012), in which the rank correlations between the five insurance domains range from 0.174 to 0.400.

Next, we show that it is the rationalizable households with stable risk preferences under the probability distortion model who are driving these rank correlations. Table 5,

---

[56]More specifically, Einav et al. (2012) investigate the stability in ranking across domains of an individual's willingness to bear risk relative to his or her peers. They rank by risk the options within each domain and compute the pairwise rank correlations in the individuals' choices across domains. They find that an individual's choice in every domain is positively correlated to some extent with his or her choice in every other domain. For a more detailed discussion of Barseghyan, Prince, and Teitelbaum (2011) and Einav et al. (2012), see Teitelbaum (2015).

TABLE 5. Rank correlation of deductible choices. Rationalizable subsample (3629 households).

	(a)	(b)	
	All Rationalizable Households ($N = 3629$)	Rationalizable Households That Satisfy Monotonicity ($N = 3079$)	Rationalizable Households That Violate Monotonicity ($N = 550$)
Auto collision and auto comprehensive	0.490*	0.553*	0.335*
Auto collision and home	0.290*	0.363*	−0.019
Auto comprehensive and home	0.285*	0.352*	0.029

Note: Each cell reports a pairwise Spearman rank correlation coefficient. *Significant at the 1 percent level.

column (b) breaks out the rank correlations for the 3079 rationalizable households that satisfy monotonicity and the 550 rationalizable households that violate monotonicity. Relative to the overall rank correlations, the rank correlations are stronger among households that satisfy monotonicity and weaker among households that violate monotonicity. Indeed, among households that violate monotonicity, the rank correlations between auto collision and home and between auto comprehensive and home are statistically indistinguishable from zero at conventional levels of significance, suggesting that the corresponding overall rank correlations are being driven entirely by the households that satisfy monotonicity.[57]

In sum, we find that stable probability distortions are the domain-general component of risk preferences that account for the rank correlation of choices across contexts in our data. The choices of the rationalizable households that satisfy stability under the probability distortion model are strongly rank correlated, while the choices of the rationalizable households that violate stability are weakly rank correlated, if at all.

9. ASYMMETRIC INFORMATION

In Section I of the Supplement, we address the concern that the asymmetric information twins—moral hazard (unobserved action) and adverse selection (unobserved type)—may be biasing our claim rate estimates and hence our results. With respect to moral hazard, we consider both ex ante and ex post moral hazard, and we conclude that neither is a significant issue in our data. With respect to adverse selection, we consider two possibilities—(i) there is heterogeneity in claim risk that is observed by the households but unobserved by the econometrician or (ii) there is heterogeneity in claim risk that is observed by the econometrician but unobserved by the households—and we show that our results and conclusions regarding probability distortions are robust to either possibility.[58]

---

[57]The results are very similar if we instead look at quadraticity or linearity; see Table S5 in the Supplement.

[58]In addition, we explore further the sensitivity of our results to our claim risk estimates. In particular, we consider three alternatives: (i) claim probabilities that are derived from fitted claim rates that do not condition on ex post claims experience; (ii) claim probabilities that are half as large as our estimates; and (iii) claim probabilities that are twice as large as our estimates. The results are qualitatively similar.

10. Concluding remarks

We take a partial identification approach to learning about the structure of household-specific risk preferences. Our principal identifying restriction is the assumption that preferences are stable across closely related domains. We show how one can combine stability and other structural assumptions with revealed preference arguments to conduct inference on the functionals of a generalized expected utility model that features probability distortions. A key advantage of our approach is that it does not entail making distributional assumptions to complete the model. It thus yields more credible inferences than standard approaches to identification and estimation that rely on such assumptions.

In addition to basic inference, we apply our approach to two important problems: (i) classifying households into preference types, where each type corresponds to a special case of the general model that we consider, and (ii) estimating the single parameterization of the model that best fits the data. In connection with the latter, we propose an estimator that suits our approach. Our estimator has several attractive properties, including notably the fact that, given the form of probability distortions, the resulting residuals give us precisely the lower bound on the degree of heterogeneity in probability distortions that is required to explain the data. In connection with the former, we utilize our approach to bound the prevalence of various types within the class of preferences that we consider. These bounds serve to quantify not only the explanatory potential of various models within the class of models that we consider (i.e., the fraction of households that each model can explain), but also the marginal contribution of each model to the class (i.e., the fraction of households that only the given model can explain).

The approach we develop in this paper is generalizable to other models or classes of models and can be readily applied in empirical research that investigates questions similar to those that we examine herein. In work currently in progress, for instance, Barseghyan, Molinari, O'Donoghue, and Teitelbaum (2015a) extend the approach developed here to investigate the question of narrow versus broad bracketing of risky choices. Using the same data, they study a probability distortion model that allows for the possibility that a household treats its three deductible choices as a joint decision. They perform a similar revealed preference analysis (which, under broad bracketing, maps households' choices into heptahedron) and a similar classification exercise so as to learn about the prevalence of the two forms of choice bracketing.

References

Afriat, S. N. (1967), "The construction of utility functions from expenditure data." *International Economic Review*, 8 (1), 67–77. [370, 374, 390]

Afriat, S. N. (1972), "Efficiency estimation of production functions." *International Economic Review*, 13 (3), 568–598. [370, 374, 390]

Andreoni, J., B. J. Gillen, and W. T. Harbaugh (2013), "The power of revealed preference tests: ex post evaluation of experimental design." Report, available at http://www.hss.caltech.edu/~bgillen/papers/GARPPower.pdf. [374, 396, 397]

Andrikogiannopoulou, A. and F. Papakonstantinou (2013), "Estimating heterogeneous risk preferences from a panel of real-world betting choices." Research Paper No. 13-53, Swiss Finance Institute. [373]

Barseghyan, L., F. Molinari, T. O'Donoghue, and J. C. Teitelbaum (2013a), "Distinguishing probability weighting from risk misperceptions in field data." *American Economic Review: Papers and Proceedings*, 103 (3), 580–585. [370]

Barseghyan, L., F. Molinari, T. O'Donoghue, and J. C. Teitelbaum (2013b), "The nature of risk preferences: Evidence from insurance choices." *American Economic Review*, 103 (6), 2499–2529. [373]

Barseghyan, L., F. Molinari, T. O'Donoghue, and J. C. Teitelbaum (2015a), "Bracketing and risk preferences: Identification and estimation in field data." Unpublished paper, Department of Economics, Cornell University. [403]

Barseghyan, L., F. Molinari, T. O'Donoghue, and J. C. Teitelbaum (2015b), "Estimating risk preferences in the field." Report, Department of Economics, Cornell University. [369, 382]

Barseghyan, L., J. Prince, and J. C. Teitelbaum (2011), "Are risk preferences stable across contexts? Evidence from insurance data." *American Economic Review*, 101 (2), 591–631. [372, 381, 401]

Beatty, T. K. M. and I. A. Crawford (2011), "How demanding is the revealed preference approach to demand?" *American Economic Review*, 101 (6), 2782–2795. [371, 374, 396]

Becker, G. S. (1962), "Irrational behavior and economic theory." *Journal of Political Economy*, 70 (1), 1–13. [396]

Beresteanu, A. and F. Molinari (2008), "Asymptotic properties for a class of partially identified models." *Econometrica*, 76 (4), 763–814. [372, 386, 387]

Blundell, R. W., M. Browning, and I. A. Crawford (2003), "Nonparametric Engel curves and revealed preference." *Econometrica*, 71 (1), 205–240. [374, 396]

Blundell, R. W., M. Browning, and I. A. Crawford (2008), "Best nonparametric bounds on demand responses." *Econometrica*, 76 (6), 1227–1262. [374]

Bronars, S. G. (1987), "The power of nonparametric tests of preference maximization." *Econometrica*, 55 (3), 693–698. [374, 396]

Brown, D. J. and C. Calsamiglia (2007), "The nonparametric approach to applied welfare analysis." *Economic Theory*, 31 (1), 183–188. [374]

Brown, D. J. and R. L. Matzkin (1996), "Testable restrictions on the equilibrium manifold." *Econometrica*, 64 (6), 1249–1262. [374]

Browning, M. (1989), "A nonparametric test of the life-cycle rational expectations hypothesis." *International Economic Review*, 30 (4), 979–992. [374]

Bruhin, A., H. Fehr-Duda, and T. Epper (2010), "Risk and rationality: Uncovering heterogeneity in probability distortion." *Econometrica*, 78 (4), 1375–1412. [373]

Cameron, A. C. and P. K. Trivedi (1998), *Regression Analysis of Count Data*. Econometric Society Monographs, Vol. 30. Cambridge University Press, Cambridge. [377]

Cherchye, L., I. Crawford, B. De Rock, and F. Vermeulen (2009), "The revealed preference approach to demand." In *Quantifying Consumer Preferences* (D. J. Slottje, ed.), Contributions to Economic Analysis, Vol. 288, 247–279, Emerald, Bingley. [374]

Cherchye, L., B. De Rock, and F. Vermeulen (2007), "The collective model of household consumption: A nonparametric characterization." *Econometrica*, 75 (2), 553–574. [374]

Cherchye, L., T. Demuynck, and B. De Rock (2015), "Is utility transferable? A revealed preference analysis." *Theoretical Economics*, 10 (1), 51–65. [374]

Chetty, R. (2012), "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply." *Econometrica*, 80 (3), 969–1018. [373]

Chiappori, P.-A. (1988), "Rational household labor supply." *Econometrica*, 56 (1), 63–89. [374]

Chiappori, P.-A., A. Gandhi, B. Salanié, and F. Salanié (2012), "From aggregate betting data to individual risk preferences." Discussion Paper 1213-08, Department of Economics, Columbia University. [373]

Chiappori, P.-A. and J.-C. Rochet (1987), "Revealed preferences and differentiable demand." *Econometrica*, 55 (3), 687–691. [374]

Choi, S., R. Fisman, D. Gale, and S. Kariv (2007), "Consistency and heterogeneity of individual behavior under uncertainty." *American Economic Review*, 97 (5), 1921–1938. [396]

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 (1), 169–186. [372]

Ciliberto, F. and E. Tamer (2009), "Market structure and multiple equilibria in airline markets." *Econometrica*, 77 (6), 1791–1828. [373]

Cohen, A. and L. Einav (2007), "Estimating risk preferences from deductible choice." *American Economic Review*, 97 (3), 745–788. [373, 377, 381, 384]

Conte, A., J. D. Hey, and P. G. Moffatt (2011), "Mixture models of choice under risk." *Journal of Econometrics*, 162 (1), 79–88. [373]

Cosaert, S. and T. Demuynck (2015), "Revealed preference theory for finite choice sets." *Economic Theory*, 59 (1), 169–200. [374]

Crawford, I. and B. De Rock (2014), "Empirical revealed preference." *Annual Review of Economics*, 6, 503–524. [374]

Dean, M. and D. Martin (forthcoming), "Measuring rationality with the minimum cost of revealed preference violations." *Review of Economics and Statistics*. [397]

Diewert, W. E. (1973), "Afriat and revealed preference theory." *Review of Economic Studies*, 40 (3), 419–425. [374]

Dominitz, J. and C. F. Manski (2011), "Measuring and interpreting expectations of equity returns." *Journal of Applied Econometrics*, 26 (3), 352–370. [373]

Echenique, F. and K. Saito (2015), "Savage in the market." *Econometrica*, 83, 1467-1495. [374]

Einav, L., A. Finkelstein, I. Pascu, and M. R. Cullen (2012), "How general are risk preferences? Choice under uncertainty in different domains." *American Economic Review*, 102 (6), 2606–2638. [371, 372, 401]

Famulari, M. (1995), "A household-based, nonparametric test of demand theory." *Review of Economics and Statistics*, 77 (2), 372–382. [374]

Forges, F. and V. Iehlé (2014), "Afriat's theorem for indivisible goods." *Journal of Mathematical Economics*, 54, 1–6. [374]

Gandhi, A. and R. Serrano-Padial (2015), "Does belief heterogeneity explain asset prices: The case of the longshot bias." *Review of Economic Studies*, 82 (1), 156–186. [373]

Green, E. J. and K. Osband (1991), "A revealed preference theory for expected utility." *Review of Economic Studies*, 58 (4), 677–695. [374]

Green, R. C. and S. Srivastava (1986), "Expected utility maximization and demand behavior." *Journal of Economic Theory*, 38 (2), 313–323. [374]

Gross, J. (1995), "Testing data for consistency with revealed preference." *Review of Economics and Statistics*, 77 (4), 701–710. [374]

Gul, F. (1991), "A theory of disappointment aversion." *Econometrica*, 59 (3), 667–686. [368]

Haile, P. A. and E. Tamer (2003), "Inference with an incomplete model of English auctions." *Journal of Political Economy*, 111 (1), 1–51. [373]

Handel, B. R. (2013), "Adverse selection and inertia in health insurance markets: When nudging hurts." *American Economic Review*, 103 (7), 2643–2682. [375]

Horowitz, J. L. (2009), *Semiparametric and Nonparametric Methods in Econometrics*. Springer, New York. [386]

Houtman, M. and J. A. Maks (1985), "Determining all maximal data subsets consistent with revealed preference." *Kwantitatieve Methoden*, 19, 89–104. [374]

Imbens, G. W. and C. F. Manski (2004), "Confidence intervals for partially identified parameters." *Econometrica*, 72 (6), 1845–1857. [372, 394]

Ingersoll, J. (2008), "Non-monotonicity of the Tversky–Kahneman probability-weighting function: A cautionary note." *European Financial Management*, 14 (3), 385–390. [382]

Jones, M. C. (1995), "On higher order kernels." *Journal of Nonparametric Statistics*, 5 (2), 215–221. [386]

Jullien, B. and B. Salanié (2000), "Estimating preferences under risk: The case of racetrack bettors." *Journal of Political Economy*, 108 (3), 503–530. [372]

Kahneman, D. (2003), "Maps of bounded rationality: Psychology for behavioral economics." *American Economic Review*, 93 (5), 1449–1475. [379]

Kahneman, D. and A. Tversky (1979), "Prospect theory: An analysis of decision under risk." *Econometrica*, 47 (2), 263–291. [368, 378, 382, 387]

Karni, E. (1985), *Decision Making Under Uncertainty: The Case of State-Dependent Preferences*. Harvard University Press, Cambridge, MA. [382]

Kliger, D. and O. Levy (2009), "Theories of choice under risk: Insights from financial markets." *Journal of Economic Behavior and Organization*, 71 (2), 330–346. [373]

Knoblauch, V. (1992), "A tight upper bound on the money metric utility function." *American Economic Review*, 82 (3), 660–663. [374]

Kőszegi, B. and M. Rabin (2006), "A model of reference-dependent preferences." *Quarterly Journal of Economics*, 121 (4), 1133–1166. [368]

Kőszegi, B. and M. Rabin (2007), "Reference-dependent risk attitudes." *American Economic Review*, 97 (4), 1047–1073. [368]

Kubler, F., L. Selden, and X. Wei (2014), "Asset demand based tests of expected utility maximization." *American Economic Review*, 104 (11), 3459–3480. [374]

Manski, C. F. (1989), "Anatomy of the selection problem." *Journal of Human Resources*, 24 (3), 343–360. [372]

Manski, C. F. (2003), *Partial Identification of Probability Distributions*. Springer, New York. [368, 372]

Manski, C. F. (2014), "Identification of income-leisure preferences and evaluation of income tax policy." *Quantitative Economics*, 5 (1), 145–174. [373]

Manski, C. F. and J. V. Pepper (2000), "Monotone instrumental variables: With an application to the returns to schooling." *Econometrica*, 68 (4), 997–1010. [373]

Matzkin, R. L. and M. K. Richter (1991), "Testing strictly concave rationality." *Journal of Economic Theory*, 53 (2), 287–303. [374]

Paravisini, D., V. Rappoport, and E. Ravina (2015), "Risk aversion and wealth: Evidence from person-to-person lending portfolios." Report, available at http://ssrn.com/abstract=1507902. [373]

Polisson, M. and J. K.-H. Quah (2013), "Revealed preference in a discrete consumption space." *American Economic Journal: Microeconomics*, 5 (1), 28–34. [374]

Polisson, M., J. K.-H. Quah, and L. Renou (2015), "Revealed preferences over risk and uncertainty." Discussion Paper 740, Department of Economics, University of Oxford. [374]

Prelec, D. (1998), "The probability weighting function." *Econometrica*, 66 (3), 497–527. [388]

Quiggin, J. (1982), "A theory of anticipated utility." *Journal of Economic Behavior and Organization*, 3 (4), 323–343. [368]

Rabin, M. (2000), "Risk aversion and expected-utility theory: A calibration theorem." *Econometrica*, 68 (5), 1281–1292. [378, 384]

Read, D. (2009), "Experimental tests of rationality." In *The Handbook of Rational and Social Choice: An Overview of New Foundations and Applications* (P. Anand, P. K. Pattanaik, and C. Puppe, eds.), Chapter 8, Oxford University Press, Oxford. [378]

Read, D., G. Loewenstein, and M. Rabin (1999), "Choice bracketing." *Journal of Risk and Uncertainty*, 19 (1–3), 171–197. [377]

Savage, L. J. (1954), *The Foundation of Statistics*. Wiley, New York. [378]

Selten, R. (1991), "Properties of a measure of predictive success." *Mathematical Social Sciences*, 21 (2), 153–167. [396]

Stoye, J. (2009), "More on confidence intervals for partially identified parameters." *Econometrica*, 77 (4), 1299–1315. [372, 394]

Swofford, J. L. and G. A. Whitney (1987), "Nonparametric tests of utility maximization and weak separability for consumption, leisure and money." *Review of Economics and Statistics*, 69 (3), 458–464. [374]

Sydnor, J. (2010), "(Over)insuring modest risks." *American Economic Journal: Applied Economics*, 2 (4), 177–199. [372]

Tamer, E. (2010), "Partial identification in econometrics." *Annual Review of Economics*, 2, 167–195. [372]

Teitelbaum, J. C. (2015), "Risk preference and risk type: Aggregating data from multiple domains." In *Big Data, Big Challenges in Evidence-Based Policy Making* (H. K. Jayasuriya and K. A. Ritcheske, eds.), 149–159, West Academic, St. Paul, MN. [401]

Tversky, A. and D. Kahneman (1992), "Advances in prospect theory: Cumulative representation of uncertainty." *Journal of Risk and Uncertainty*, 5 (4), 297–323. [379, 388]

Varian, H. R. (1982), "The nonparametric approach to demand analysis." *Econometrica*, 50 (4), 945–973. [374]

Varian, H. R. (1983), "Non-parametric tests of consumer behaviour." *Review of Economic Studies*, 50 (1), 99–110. [374]

Varian, H. R. (1985), "Non-parametric analysis of optimizing behavior with measurement error." *Journal of Econometrics*, 30 (1-2), 445–458. [374]

Varian, H. R. (1988), "Estimating risk aversion from Arrow–Debreu portfolio choice." *Econometrica*, 56 (4), 973–979. [374]

Varian, H. R. (1990), "Goodness-of-fit in optimizing models." *Journal of Econometrics*, 46 (1–2), 125–140. [370, 374, 390]

Varian, H. R. (1993), "Goodness-of-fit for revealed preference tests." Working paper, University of Michigan, availabe at http://econpapers.repec.org/RePEc:wpa:wuwpem:9401001. [370, 374, 390]

Varian, H. R. (2005), "Revealed preference." In *Samuelsonian Economics and the Twenty-First Century* (M. Szenberg, L. Ramrattan, and A. A. Gottesman, eds.), 99–115, Oxford University Press, Oxford. [374]

Yaari, M. E. (1987), "The dual theory of choice under risk." *Econometrica*, 55 (1), 95–115. [370]