

Ever since Ellsberg^{*}

Aluma Dembo, Shachar Kariv, Matthew Polisson, and John K.-H. Quah[†]

July 10, 2026

Ellsberg’s famous paradox challenged Savage’s subjective expected utility theory (EUT)—which reduces uncertainty to risk—by suggesting an aversion toward ambiguity. We provide a revealed preference test of the full set of axioms underpinning subjective EUT under uncertainty and compare it to an analogous test of objective EUT under risk. We find that individual choices are as consistent with utility maximization and expected utility maximization under uncertainty as they are under risk. Nevertheless, there is greater empirical scope for non-EUT models under uncertainty than under risk, and the absolute and relative consistency of EUT and non-EUT models vary considerably across subjects.

JEL CODES: D81, C91.

KEYWORDS: revealed preference, rationality, risk, uncertainty, ambiguity, multiple priors, subjective expected utility, experiment.

^{*}The experiments reported in this paper were conducted in the Experimental Social Science Laboratory (Xlab) at UC Berkeley and the California Social Science Experimental Laboratory (CASSEL) at UCLA. This research has made use of the ALICE High Performance Computing Facility at the University of Leicester and BlueCrystal/BluePebble High Performance Computing at the University of Bristol. Financial support was provided by the National Science Foundation (Grant No. SES-0962543) and the British Academy/Leverhulme Trust (Grant No. SRG21\211614).

[†]Dembo: Reichman University (aluma.dembo@runi.ac.il); Kariv: University of California, Berkeley (kariv@berkeley.edu); Polisson: University of Leicester (matt.polisson@leicester.ac.uk); Quah: National University of Singapore (ecsqkhj@nus.edu.sg).

1 INTRODUCTION

The distinction between risk (known probabilities) and uncertainty (unknown probabilities) dates back to at least Knight (1921). In Savage’s (1954) celebrated theory of subjective expected utility (EUT), however, the decision maker acts as if a single probability measure governs the states of the world, effectively reducing uncertainty to risk. Ellsberg (1961) countered this reduction using thought experiments suggesting an aversion toward ambiguity. Subsequently, a large experimental literature has documented Ellsberg-type behavior, while a large theoretical literature has developed models to accommodate it.

The objective of this paper is to provide a positive account of choice behavior under uncertainty—and to compare it to analogous behavior under risk—by evaluating the performance of subjective EUT as well as non-EUT models of ambiguity aversion. We do so in a choice environment where *all* the axioms underpinning these models can be assessed jointly. Our tests are both *comprehensive* and *nonparametric*: comprehensive in that we evaluate complete preference representations rather than individual axioms, and nonparametric in that we rely only on revealed preference relations, making no assumptions about the functional form of the underlying utility or on the structure of the errors. Such tests are feasible only with rich individual-level data—analysis at the individual level, as opposed to pooling observations or assuming homogeneity across subjects, is crucial given the well-documented heterogeneity in attitudes toward ambiguity.

Our data come from the portfolio choice experiment of Ahn *et al.* (2014). A portfolio choice problem is a choice over state-contingent commodities: there are three states of nature, denoted $s = 1, 2, 3$, and for each state s an Arrow security that pays one token (the experimental currency) if state s is realized and nothing otherwise. A portfolio therefore specifies a payoff in each state, so that choosing one amounts to allocating wealth across the three states. State 2 is assigned an objectively known probability $\pi_2 = \frac{1}{3}$, while states 1 and 3 have unknown probabilities $\pi_1 > 0$ and $\pi_3 > 0$ satisfying $\pi_1 + \pi_3 = \frac{2}{3}$. Letting $x_s \geq 0$ denote the demand for the security paying off in state s and $p_s > 0$ its price, the budget constraint is $\mathbf{p} \cdot \mathbf{x} = 1$, where $\mathbf{x} = (x_1, x_2, x_3) \geq \mathbf{0}$ and $\mathbf{p} = (p_1, p_2, p_3) \gg \mathbf{0}$, and the subject may choose any non-negative portfolio $\mathbf{x} \geq \mathbf{0}$ satisfying the constraint.

Each such problem is a standard economic decision problem that can be interpreted

either as a portfolio choice problem (the allocation of wealth among three assets) or as a consumer decision problem (the selection of a bundle of state-contingent commodities from a budget set). They are presented through a user-friendly graphical interface, which allows each subject to make many choices from widely varying budget sets—generating the rich individual-level data on which our analysis depends. To isolate the role of uncertainty, we also draw on data from a previous experiment of ours in Dembo *et al.* (2026) that is otherwise identical to the experiment of Ahn *et al.* (2014), except that $\pi_1 = \pi_3 = \frac{1}{3}$ is objectively known and announced to subjects, eliminating any uncertainty.

A key feature of this decision problem is that subjects can limit their exposure to ambiguity by choosing portfolios whose payoffs depend less on the uncertain states; in the limit, a portfolio with $x_1 = x_3$ pays the same regardless of which uncertain state occurs, leaving the subject with no exposure to ambiguity at all. To see how such choices can reveal attitudes toward ambiguity, consider the following version of the Ellsberg paradox in the portfolio choice setting, involving two budget sets. Suppose the allocation $\mathbf{x}^1$ (resp. $\mathbf{x}^2$) is chosen at the prices $\mathbf{p}^1$ (resp. $\mathbf{p}^2$) as follows:

	$\mathbf{p}$	$\mathbf{x}$
1	(p_L, p_M, p_H)	(x_M, x_H, x_L)
2	(p_H, p_M, p_L)	(x_L, x_H, x_M)

where $p_H > p_M > p_L$ and $x_H > x_M > x_L$. The price of the security that pays off in the risky state (state 2) is the intermediate of the three prices, p_M , in both budgets, but its associated demand is the highest, x_H , in each; the two securities that pay off in the uncertain states (states 1 and 3), by contrast, have their prices swapped between the first and second budgets, with their associated demands swapping accordingly.

These choices are not compatible with the maximization of a utility function that is monotonic with respect to first-order stochastic dominance (FOSD) under any single prior $\boldsymbol{\pi} = (\pi_1, \frac{1}{3}, \pi_3)$ satisfying $\pi_1 + \pi_3 = \frac{2}{3}$, and therefore, also incompatible with subjective EUT. In both budgets, the subject demands more of the security that pays off in the risky state than of the cheaper of the two uncertain securities, despite its higher price. In the first budget, where the cheaper uncertain security pays off in state 1, this is optimal under a single prior only if $\pi_1 < \frac{1}{3}$; in the second, where it pays off in state 3, it requires that

$\pi_3 < \frac{1}{3}$. These cannot both hold, since $\pi_1 + \pi_3 = \frac{2}{3}$. Yet neither $\mathbf{x}^1$ is revealed preferred to $\mathbf{x}^2$ nor vice versa, since neither was affordable at the prices prevailing when the other was chosen ($\mathbf{p}^1 \cdot \mathbf{x}^2 > \mathbf{p}^1 \cdot \mathbf{x}^1$ and $\mathbf{p}^2 \cdot \mathbf{x}^1 > \mathbf{p}^2 \cdot \mathbf{x}^2$), so these choices are consistent with utility maximization; in fact, one could even choose the rationalizing utility function to belong to a class of multiple-prior models of ambiguity aversion.

More broadly, revealed preference theory allows us to test the full hierarchy of *rationalizability* conditions on a subject’s entire dataset of many budget sets with widely varying prices. We do so by extending to uncertainty the nonparametric revealed preference approach to testing EUT against non-EUT alternatives under risk, developed by Polisson, Quah, and Renou (2020) and employed by Dembo *et al.* (2026). For each subject, we first ask whether a dataset is *rationalizable*, in the sense of being consistent with the maximization of a well-behaved (continuous and increasing) utility function. By Afriat’s (1967) theorem, a dataset is rationalizable in this basic sense if and only if it obeys the generalized axiom of revealed preference (GARP). Extensions to Afriat’s (1967) theorem then allow us to test whether a dataset is consistent with utility functions that comply with a stronger set of axioms.

In the Ahn *et al.* (2014) experiment, the probability of state 2 and the total probability of states 1 and 3 are commonly known to be $\frac{1}{3}$ and $\frac{2}{3}$, respectively. Furthermore, states 1 and 3 are treated symmetrically in the experiment, so there is no reason for any subject to believe that either uncertain state is more likely. Given these features of the experiment, it is then natural to require that the rationalizing utility function is (in addition to being well-behaved) symmetric in states 1 and 3 and increasing with respect to FOSD given the objectively known features of the underlying probability distribution. A dataset that can be rationalized by a utility function satisfying these properties is said to be *FOSD-rationalizable under multiple priors* because these requirements are compatible with multiple-prior models that treat states 1 and 3 symmetrically.

A stronger hypothesis is that a subject is compatible with *probabilistic sophistication* (Machina and Schmeidler, 1992; Epstein, 2000), which requires that a dataset can be rationalized by a well-behaved utility function that is increasing with respect to FOSD for a fixed probability distribution on all three states. In the Ahn *et al.* (2014) experiment, the symmetric treatment of states 1 and 3 means that the natural distribution to choose is the

uniform prior $\pi = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. A dataset is said to be *FOSD-rationalizable under the uniform prior* if it can be rationalized with respect to the uniform prior; this is equivalent to requiring that the rationalizing utility function is increasing and symmetric across all three states. Lastly, we can also check whether a dataset can be rationalized by a well-behaved utility function taking the expected utility form (Savage, 1954), again with a uniform distribution on the states. A dataset is said to be *EUT-rationalizable* if it is rationalizable in this sense.¹

Since GARP provides an exact test of basic rationalizability—either the data satisfy it or they do not—and rich individual-level data almost always contain at least some violations, we assess how nearly the data comply with GARP using Afriat’s (1973) critical cost-efficiency index (CCEI). The CCEI, denoted e^* and bounded between 0 and 1, is the proportion by which each budget constraint must be reduced to remove all violations of GARP—the least costly adjustment that renders the data rationalizable. It can be interpreted as the decision maker leaving up to a fraction $1 - e^*$ of the budget on the table through inconsistent choices. The notion of CCEI can be extended to measure the proportion by which budget constraints must be reduced for the data to be FOSD-rationalizable or EUT-rationalizable, and recent advances in revealed preference analysis (see Nishimura, Ok, and Quah (2017) and Polisson, Quah, and Renou (2020)) allow for the calculation of these indices. For each subject, we thus report four CCEI-type scores: e^* for (basic) rationalizability, $\tilde{e}^{**} \leq e^*$ for FOSD-rationalizability under multiple priors, $e^{**} \leq \tilde{e}^{**}$ for FOSD-rationalizability under the uniform prior, and $e^{***} \leq e^{**}$ for EUT-rationalizability.

The nested CCEI-type scores let us decompose violations of subjective EUT and its non-EUT alternatives, comparing the magnitudes of these violations across the full set of axioms underpinning these models. Our test is thus both nonparametric (based solely on revealed preference relations) and comprehensive (covering the entire set of axioms): perfect consistency with a non-EUT model that obeys ordering and monotonicity with respect to FOSD under multiple priors—but not under the uniform prior and therefore not subjective EUT itself—would yield

$$1 = e^* = \tilde{e}^{**} > e^{**} \geq e^{***}.$$

¹It is possible to check whether a dataset can be rationalized by a utility function that is increasing with respect to FOSD under a *non*-uniform distribution on the states (and the same is true for subjective EUT). We perform these tests and discuss the results in Section 7.2. It suffices to say at this point that the additional flexibility from non-uniform priors does not materially improve the performance of these models.

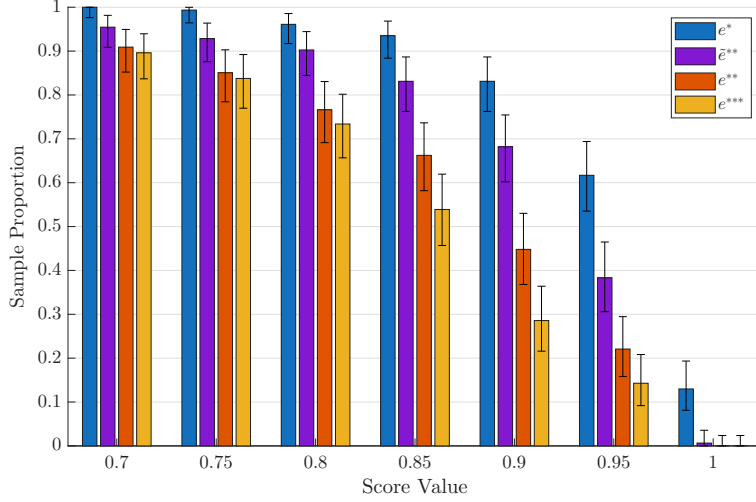


Figure 1: Distributions of Rationalizability Scores

The horizontal axis shows the score values and the vertical axis presents the fraction of subjects with scores above each value for rationalizability (e^*), FOSD-rationalizability ($\tilde{e}^{**}, e^{**}$), and EUT-rationalizability (e^{***}). The braces indicate 95 percent confidence intervals on these proportions.

Perfect consistency with subjective EUT, by contrast, would mean that all four rationalizability scores are equal to 1.

Because our experiment provides a powerful test of revealed preference, very few subjects are perfectly rationalizable at any level: out of 154 subjects only 20 (13.0 percent) are perfectly rationalizable ($e^* = 1$), one is perfectly FOSD-rationalizable under multiple priors ($\tilde{e}^{**} = 1$), none are perfectly FOSD-rationalizable under the uniform prior ($e^{**} = 1$)—and hence none are perfectly subjective EUT-rationalizable ($e^{***} = 1$). Figure 1 plots the survival functions of the e^* , $\tilde{e}^{**}$, e^{**} , and e^{***} rationalizability scores. For each score value on the horizontal axis, the vertical axis gives the fraction of subjects at or above that value.

Figure 1 reveals a nuanced pattern in which different gaps are more pronounced at different points of the score distribution. In the middle of the distribution—roughly the 0.8 to 0.9 range where the bulk of subjects are concentrated—the gap between $\tilde{e}^{**}$ and e^{**} is the most prominent, reflecting the drop in consistency from multiple-prior to uniform-prior FOSD-rationalizability, and pointing to some empirical scope for non-EUT models. Around the 0.9 level, however, a gap between e^{**} and e^{***} emerges, marking the further drop to EUT-rationalizability. At the upper tail, near 0.95 and above, both the gap between 1 and e^* and the gap between e^* and $\tilde{e}^{**}$ become prominent, meaning that at this level of consistency even basic rationalizability and multiple-prior FOSD-rationalizability are bind-

ing constraints. Taken together, the picture is one of layered complexity, with different rationalizability requirements binding for different subsets of subjects, which highlights the importance of our individual-level and comprehensive tests of all the axioms underpinning the different theories of choice under uncertainty.

We also compare our results to those obtained from the risk experiment in Dembo *et al.* (2026), which is identical to the Ahn *et al.* (2014) experiment except that the probability of all three states is commonly known to be $\frac{1}{3}$. In the Dembo *et al.* (2026) experiment, the e^* , e^{**} , and e^{***} scores are also calculated for each subject ($\tilde{e}^{**}$ has no risk counterpart). Compared to the risk experiment, the distributions of rationalizability (e^*) and EUT-rationalizability (e^{***}) scores under uncertainty are strikingly similar, implying that choices under uncertainty are as rationalizable and as EUT-rationalizable as choices under risk. The key difference is that under risk the gap between e^{**} and e^{***} is relatively modest for the vast majority of subjects, indicating little scope for the canonical non-EUT models that respect FOSD to outperform EUT. Under uncertainty, by contrast, the $\tilde{e}^{**}$ score introduces a new layer in the distribution, and the gap between $\tilde{e}^{**}$ and e^{**} is the dominant feature across much of the score range. This gap is precisely what creates the greater empirical scope for non-EUT models under uncertainty: while violations of ordering and monotonicity remain a primary source of departures from EUT in both domains, the additional wedge between multiple-prior and uniform-prior FOSD-rationalizability under uncertainty opens room for non-EUT models that find little scope under risk.

There are a variety of non-EUT models of attitudes toward ambiguity that are FOSD-rationalizable under multiple priors, though, as Ahn *et al.* (2014) shows, they all give rise to either the so-called “kinked” or “smooth” utility/demand specification. The kinked specification can be interpreted through various utility models under the broad umbrella of rank-dependent utility (RDU), while the smooth specification rests on the more recent class of recursive expected utility (REU) preferences; we discuss both of these classes in detail in Section 4.² Because both classes are accommodated by FOSD-rationalizability under multiple priors, our nonparametric approach does not distinguish between them: $\tilde{e}^{**}$ bounds the performance of the entire multiple-prior class without committing to either specification.

²RDU: Gilboa and Schmeidler (1989); Schmeidler (1989); Ghirardato, Maccheroni, and Marinacci (2004) and Olszewski (2007); Gajdos *et al.* (2008). REU: Ergin and Gul (2009), Klibanoff, Marinacci, and Mukerji (2005), Nau (2006), Seo (2009), and also Halevy and Feltkamp (2005), Giraud (2005), and Ahn (2008).

Separating RDU from REU would require a finer, computationally demanding test and even richer individual-level datasets and more powerful computational tools—an important challenge we aim to take up in future work.³

The rest of the paper is organized as follows. Section 2 provides background and situates our contribution within the wider literature. Section 3 introduces the experimental setting, and Section 4 outlines the main models of ambiguity aversion. Section 5 develops our revealed preference approach. Section 6 discusses some illustrative subjects to highlight basic features of the individual-level data, and Section 7 presents the results. Section 8 concludes.

2 BACKGROUND AND MOTIVATION

We will not attempt to review the large and growing theoretical-experimental literature on ambiguity aversion. Instead, we focus on the papers most relevant to our study. Etner, Jeleva, and Tallon (2012), Wakker (2012), and Machina and Siniscalchi (2014) provide comprehensive treatments of both theory and experimental evidence, while Trautmann and van de Kuilen (2015) provides a recent survey focused primarily on the experimental literature, and Camerer and Weber (1992) and Camerer (1995) offer excellent, if now somewhat dated, earlier ones. Key experimental papers include Halevy (2007), Bossaerts *et al.* (2010), Hey, Lotito, and Maffioletti (2010), Conte and Hey (2013), Ahn *et al.* (2014), Baillon and Bleichrodt (2015), Dimmock, Kouwenberg, and Wakker (2016), and Baillon *et al.* (2018).

Halevy (2007) develops an experiment specifically designed to examine the connection between the reduction of objective compound lotteries and attitudes toward uncertainty. Four different urns are used to elicit choices in the presence of risk, uncertainty, and two degrees of compound uncertainty. EUT and RDU under uncertainty, as well as under both risk and uncertainty (Segal, 1987, 1990), and REU generate different predictions about how reservation values for these four urns will be ordered, allowing each subject to be classified according to which model best predicts his ordering. Halevy (2007) concludes that no single model accounts for all observed behaviors—all models are represented in the subject pool.

³The difficulty in distinguishing RDU from REU stems largely from REU’s dependence on the cardinal scale of utility, as pointed out by Ahn *et al.* (2014): its ambiguity and scale parameters enter only as a product and so are not separately identified absent a normalization. This cardinal indeterminacy, which they flag as a parametric identification problem, reappears at the revealed preference level as a non-characterizability problem. We elaborate on this issue in Section 5.

Ahn *et al.* (2014) shares Halevy’s (2007) starting point that different models may be appropriate for describing different subjects’ behavior, but takes this view of heterogeneity further: not only is there variation in the appropriateness of different models across subjects, but also variation in the degree of ambiguity aversion among subjects who conform to the same model. Specifically, the experiment of Ahn *et al.* (2014) enables estimation of parametric models at the level of the individual subject: RDU under uncertainty (the “kinked” specification), REU (the “smooth” specification), and RDU under both risk and uncertainty (the “generalized kinked” specification). The first two are each characterized by two parameters, one measuring risk aversion and the other ambiguity aversion, while the generalized kinked specification adds a third parameter for pessimism/optimism and can be interpreted as a recursive non-EUT model in the sense of Segal (1987, 1990). Choices consistent with any of these specifications are FOSD-rationalizable under multiple priors but not under a single prior. Ahn *et al.* (2014) finds that the parameter estimates exhibit considerable heterogeneity across all specifications, but subjective EUT cannot be rejected for a majority of subjects, while most of the remaining subjects exhibit statistically significant ambiguity aversion (or seeking) and/or pessimism (or optimism).

A limitation shared by Halevy (2007), Ahn *et al.* (2014), and other work measuring ambiguity aversion, is that assigning each subject to one of the models risks misclassification—specifically of those who are not FOSD-rationalizable under multiple priors and/or not even rationalizable in the sense of ordering (completeness and transitivity)—conflating ambiguity attitudes with failures of the more basic axioms that underpin all these models. Halevy, Persitz, and Zrill (2018) makes the analogous point in the domain of risk: structural estimation cannot separate risk preferences from model misspecification and choice inconsistency, so that failures of basic rationalizability are absorbed into the estimated parameters rather than detected. The nonparametric revealed preference approach we adopt confronts precisely this difficulty under uncertainty, as Dembo *et al.* (2026) shows under risk, isolating failures of rationalizability and FOSD-rationalizability before any residual departure from subjective EUT is attributed to ambiguity attitudes.

Concretely, using the data of Ahn *et al.* (2014), we extend the revealed preference framework of Dembo *et al.* (2026) from risk to uncertainty, providing the first comprehensive, nonparametric, individual-level test of the full set of axioms underlying subjective EUT and

its prominent non-EUT alternatives. Just as Dembo *et al.* (2026) shows that violations of EUT under risk run much deeper than violations of independence alone—with departures from ordering and monotonicity dwarfing departures from the independence axiom—we develop analogous tests to an otherwise identical setting under uncertainty. We thus decompose violations of subjective EUT and its non-EUT alternatives in a way that no prior study has attempted, offering new guidance for theorists building models of ambiguity aversion and for experimentalists designing tests of such models.

3 EXPERIMENTAL DATA

The experimental data we analyze in this paper were originally collected by Ahn *et al.* (2014) to estimate parametric models of ambiguity aversion at the individual level. The key feature of the experimental design is that it allows subjects to make numerous choices over a wide range of budget sets, yielding rich datasets well-suited to individual-level analysis without the need to pool data or assume homogeneity across subjects. Because choices are from standard budget sets, we are able to use revealed preference analysis to assess the consistency of the individual-level data with the entire set of axioms underlying the canonical models of choice under uncertainty.

In the experiment, Ahn *et al.* (2014) presented subjects with a sequence of portfolio choice problems involving three assets—the selection of bundles of contingent commodities from three-dimensional budget sets. There are three states of nature denoted by $s = 1, 2, 3$, and for each state s , an Arrow security that pays one token (the experimental currency) in state s and nothing in the other states. To distinguish between the effects of risk and uncertainty, state 2 has an objectively known probability, whereas the probabilities of states 1 and 3 are unknown. Specifically, state 2 occurs with probability $\pi_2 = \frac{1}{3}$ and states 1 and 3 occur with unknown probabilities $\pi_1 > 0$ and $\pi_3 > 0$. Subjects were only informed that $\pi_1 + \pi_3 = \frac{2}{3}$. Except for the presence of uncertainty, the experiment of Ahn *et al.* (2014) is otherwise identical to the risk experiment of Dembo *et al.* (2026), where all states are equally likely: $\pi_1 = \pi_2 = \pi_3 = \frac{1}{3}$. We use the data of Dembo *et al.* (2026) for comparison in our analysis below.

Letting $x_s \geq 0$ denote the demand for the security paying off in state s and $p_s > 0$ denote

its corresponding price, and (without loss of generality) normalizing the experimental budget to 1, the budget set can be written as

$$\mathcal{B} = \{\mathbf{x} \in \mathbb{R}_+^3 : \mathbf{p} \cdot \mathbf{x} \leq 1\},$$

where $\mathbf{x} = (x_1, x_2, x_3)$ and $\mathbf{p} = (p_1, p_2, p_3)$. Decision problems were presented using the graphical interface developed by Choi *et al.* (2007b) and first used in Choi *et al.* (2007a), through which subjects could choose any portfolio $\mathbf{x} \geq \mathbf{0}$ on the budget constraint $\mathbf{p} \cdot \mathbf{x} = 1$.⁴

The experiment consisted of 50 decision rounds. In each round, subjects allocated tokens between three token accounts (X , Y , and Z), each corresponding to an axis in a three-dimensional graph scaled from 0 to 100 tokens.⁵ Each choice involved selecting a point on a budget set of feasible token allocations. For each round, the computer randomly selected a budget set from those intersecting at least one axis at 50 or more tokens, with no intercept below 10 tokens or above 100 tokens. The budget sets selected for each subject in his decision problems were independent of each other and of the budget sets selected for other subjects in their decision problems. Subjects were not informed of which account was actually selected at the end of each round. At the end of the experiment, one decision round was randomly selected for payment, with each round having an equal probability of being chosen.⁶

Let $\mathcal{D} := \{(\mathbf{p}^i, \mathbf{x}^i)\}_{i=1}^{50}$ be the data generated by a given subject’s choices, where $\mathbf{p}^i$ denotes the i -th observation of the price vector and $\mathbf{x}^i$ denotes the corresponding allocation.

⁴Choi *et al.* (2014) uses the same graphical methodology to study risk preferences in a nationally representative sample from the Netherlands, and Cappelen *et al.* (2023) in student samples from the US and Tanzania. Three-dimensional budget sets were first used by Fisman, Kariv, and Markovits (2007) to study social preferences. A series of subsequent papers have applied similar methodologies to different pools of subjects (Fisman, Jakiela, and Kariv, 2015, 2017; Fisman *et al.*, 2015, 2023; Li, Dow, and Kariv, 2017; Li *et al.*, 2022). Zame *et al.* (2026) combines the risk- and social-preference designs, eliciting choices across personal-risk, social-choice, and social-risk domains from the same subjects to link risk and social preferences.

⁵In Ahn *et al.* (2014), which account X , Y , or Z paid off in the risky state was rotated across experimental sessions, counterbalancing any framing or position effects tied to a particular account.

⁶The random payoff method—paying one randomly selected decision—has been much debated. A subject who integrates decisions together with the randomizing device, rather than evaluating each in isolation, can use the objective randomization to hedge ambiguity (Oechssler and Roomets, 2014; Bade, 2015; Saito, 2015; Kuzmics, 2017; Azrieli, Chambers, and Healy, 2018; Ke and Zhang, 2020). In an Ellsberg design, Baillon, Halevy, and Li (2022b) find that the random incentive system roughly halves measured ambiguity aversion, with about half of ambiguity-averse subjects integrating. These concerns bear much less directly on the Ahn *et al.* (2014) data analyzed here: subjects make 50 portfolio choices over heterogeneous budget sets, presented sequentially and with only the random price-generating process disclosed, so the cross-decision coordination needed to hedge is largely unavailable. Ahn *et al.* (2014) makes this point themselves, as do Baillon, Halevy, and Li (2022a), who place the portfolio design among the tasks least susceptible to random-incentive hedging.

We combine two datasets: Ahn *et al.* (2014), with 154 subjects making decisions under uncertainty, and Dembo *et al.* (2026), with 168 subjects making decisions under risk. We emphasize that the subjects of Ahn *et al.* (2014) and Dembo *et al.* (2026) made choices over three-dimensional budget sets, whereas the subjects of Choi *et al.* (2007a) and Choi *et al.* (2014) made choices over two-dimensional budget lines—datasets that have themselves been analyzed by Halevy, Persitz, and Zrill (2018), Polisson, Quah, and Renou (2020), de Clippel and Rozen (2023), and Echenique, Imai, and Saito (2023), among others. The three-dimensional design has two important advantages over the two-dimensional one: with three goods, the weak and generalized axioms of revealed preference—WARP and GARP—are no longer observationally equivalent, so the design can distinguish them and thereby separate incompleteness from intransitivity; and it provides a stronger test of the revealed preference conditions. We refer the reader to Dembo *et al.* (2026) for a detailed discussion.

4 MODELS OF AMBIGUITY

There are a variety of theoretical models that embody attitudes toward ambiguity, but they all give rise to one of two main utility specifications, termed “kinked” and “smooth” by Ahn *et al.* (2014). We refer to these as the rank-dependent utility (RDU) and recursive expected utility (REU) classes under uncertainty, respectively; subjective EUT is nested within both. In this section, we outline the RDU and REU classes, focusing on their functional form assumptions as a preface to the next section, where we show that any behavior consistent with these models is FOSD-rationalizable under multiple priors but not under the uniform prior. We refer the interested reader to Appendix V of Ahn *et al.* (2014) for a full discussion.

Concretely, the RDU specification takes the following form:

$$U(\mathbf{x}) = \frac{2}{3} \alpha u(x_{\min}) + \frac{1}{3} u(x_2) + \frac{2}{3} (1 - \alpha) u(x_{\max}),$$

where $u : \mathbb{R}_+ \rightarrow \mathbb{R}$ is the Bernoulli index, $x_{\min} = \min \{x_1, x_3\}$ and $x_{\max} = \max \{x_1, x_3\}$, and $\alpha \in (0, 1)$ is the ambiguity parameter. Utility is thus a weighted average of expected utility in the worst- and best-case scenarios across the two uncertain states, with $\alpha > 1/2$ (resp. $\alpha < 1/2$) indicating ambiguity aversion (resp. seeking) and $\alpha = 1/2$ reducing the RDU

specification to standard subjective EUT. The specification is so-called “kinked” because for any $\alpha \neq 1/2$ the indifference curves have a kink along the $x_1 = x_3$ diagonal, where the portfolios are ambiguity-free.

In contrast, the REU specification is everywhere differentiable—and hence the term “smooth”—so that increased ambiguity aversion affects behavior in a manner qualitatively similar to increased risk aversion. It takes the following recursive double-expectation form:

$$U(\mathbf{x}) = \int_{-1/3}^{1/3} \varphi \left(\left(\frac{1}{3} - \varepsilon \right) u(x_1) + \frac{1}{3} u(x_2) + \left(\frac{1}{3} + \varepsilon \right) u(x_3) \right) \mu(\varepsilon) d\varepsilon,$$

where $u : \mathbb{R}_+ \rightarrow \mathbb{R}$ is the Bernoulli index, $\mu : (0, 2/3) \rightarrow (0, 1)$ is a (second-order) distribution over possible priors π_1 , and $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ is a possibly nonlinear transformation over expected utility levels. The concavity (resp. convexity) of φ determines the degree of ambiguity aversion (resp. seeking), with linear φ reducing the REU class to standard subjective EUT.

Our focus here is to lay the groundwork for showing in the next section that ambiguity-averse choices consistent with the RDU and REU specifications are FOSD-rationalizable under multiple priors but not under the uniform prior. Intuitively, both specifications encode ambiguity aversion by tilting weight toward the uncertain state in which the decision maker’s demand is lower—the smaller of x_1 and x_3 . Because the identity of this state—and hence the implied weighting—shifts with the prices and the resulting portfolio, no single prior can reproduce the behavior; this is what the Ellsberg-type example above reveals, where the choices imply that $\pi_1 < \frac{1}{3}$ in one budget and $\pi_3 < \frac{1}{3}$ in another, contradicting $\pi_1 + \pi_3 = \frac{2}{3}$.⁷

5 REVEALED PREFERENCE ANALYSIS

Our analysis is nonparametric and rests entirely on revealed preference: for each subject we ask which of several nested rationalizability concepts, from basic rationalizability to EUT-rationalizability, is compatible with the observed choices, and by how much. Section 5.1 defines these rationalizability concepts and Section 5.2 turns them into cost-efficiency scores that gauge approximate compatibility. The framework extends Dembo *et al.* (2026) from

⁷Beyond the RDU and REU classes, Ahn *et al.* (2014) also considers a generalized kinked specification—a more general RDU model with an additional pessimism/optimism parameter—which can be interpreted as a recursive nonexpected utility model in the sense of Segal (1987, 1990) (their Appendix VIII). Like the RDU and REU classes, it is FOSD-rationalizable under multiple priors but not under the uniform prior.

risk to uncertainty so we keep the treatment brief, referring the reader there for the revealed preference machinery and the formal results; one new ingredient is FOSD-rationalizability under multiple priors, which accommodates ambiguity aversion (and seeking).

5.1 Rationalizability Concepts

A subject's data are the price-allocation pairs $\mathcal{D} := \{(\mathbf{p}^i, \mathbf{x}^i)\}_{i=1}^{50}$ described in Section 3, and a utility function $U : \mathbb{R}_+^3 \rightarrow \mathbb{R}$ is well-behaved if it is continuous and increasing.⁸ We consider four families of well-behaved utility functions, each nested within the last.

5.1.1 Basic Rationalizability

A dataset $\mathcal{D}$ is rationalizable if it is compatible with some well-behaved utility function U , in the sense that $U(\mathbf{x}^i) \geq U(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{B}^i = \{\mathbf{x} \in \mathbb{R}_+^3 : \mathbf{p}^i \cdot \mathbf{x} \leq 1\}$. By Afriat's (1967) Theorem, $\mathcal{D}$ is rationalizable if and only if it satisfies the generalized axiom of revealed preference (GARP). As noted in Section 3, the three-dimensional budgets make GARP and the weak axiom (WARP) observationally distinct—a WARP violation is incompatible with any complete preference and so points to incompleteness, though the underlying preference may still be transitive; a GARP violation that respects WARP is, by contrast, compatible with a complete preference but rules out transitivity. The three-dimensional design therefore separates failures of completeness from failures of transitivity. See Dembo *et al.* (2026) for the revealed preference relations and precise statements.

5.1.2 FOSD-Rationalizability

Between the basic rationalizability and EUT-rationalizability concepts lies rationalizability by a family of utility functions that are monotone with respect to first-order stochastic dominance (FOSD): a well-behaved utility function U is FOSD-increasing if $U(\mathbf{x}'') \geq U(\mathbf{x}')$ whenever the payoff distribution induced by $\mathbf{x}''$ first-order stochastically dominates that induced by $\mathbf{x}'$, and strictly when the dominance is strict. Violations of FOSD are naturally interpreted as errors rather than as expressions of risk or ambiguity attitudes, which is why

⁸By increasing, we mean that $U(\mathbf{x}') > U(\mathbf{x})$ whenever $\mathbf{x}' > \mathbf{x}$; note that $\mathbf{x}' \geq \mathbf{x}$ if $x'_s \geq x_s$ for all s , and $\mathbf{x}' > \mathbf{x}$ if $\mathbf{x}' \geq \mathbf{x}$ and $\mathbf{x}' \neq \mathbf{x}$.

this monotonicity condition is widely imposed. Recall that state 2 has the objectively known probability $\frac{1}{3}$, while the uncertain states 1 and 3 have unknown probabilities that sum to $\frac{2}{3}$; because the latter are unknown, the requirement takes two forms:

- **Multiple priors.** A dataset $\mathcal{D}$ is *FOSD-rationalizable under multiple priors* if it is compatible with a well-behaved utility function U that is (i) symmetric across the two uncertain states ($U(a, b, c) = U(c, b, a)$), and (ii) FOSD-increasing using only the objectively known structure (state 2 occurs with probability $\frac{1}{3}$, while states 1 and 3 share the remaining $\frac{2}{3}$). For a well-behaved utility function U , (ii) is equivalent to requiring $U(b, a, b) < U(a, b, a)$ whenever $b < a$. FOSD-rationalizability under multiple priors allows a subject to behave as though the division of probability between states 1 and 3 varies from choice to choice; it accommodates ambiguity aversion and seeking and contains the RDU and REU classes of Section 4.
- **Uniform prior.** A dataset $\mathcal{D}$ is *FOSD-rationalizable under the uniform prior* if it is compatible with a well-behaved utility function U that is FOSD-increasing with respect to $\pi = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, or equivalently, increasing and symmetric across all three states. In this case, we require that the subject satisfies *probabilistic sophistication* in the sense of Machina and Schmeidler (1992) and Epstein (2000), where a single division of the probability shared between states 1 and 3 is applied across all choice problems; since we also impose symmetry between states 1 and 3, this leads to the uniform distribution across all three states. For a well-behaved utility function, FOSD-rationalizability in this sense is then equivalent to requiring that the utility function is symmetric across all three states. Note that in the experiment conducted by Dembo *et al.* (2026), the distribution on the three states is commonly known to be uniform, and what they simply refer to as “FOSD-rationalizability” corresponds exactly to what we refer to as “FOSD-rationalizability under the uniform prior.” By this we mean that both definitions require a dataset to be rationalized by a well-behaved utility function that is FOSD-increasing under the uniform distribution on the three states.

Clearly, FOSD-rationalizability under the uniform prior is more demanding than FOSD-rationalizability under multiple priors. The gap between the two concepts is precisely what the canonical non-EUT models of ambiguity exploit: they are FOSD-rationalizable under

multiple priors but not under the uniform prior. We test both notions using the results of Nishimura, Ok, and Quah (2017), which strengthen GARP to rule out a larger class of revealed preference cycles; see also Dembo *et al.* (2026).

We have defined FOSD-rationalizability for the uniform prior $\boldsymbol{\pi} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$, and EUT-rationalizability below is also stated for the uniform prior. Both rationalizability concepts extend to any single prior $\boldsymbol{\pi} = (\pi_1, \frac{1}{3}, \pi_3)$ such that $\pi_1 + \pi_3 = \frac{2}{3}$, and we analyze and present results for the full range of single priors in Section 7.

5.1.3 EUT-Rationalizability

Naturally, the most stringent rationalizability concept that we consider is subjective EUT-rationalizability. Under the uniform prior, EUT requires a continuous and increasing Bernoulli index $u : \mathbb{R}_+ \rightarrow \mathbb{R}$ over monetary consumption with $U(\mathbf{x}) = \frac{1}{3} \sum_s u(x_s)$, and a dataset $\mathcal{D}$ is EUT-rationalizable if it is compatible with a well-behaved utility function U of this form. Every such U is FOSD-increasing and symmetric, so EUT-rationalizability is more demanding than FOSD-rationalizability under the uniform prior, which is in turn more demanding than FOSD-rationalizability under multiple priors. We test EUT-rationalizability using the generalized restriction of infinite domains (GRID) method of Polisson, Quah, and Renou (2020), which converts the problem into a finite system of linear inequalities; once again, we refer to Dembo *et al.* (2026) for the details.⁹

5.1.4 On Distinguishing RDU and REU

We develop tests for rationalizability, FOSD-rationalizability, and EUT-rationalizability, but not for REU-rationalizability, and accordingly we do not compare RDU and REU head to head. Since the REU class is FOSD-rationalizable under multiple priors and contains EUT as the special case of a linear aggregator, REU-rationalizability is nested between EUT-rationalizability and FOSD-rationalizability under multiple priors (and the same holds for RDU). As discussed in Section 4 (and in Ahn *et al.* (2014)), REU is identified only up to the product of its ambiguity and scale parameters. This cardinal indeterminacy reappears at the revealed preference level as a non-characterizability problem, implying that—unlike

⁹Other revealed preference approaches include Diewert (2012), Bayer *et al.* (2013), Echenique and Saito (2015), Chambers, Echenique, and Saito (2016), and Echenique, Imai, and Saito (2021).

rationalizability, FOSD-rationalizability, and EUT-rationalizability—REU-rationalizability, to our knowledge, does not admit a computationally feasible nonparametric characterization. We note that both the RDU and REU classes are FOSD-rationalizable under multiple priors, so this family already captures the performance of the entire multiple-prior class—RDU and REU alike. Distinguishing between the two remains a question that we are forced to leave for future work.

5.2 Rationalizability Scores

A subject’s data need not be exactly rationalizable by any of the rationalizability concepts outlined in Section 5.1, so we measure how close they come with using critical cost-efficiency index (CCEI) of Afriat (1972, 1973). For $e \in (0, 1]$, a dataset $\mathcal{D}$ is rationalizable at cost-efficiency e if it is compatible with some well-behaved utility function U in the relevant family once each budget is shrunk to $\mathcal{B}^i(e) = \{\mathbf{x} \in \mathbb{R}_+^3 : \mathbf{p}^i \cdot \mathbf{x} \leq e\}$, and the CCEI is the largest such e . The CCEI has the usual money-metric reading: a subject leaves up to the fraction one minus the CCEI of each budget on the table through inconsistent choices. We refer the interested reader to Dembo *et al.* (2026) for the construction and for a fuller discussion of the interpretation.

Computing the CCEI for each rationalizability concept in Section 5.1 gives four scores per subject: e^* for rationalizability, $\tilde{e}^{**}$ for FOSD-rationalizability under multiple priors, e^{**} for FOSD-rationalizability under the uniform prior, and e^{***} for EUT-rationalizability. Since each family is contained in the previous one, the scores are nested:

$$1 \geq e^* \geq \tilde{e}^{**} \geq e^{**} \geq e^{***} > 0,$$

so that each consecutive gap measures the additional cost of imposing the corresponding requirement, placing every subject’s departure from EUT-rationalizability on a single scale.

Finally, we note that the CCEI is not the only way to gauge how far a subject’s choices fall short of exact rationalizability. Other efficiency-based indices exist, and Dembo *et al.* (2026) also develops an alternative, distance-based measure, which instead asks how far the observed choices would have to be moved on the budget constraint to be rationalized. We use the CCEI for the reasons they give: it has a natural economic interpretation, it is the

standard measure in the revealed preference literature, and it can be computed for each of our nested concepts, whereas no exact procedure is known for the distance-based measure for EUT-rationalizability. Their two approaches deliver broadly the same conclusions in any case. We refer the reader to Dembo *et al.* (2026) for the comparison and details.

6 ILLUSTRATIVE SUBJECTS

The dataset from the experiment of Ahn *et al.* (2014) contains observations on 154 individual subjects. For each subject, we have 50 observations $\mathcal{D} := \{(\mathbf{p}^i, \mathbf{x}^i)\}_{i=1}^{50}$, where $\mathbf{p}^i = (p_1^i, p_2^i, p_3^i)$ denotes the i -th observation of the price vector and $\mathbf{x}^i = (x_1^i, x_2^i, x_3^i)$ denotes the corresponding demand allocation. Recall that state 2 is risky and has a $\frac{1}{3}$ chance of occurring, and states 1 and 3 are uncertain and occur according to unknown probabilities summing to $\frac{2}{3}$.

Ahn *et al.* (2014) provides an overview of the important aggregate features of their data, but these can tell us little about the choice behavior of individual subjects. To get a better sense of the wide range of observed behavior, we next provide an overview of some key features of the individual-level data. We focus on subjects whose behavior corresponds to one of several prototypical attitudes toward risk and/or uncertainty, in order to illustrate the regularity within subjects and heterogeneity across subjects that is characteristic of our data. These selected cases also bring out the heterogeneity in subjects' consistency with the different rationalizability concepts. We also report and discuss these subjects' rationalizability scores, which helps build intuition for how the scores work and how they pinpoint the sources of departure from EUT-rationalizability. Note that these are special cases where the regularities in the data are particularly clear; scatterplots for the full set of subjects are available in Appendix III of Ahn *et al.* (2014).

Figure 2 depicts the allocations chosen by six individual subjects, expressed either as token shares or budget shares across the three securities, as points in the unit simplex. The vertices of the unit simplex correspond to allocations consisting entirely of one of the three securities. Each interior point represents an allocation as a convex combination of the vertices. For any allocation $\mathbf{x}^i = (x_1^i, x_2^i, x_3^i)$, we define the *token share* of the security that pays off in state s to be the number of tokens payable in state s as a fraction of the sum of tokens payable in all three states $x_s^i / (x_1^i + x_2^i + x_3^i)$. We also define the *budget share* (or

expenditure share) of the security that pays off in state s to be the expenditure on tokens invested in this security as a fraction of total expenditure; since prices are normalized so that total expenditure equals unity, the budget share is simply $p_s^i x_s^i$.

Before proceeding to the scatterplots, we tabulate the CCEI-type scores for each of the six illustrative subjects whose choices are depicted in Figure 2: e^* for (basic) rationalizability, $\tilde{e}^{**} \leq e^*$ for FOSD-rationalizability under multiple priors, and $e^{***} \leq \tilde{e}^{**}$ for EUT-rationalizability. We do not report the uniform-prior FOSD-rationalizability scores $e^{**} \leq \tilde{e}^{**}$ for these six subjects, as they are all equal to their EUT-rationalizability scores e^{***} . Columns (a)–(f) correspond to the panels of Figure 2.

	(a)	(b)	(c)	(d)	(e)	(f)
e^*	0.994	1.000	1.000	0.943	0.981	0.934
$\tilde{e}^{**}$	0.966	0.994	0.999	0.928	0.952	0.848
e^{***}	0.965	0.994	0.991	0.660	0.801	0.750

Note that the choices of all six illustrative subjects are nearly rationalizable ($e^* \approx 1$), but only the first three subjects, whose choices are depicted in panels (a)–(c), are also nearly EUT-rationalizable ($e^{***} \approx 1$). The subjects whose choices are depicted in panels (d)–(f) are rationalizable but display meaningful departures from EUT-rationalizability.

In all panels of Figure 2, choices are depicted in terms of token shares, except in panel (c), where they are depicted in terms of budget shares. Figure 2a depicts the choices of a subject (ID 703) who always chooses nearly equal amounts $x_1^i = x_2^i = x_3^i$, suggesting infinite risk aversion. Figure 2b shows a very different case, the choices of a subject (ID 629) who, with a few exceptions, invests all his tokens in the cheapest security, behavior consistent with risk neutrality. Figure 2c depicts the choices of a subject (ID 921) who equalizes expenditures $p_1^i x_1^i = p_2^i x_2^i = p_3^i x_3^i$, rather than tokens, across the three securities. This behavior is consistent with maximizing a logarithmic expected utility function.

Figure 2d shows the choices of a subject (ID 340) who, with very few exceptions, allocates nearly equal amounts to the securities that pay off in the uncertain states $x_1^i = x_3^i \neq x_2^i$. Figure 2e shows the choices of a subject (ID 407) who does not demand nearly equal amounts of the securities that pay off in the uncertain states $x_1^i \neq x_3^i$, but these demands are much closer to each other than to the demand for the security that pays off in the risky state x_2^i .

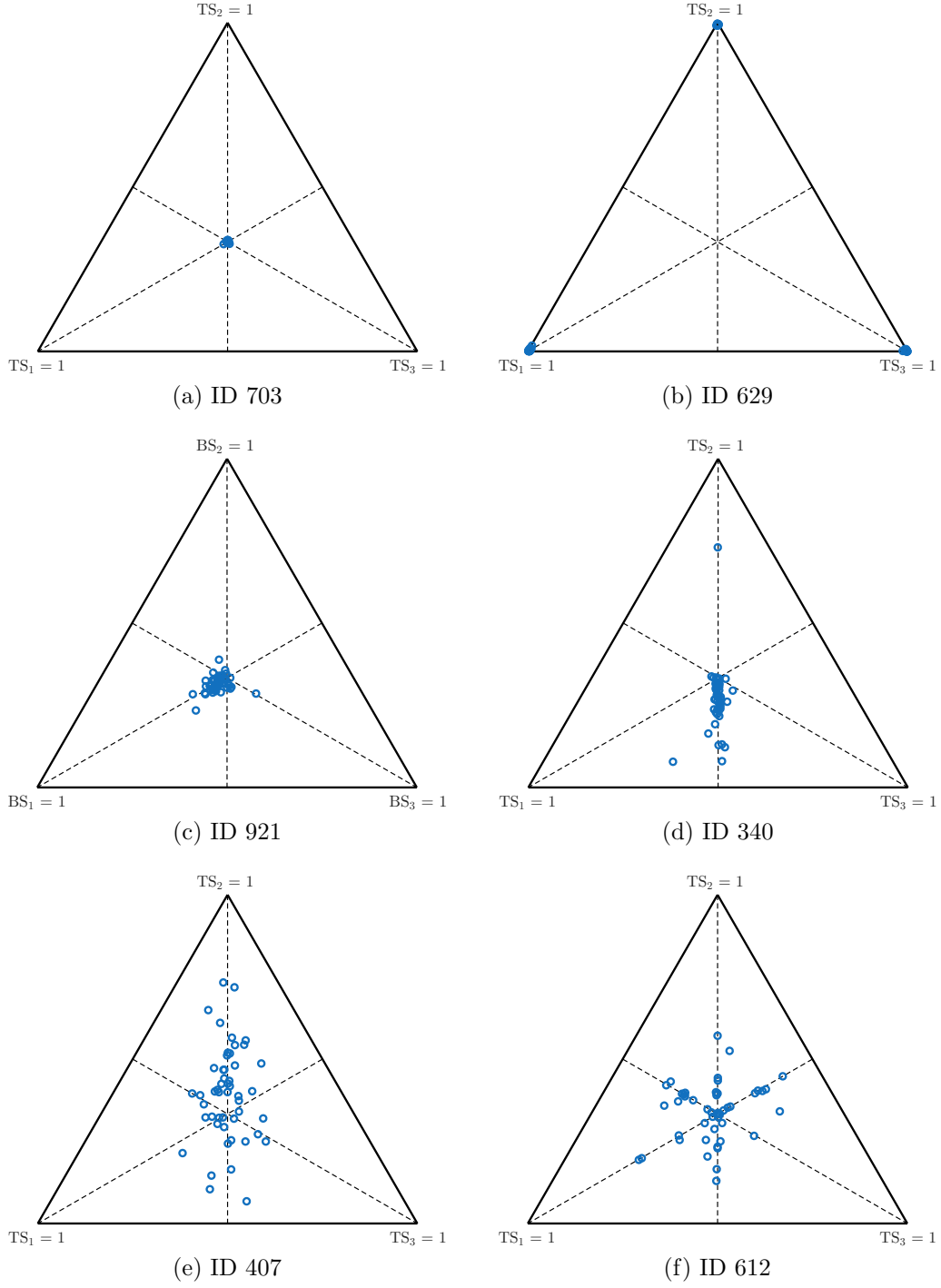


Figure 2: Subject Behavior

Each panel shows all 50 choices for a single subject in terms of token shares, except for panel (c) which is in terms of budget shares. Each vertex of the unit simplex corresponds to a full allocation to one of the three securities. (a) ID 703 is consistent with infinite risk aversion; (b) ID 629 is consistent with risk neutrality; (c) ID 921 is consistent with the maximization of logarithmic von Neumann-Morgenstern subjective expected utility; (d) ID 340 and (e) ID 407 are consistent with ambiguity aversion; and (f) ID 612 is consistent with loss aversion.

The behavior of these subjects suggests ambiguity aversion, in the sense that they are trying to reduce the sensitivity of their payoffs to states with unknown probabilities.

The main evidence for ambiguity aversion in the data is the tendency for subjects to equate, or at least reduce the difference in, their demands for the securities that pay off in the uncertain states. However, there is a similar, though weaker, tendency to equate the demands for the securities that pay off in states where one state is uncertain and the other is risky, which indicates loss aversion. Figure 2f shows the choices of one such subject (ID 612), who also equalizes the demands for the other pairs of securities where only one of the securities pays off in an uncertain state $x_1^i = x_2^i$ and $x_2^i = x_3^i$.

The subjects depicted in Figure 2d (ID 340) and Figure 2e (ID 407) are almost as FOSD-rationalizable under multiple priors as they are rationalizable ($\tilde{e}^{**} \approx e^*$), but have substantially lower EUT-rationalizability scores ($e^{***} < \tilde{e}^{**}$), consistent with strong ambiguity aversion reflected in the near-equalization of demands for the uncertain securities. For the subject depicted in Figure 2f (ID 612), by contrast, departures from EUT-rationalizability are also driven by FOSD violations even under multiple priors ($e^{***} < \tilde{e}^{**} < e^*$). This heterogeneity underscores that different subjects deviate differently from the axioms underlying models of choice under uncertainty, which is precisely why our comprehensive individual-level nonparametric tests are important: pooling data or focusing on a single axiom would obscure these distinct patterns of behavior and mischaracterize the sources of departure.

7 EXPERIMENTAL RESULTS

In this section, we present our empirical results. Throughout, we report the results from the uncertainty experiment of Ahn *et al.* (2014) alongside the corresponding results from the risk experiment of Dembo *et al.* (2026), so that every finding under uncertainty can be read directly against its counterpart under risk. The section proceeds from coarse to fine and then to robustness. Section 7.1 draws the central comparison between the two experiments: we first summarize the underlying WARP and GARP violations and the CCEI scores they generate, and then compare the full distributions of the scores for rationalizability (e^*), FOSD-rationalizability under the uniform prior (e^{**}), and EUT-rationalizability under the uniform prior (e^{***}) across the two decision domains; we also discuss FOSD-rationalizability

under multiple priors in the domain of uncertainty ($\tilde{e}^{**}$). Section 7.2 asks whether our findings depend on the uniform prior we adopt under uncertainty, and shows that the uniform prior approximately maximizes single-prior rationalizability, so that no other single prior would materially change our conclusions. Section 7.3 moves from the aggregate distributions to the individual level, locating the source of each subject’s departure from subjective EUT-rationalizability and testing, subject by subject, the magnitudes of the gaps between scores. Finally, Section 7.4 evaluates the power of our tests by comparing the observed scores to those generated from simulated data.

7.1 *Rationalizability under Uncertainty versus Risk*

First, we compare the rationalizability of choices under uncertainty with the rationalizability of choices under risk. We begin with the most basic revealed preference diagnostics—the WARP and GARP violations that underlie the CCEI rationalizability scores—and then compare the distributions of rationalizability scores across the two experiments. Throughout, we take the uniform prior $\boldsymbol{\pi} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ as the natural single-prior benchmark for the uncertainty experiment, given its symmetric design; Section 7.2 confirms that this choice is innocuous. Under the uniform prior, the scores are nested: $1 \geq e^* \geq \tilde{e}^{**} \geq e^{**} \geq e^{***} > 0$.

Before turning to the rationalizability scores, we describe the underlying revealed preference violations exhibited by individual subjects. Table 1 reports percentile values of the number of WARP violations, GARP violations, and GARP violations that do not contain a WARP violation, alongside the CCEI scores required to remove all violations of WARP and GARP. The upper panel of Table 1 reports the results under uncertainty and the lower panel under risk. The number of GARP violations is the number of distinct revealed preference cycles; a WARP violation is a special case of a GARP violation involving a pairwise cycle.

As shown in the upper panel of Table 1, the median number of WARP violations is 4 and the median number of GARP violations is 10. For a few subjects the number of GARP violations is high, but the vast majority of these GARP violations also contain WARP violations: only 25 subjects (16.2 percent) have GARP violations that do not contain one or more WARP violations. Moreover, all 20 subjects (13.0 percent) with no WARP violations also have no GARP violations, so every subject who violates GARP also violates WARP.

Uncertainty						
		Violations			1 – CCEI	
		WARP	GARP	GARP not WARP	WARP	GARP
Percentile Values	1	0	0	0	0.000	0.000
	5	0	0	0	0.000	0.000
	10	0	0	0	0.000	0.000
	25	2	2	0	0.009	0.009
	50	4	10	0	0.036	0.036
	75	9	262	0	0.076	0.076
	90	16	111,506	2	0.134	0.136
	95	21	9,057,541	6	0.163	0.163
	99	52	$\geq 10^7$	235	0.238	0.238

Risk						
		Violations			1 – CCEI	
		WARP	GARP	GARP not WARP	WARP	GARP
Percentile Values	1	0	0	0	0.000	0.000
	5	0	0	0	0.000	0.000
	10	0	0	0	0.000	0.000
	25	1	1	0	0.004	0.004
	50	4	7	0	0.031	0.032
	75	8	318	0	0.084	0.087
	90	14	181,655	1	0.172	0.172
	95	25	$\geq 10^7$	7	0.213	0.213
	99	55	$\geq 10^7$	24	0.332	0.332

Table 1: Number of WARP/GARP Violations and CCEI Scores (Uncertainty vs. Risk)

Percentile values of the number of WARP violations, GARP violations, and GARP violations that do not contain a WARP violation, along with the CCEI scores required to eliminate all WARP and GARP violations. The lower panel (risk) is reproduced from Dembo *et al.* (2026).

For each subject, we can also calculate the CCEI measuring the amount by which each budget constraint must be reduced to remove all violations of WARP, which must be weakly greater than the usual CCEI (measuring the amount by which each budget constraint must be reduced to remove all violations of GARP). These two scores are identical for all but 4 subjects (2.6 percent), and as shown in the last two columns of the upper panel of Table 1, the distributions are nearly identical and concentrated near one.

Overall, departures from rationalizability are small—most subjects are close to being rationalizable—but where deviations occur, they arise from both incompleteness and non-

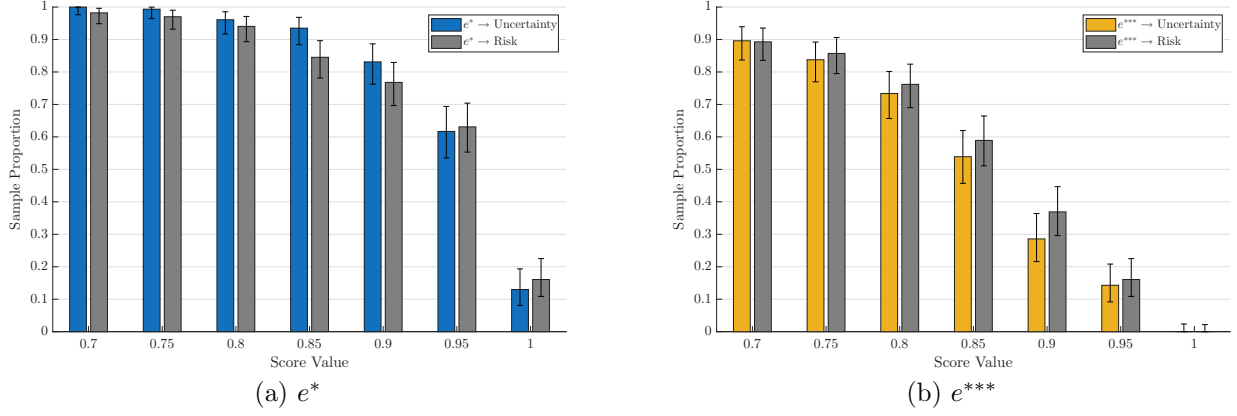


Figure 3: Distributions of Rationalizability Scores (Uncertainty vs. Risk)

The horizontal axis shows the score values and the vertical axis presents the fraction of subjects with scores above each value for (a) rationalizability (e^*) and (b) EUT-rationalizability (e^{***}). The braces indicate 95 percent confidence intervals on these proportions.

transitivity: WARP violations (occurring whenever subjects depart from rationalizability) reflect incompleteness, while GARP violations that do not reduce to WARP violations (occurring for only a minority of subjects) reflect nontransitivity. Finally, for comparison, the lower panel of Table 1 reproduces the analogous table of revealed preference violations and CCEI scores from the risk experiment of Dembo *et al.* (2026). The distributions of violations and scores across the two experiments are strikingly similar, which suggests that the introduction of uncertainty, over and above risk, does not materially degrade basic rationalizability—namely, the fundamental/conventional axioms of completeness and transitivity.

We now turn to the rationalizability scores themselves. We first take up basic rationalizability (e^*) and EUT-rationalizability (e^{***}), deferring our analysis of FOSD-rationalizability, because the distinction in between is specific to choice under uncertainty—that is, between rationalizability under the uniform prior (e^{**}) and under multiple priors ($\tilde{e}^{**}$). In Figure 3, we compare the distributions of e^* and e^{***} across the uncertainty and risk experiments. The distributions are strikingly similar across the two domains: a two-sample Kolmogorov–Smirnov test does not reject equality of the distributions at either the 1 or the 5 percent significance levels ($p = 0.4337$ and $p = 0.2263$ for e^* and e^{***} , respectively). One might have conjectured that EUT-rationalizability would deteriorate as a behavioral explanation in moving from risk to uncertainty, but we find no evidence of any difference across the two domains. Given this closeness—and the analogous result for e^{**} ($p = 0.6386$) and whose dis-

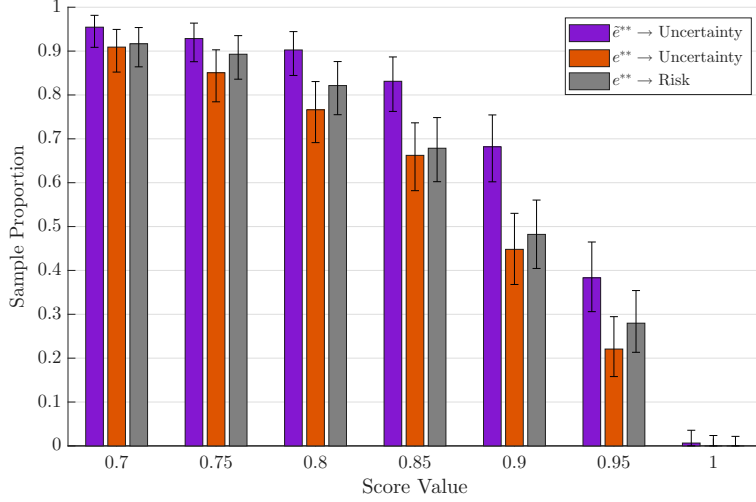


Figure 4: Distributions of FOSD-Rationalizability Scores (Uncertainty vs. Risk)

The horizontal axis shows the score values and the vertical axis presents the fraction of subjects with scores above each value for FOSD-rationalizability ($\tilde{e}^{**}$ and e^{**}). The braces indicate 95 percent confidence intervals on these proportions.

tributions are presented in Figure 4—it is no surprise that our conclusions under the uniform prior coincide with those of Dembo *et al.* (2026): for nearly all subjects there is little to no scope for single-prior non-EUT models of choice under uncertainty to explain behavior not already accounted for by subjective EUT.

Although choices under uncertainty are as rationalizable and as EUT-rationalizable as choices under risk, FOSD-rationalizability under multiple priors ($\tilde{e}^{**}$) accommodates the canonical models of decision making under uncertainty—RDU and REU—and allows for both ambiguity aversion and ambiguity seeking. Figure 4 compares the distributions of the FOSD-rationalizability scores $\tilde{e}^{**}$ and e^{**} —the former allowing for multiple priors, the latter restricted to the uniform prior—where e^{**} is shown for both the uncertainty and the risk experiments; by construction, $\tilde{e}^{**} \geq e^{**}$. Accordingly, the $\tilde{e}^{**}$ distribution in Figure 4 lies above the e^{**} distribution under uncertainty; since the vertical axis gives the fraction of subjects with a score above each value, a larger proportion of subjects clears any given threshold under multiple priors than under the uniform prior. A two-sample Kolmogorov–Smirnov test confirms that the two distributions differ significantly ($p = 0.000$). Figure 4 thus suggests that, within decision making under uncertainty, models that allow for multiple priors—whether RDU or REU in their specification—provide a better explanation of subjects’ behavior than models restricted to the uniform prior.

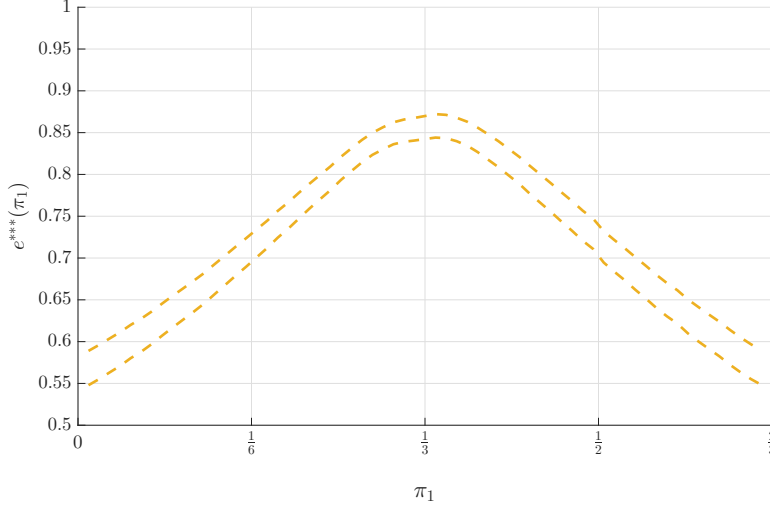


Figure 5: EUT-Rationalizability Scores under Single Priors

The horizontal axis shows the probability of state 1 (π_1) and the vertical axis presents EUT-rationalizability scores ($e^{***}(\pi_1)$). The dashed lines indicate 95 percent confidence intervals on the mean scores across subjects.

7.2 Non-Uniform Single Priors

Section 7.1 adopts the uniform prior $\boldsymbol{\pi} = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ as the natural single-prior benchmark under uncertainty. This subsection verifies that this choice does not drive our empirical findings. We ask whether some non-uniform single prior would render choices substantially more EUT-rationalizable, and show that the uniform prior approximately maximizes EUT-rationalizability. Recall that we can test for FOSD-rationalizability and EUT-rationalizability under any single prior $\boldsymbol{\pi} = (\pi_1, \frac{1}{3}, \frac{2}{3} - \pi_1)$, from which we obtain $e^{**}(\pi_1)$ and $e^{***}(\pi_1)$. The uniform prior is the natural single prior given the symmetric experimental design of Ahn *et al.* (2014), but we can test this hypothesis empirically across the full range of single priors.

Figure 5 plots $e^{***}(\pi_1)$ for different values of π_1 across subjects; since $\pi_2 = \frac{1}{3}$ is held fixed, we let π_1 (and hence π_3) vary over the interval $(0, \frac{2}{3})$ along the horizontal axis, and the vertical axis depicts 95 percent confidence intervals on the mean of $e^{***}(\pi_1)$ across subjects for each value of π_1 . Figure 5 shows that $e^{***} = e^{***}(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}) \approx \max_{\pi_1} e^{***}(\pi_1)$ on average; that is, the uniform prior approximately maximizes EUT-rationalizability under a single prior in the aggregate. The same finding holds at the individual level and under FOSD-rationalizability (omitted to economize on space). The uniform prior is therefore not merely natural given

the design—it is approximately the most favorable single prior—so the scope we document below for multiple-prior models cannot be attributed to an unfavorable choice of single prior.

7.3 Individual-Level Rationalizability Scores

Having established that the distributions of rationalizability (e^*), FOSD-rationalizability (e^{**}), and EUT-rationalizability (e^{***}) scores are indistinguishable across the uncertainty and risk experiments, while FOSD-rationalizability under multiple priors ($\tilde{e}^{**}$) accounts for a larger proportion of subjects than the uniform prior, we now examine the rationalizability scores at the individual level. For each subject the scores e^* , $\tilde{e}^{**}$, and e^{***} identify the source of any departure from subjective EUT: the values of e^* and $\tilde{e}^{**}$ reveal whether the departure stems from violations of the basic ordering axioms, from violations of first-order stochastic dominance that even multiple priors cannot accommodate, or from behavior consistent with multiple-prior non-EUT models. We first plot the scores against one another, and then compare, subject by subject, the magnitudes of the gaps between them, testing their difference using a nonparametric difference-in-differences procedure.

Figure 6a shows a scatterplot of rationalizability (e^*) against FOSD-rationalizability under multiple priors ($\tilde{e}^{**}$), and Figure 6b shows a scatterplot of FOSD-rationalizability under multiple priors ($\tilde{e}^{**}$) against EUT-rationalizability (e^{***}). By definition, $e^* \geq \tilde{e}^{**} \geq e^{***}$, so all points in both scatterplots lie on or below the 45-degree line. A subject who is perfectly subjective EUT-rationalizable is at the top-right corner of both scatterplots where $1 = e^* = \tilde{e}^{**} = e^{***}$; when $e^{***} < 1$, the corresponding values of e^* and $\tilde{e}^{**}$ identify the source of the subject's departure from subjective EUT.

Out of 154 subjects, the choices of only 20 (13.0 percent) are perfectly rationalizable ($e^* = 1$), the choices of only one (0.6 percent) are perfectly FOSD-rationalizable under multiple priors ($\tilde{e}^{**} = 1$), and the choices of none of the subjects are perfectly EUT-rationalizable ($e^{***} = 1$). Furthermore, 28 subjects (18.2 percent) fall on the 45-degree line in the scatterplot of e^* against $\tilde{e}^{**}$ (Figure 6a); the choices of these subjects need not be perfectly rationalizable, but they are no less FOSD-rationalizable under multiple priors than they are rationalizable ($e^* = \tilde{e}^{**}$). Similarly, 14 subjects (9.1 percent) fall on the 45-degree line in the scatterplot of $\tilde{e}^{**}$ against e^{***} (Figure 6b); the choices of these subjects are not perfectly

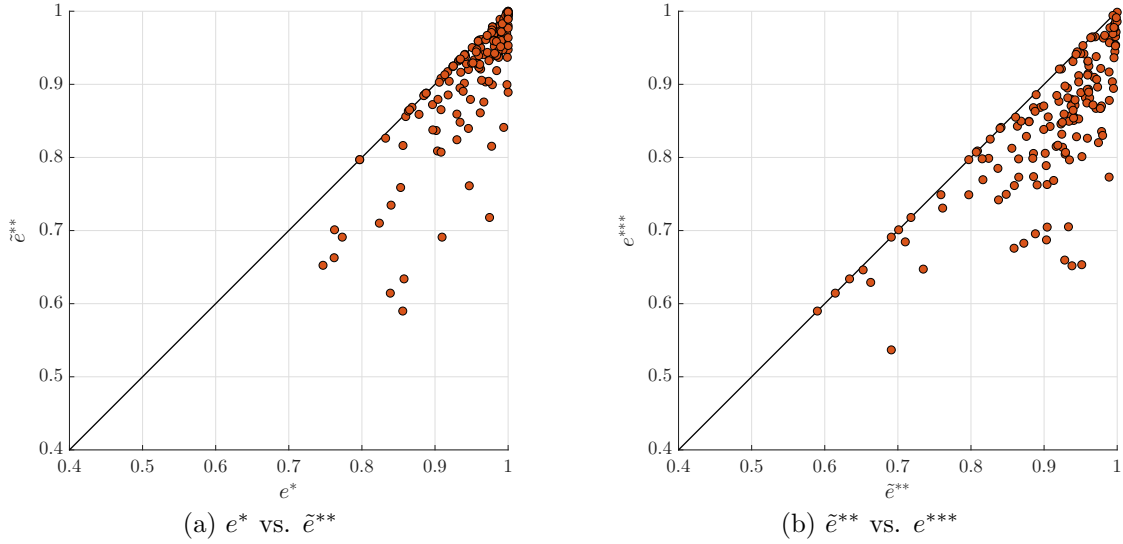


Figure 6: Individual-Level Rationalizability Scores

Rationalizability scores for individual subjects. (a) e^* (horizontal axis) vs. $\tilde{e}^{**}$ (vertical axis) and (b) $\tilde{e}^{**}$ (horizontal axis) vs. e^{***} (vertical axis). By definition, $e^* \geq \tilde{e}^{**} \geq e^{***}$, so all points in both scatterplots must lie on or below the 45-degree line.

FOSD-rationalizable under multiple priors, but they are no less EUT-rationalizable than they are FOSD-rationalizable ($\tilde{e}^{**} = e^{***}$). Finally, only 1 subject (0.6 percent) falls on the 45-degree line in both scatterplots; the choices of this subject are no less EUT-rationalizable than they are rationalizable ($e^* = \tilde{e}^{**} = e^{***}$).

Crucial to our analysis, we can also compare the magnitudes of the differences between scores. Figure 7a shows a scatterplot of the difference between perfect rationalizability and FOSED-rationalizability under multiple priors ($1 - \tilde{e}^{**}$) against the difference between FOSED-rationalizability under multiple priors and EUT-rationalizability ($\tilde{e}^{**} - e^{***}$). There is marked heterogeneity across subjects: of the 154 subjects, 80 (51.9 percent) fall above the 45-degree line ($1 - \tilde{e}^{**} < \tilde{e}^{**} - e^{***}$) and 74 (48.1 percent) fall below it ($1 - \tilde{e}^{**} > \tilde{e}^{**} - e^{***}$), so subjects are fairly evenly split on either side.

The near-even split is itself the finding. For the 80 subjects above the 45-degree line, the rationalizability that multiple-prior non-EUT models add over subjective EUT ($\tilde{e}^{**} - e^{***}$) exceeds the residual departure from perfect rationalizability that no complete and transitive preference can absorb ($1 - \tilde{e}^{**}$); for these subjects, ambiguity-sensitive preferences account for the majority of the departure from subjective EUT, rather than that departure reflecting mere inconsistency. Pooled across subjects, the two gaps are of comparable magnitude,

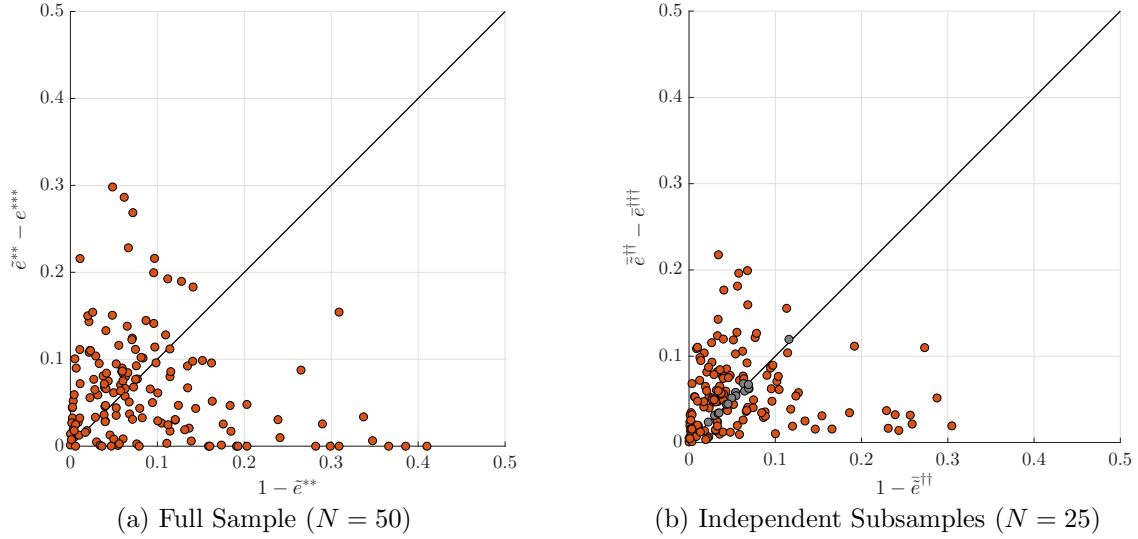


Figure 7: Differences-in-Differences in Individual-Level Rationalizability Scores

Differences between perfect rationalizability and FOSD-rationalizability (horizontal axes) vs. differences between FOSD-rationalizability and EUT-rationalizability (vertical axes). (a) $(1 - \tilde{e}^{**})$ vs. $(\tilde{e}^{**} - \tilde{e}^{***})$ using the full sample of 50 observations and (b) $(1 - \tilde{e}^{\dagger\dagger})$ vs. $(\tilde{e}^{\dagger\dagger} - \tilde{e}^{\dagger\dagger\dagger})$ using 1,000 independent subsamples of 25 observations each. In (b), subjects are depicted in red if the difference-in-difference is statistically significant at the 1 percent level and in gray if not.

so models of ambiguity carry first-order—not residual—explanatory weight. This is exactly what does not arise under risk in Dembo *et al.* (2026), where the FOSD–EUT gap $(e^{**} - e^{***})$ is roughly a fifth of the perfect–FOSD gap $(1 - e^{**})$; the even split here is direct evidence that ambiguity opens empirical space for non-EUT models that the risk setting does not.

The rich data of Ahn *et al.* (2014) also allow us to compare, for each individual subject, the gap between perfect rationalizability and FOSD-rationalizability under multiple priors with the gap between FOSD-rationalizability under multiple priors and EUT-rationalizability, using a difference-in-differences approach. The test is purely nonparametric, in that no functional form assumptions are imposed on subjects’ underlying preferences or on the error structure. To construct a test statistic and sampling distribution for each subject, we repeatedly subsample half of the subject’s data, drawing 1,000 subsamples of 25 observations each; see Dembo *et al.* (2026) for a description of the test, including the subsampling and bootstrapping procedures. For each subsample, we calculate FOSD-rationalizability under multiple priors and EUT-rationalizability scores, denoted by $\tilde{e}^{\dagger\dagger}$ and $\tilde{e}^{\dagger\dagger\dagger}$, and the resulting difference-in-differences $(1 - \tilde{e}^{\dagger\dagger}) - (\tilde{e}^{\dagger\dagger} - \tilde{e}^{\dagger\dagger\dagger})$. We use $\dagger$ rather than $*$ to denote rationalizability scores based on 25 rather than 50 observations.

Figure 7b displays the result of this test on the independent subsamples, coloring each subject red when the difference-in-differences is statistically significant at the 1 percent level and gray otherwise. The difference-in-differences fails to reach significance for only a handful of subjects—14 of the 154 (9.1 percent)—so for the remaining the test reliably places the subject on one side of the 45-degree line. This matters because subsampling uses only half of a subject’s data, resulting in fewer observed violations. That the test nonetheless rejects equality of the two gaps for nearly every subject establishes that each subject’s position relative to the 45-degree line is a genuine, individual-level feature of behavior, not an artifact of how the scores are constructed or of sampling noise.

7.4 Power

The wedge between FOSD-rationalizability under multiple priors and under the uniform prior is the empirical space in which multiple-prior non-EUT models can improve on subjective EUT. This subsection asks whether that wedge reflects genuine, systematically ambiguity averse behavior or is merely a mechanical consequence of the additional freedom that multiple priors afford. In the same spirit as Dembo *et al.* (2026), we answer this by comparing the wedge of our actual subjects with the wedge of simulated subjects whose departures from uniform-prior FOSD-rationalizability are unstructured.

Following Dembo *et al.* (2026) and in the spirit of Bronars (1987), the simulated subjects choose uniformly at random on each budget set, conditional on attaining a given level of FOSD-rationalizability under the uniform prior; that is, their choices are FOSD-rationalizable under the uniform prior up to a controlled degree of error, as measured by $e^{\dagger\dagger}$. Each simulated subject makes choices from budget sets drawn from the same distribution as the actual subjects, and for each we compute the multiple-prior FOSD-rationalizability score $\tilde{e}^{\dagger\dagger}$ and the wedge $\tilde{e}^{\dagger\dagger} - e^{\dagger\dagger}$. To compare the two groups at a common number of decisions, the actual subjects’ scores are averaged over 1,000 subsamples of 25 choices (and thus we use $\dagger$ rather than $*$ to denote rationalizability scores). See Dembo *et al.* (2026) for precise details of the simulation. Figure 8 reports the results. The horizontal axis groups subjects into bins by their uniform-prior FOSD-rationalizability score $e^{\dagger\dagger}$, and the vertical axis shows, within each bin, the mean wedge $\tilde{e}^{\dagger\dagger} - e^{\dagger\dagger}$ for the actual and the simulated subjects.

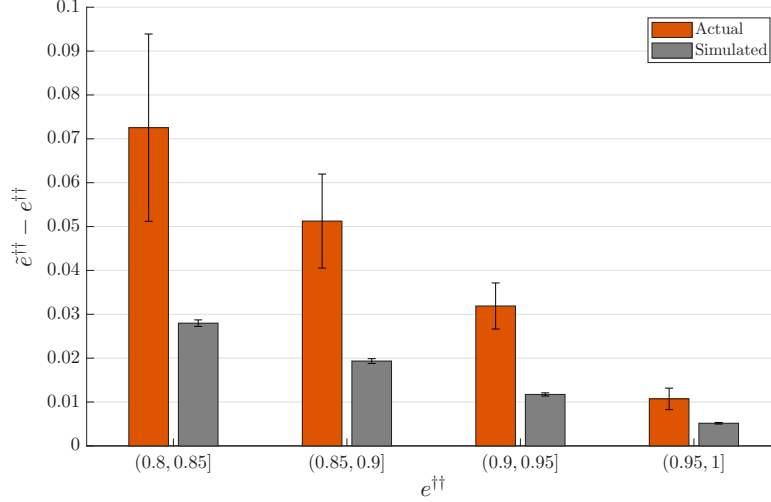


Figure 8: Power of Testing FOSD-Rationalizability

The horizontal axis shows different score value bins for FOSD-rationalizability under the uniform prior ($e^{\dagger\dagger}$). The vertical axis represents the gap between FOSD-rationalizability under multiple priors versus under the uniform prior ($e^{\dagger\dagger} - \tilde{e}^{\dagger\dagger}$). The braces represent 95 percent confidence intervals around the mean.

Across the range of uniform-prior FOSD-rationalizability scores, the wedge is larger for the actual subjects than for the simulated subjects. Because the two groups share the same uniform-prior FOSD score $e^{\dagger\dagger}$ within each bin, this gap is due entirely to the actual subjects attaining higher multiple-prior FOSD-rationalizability $\tilde{e}^{\dagger\dagger}$: their departures from uniform-prior FOSD-rationalizability are systematically of the kind that multiple priors can accommodate, whereas the simulated subjects' unstructured departures are not. The behavior of many actual subjects therefore cannot be explained as uniform-prior FOSD-rationalizability subject to error—the pattern characteristic of the simulated subjects—because such error does not open a wedge between multiple-prior and uniform-prior FOSD-rationalizability. The larger wedge among the actual subjects instead indicates genuine ambiguity aversion of the kind generated by RDU and REU models, which are FOSD-rationalizable under multiple but not under the uniform prior.

8 CONCLUDING REMARKS

This paper has carried the nonparametric revealed preference methodology of Dembo *et al.* (2026) from choice under risk to choice under uncertainty, applying it to the portfolio choice data of Ahn *et al.* (2014). As in its companion, the analysis is both nonparametric—resting

only on revealed preference relations—and comprehensive—testing the full set of axioms underpinning subjective EUT and its leading non-EUT alternatives. The central construct is a nested chain of CCEI-type scores, $1 \geq e^* \geq \tilde{e}^{**} \geq e^{**} \geq e^{***} > 0$, evaluating each subject’s data against basic rationalizability, FOSD-rationalizability, and EUT-rationalizability in turn. The score $\tilde{e}^{**}$, FOSD-rationalizability under multiple priors, has no counterpart in the risk setting; it bounds the explanatory reach of the entire multiple-prior class without committing to a particular specification.

Two findings stand out. First, choices under uncertainty are as rationalizable and as EUT-rationalizable as choices under risk: the distributions of e^* and e^{***} are statistically indistinguishable from their risk-experiment counterparts. Whatever additional cognitive burden subjective uncertainty imposes, it does not erode the consistency of behavior with utility maximization, or with expected-utility maximization, relative to objective risk. Departures from rationalizability, where they arise, reflect both incompleteness, visible as WARP violations, and nontransitivity, visible as GARP-not-WARP violations.

Second, and in sharp contrast to the risk setting, the data leave considerable room for models of ambiguity. The gap between multiple-prior and uniform-prior FOSD-rationalizability ($\tilde{e}^{**} - e^{**}$) is the dominant feature across much of the score distribution, whereas the analogous uniform-prior gap that survives under risk is modest for the vast majority of subjects. This wedge is what opens empirical space for the non-EUT models—RDU and REU—that accommodate ambiguity aversion, and it is precisely the layer the risk experiment lacks. An individual-level difference-in-differences test, applied to repeated subsamples of each subject’s data, confirms that these gaps between scores are reliable features of behavior rather than artifacts of the score construction, and a synthetic-data power analysis shows that the gap among actual subjects exceeds that which arises among simulated subjects, supporting its reading as genuine ambiguity aversion.

Our results also bear on the practice of structural estimation. When choice data depart materially from rationalizability or FOSD-rationalizability, fitting a parametric utility family risks recovering parameters of a preference that does not describe the subject, with consequences for positive prediction and welfare analysis—the concern pressed by Halevy, Persitz, and Zrill (2018). By measuring how far the data lie from each model nonparametrically and

at the individual level, before any functional form is imposed, the methodology offers a disciplined way to gauge when a given class of preferences can be taken to the data, and when it cannot. Where Savage (1954) reduced uncertainty to risk, the revealed preference record studied here shows that Ellsberg’s (1961) intuition leaves a measurable footprint—one that subjective expected utility cannot absorb but multiple-prior models can.

Finally, the departures from basic rationalizability documented in our paper are due to failures of transitivity, completeness, or both, and the literature offers models of each. On nontransitivity, the state-based regret apparatus of Bell (1982), Fishburn (1982), and Loomes and Sugden (1982) was developed for risk but transfers to the Savage setting, and Fishburn (1984, 1989), Sugden (1993), Quiggin (1994), and Nakamura (1998) supply axiomatic foundations for nontransitive subjective expected utility—a single-prior representation that relaxes transitivity alone (see also Diecidue and Somasundaram (2017)). On incompleteness, dropping completeness yields multi-utility representations under risk (Dubra, Maccheroni, and Ok, 2004) and multi-prior, multi-utility representations under uncertainty, in which non-comparability is traceable to indecisiveness in beliefs, in tastes, or in both (Ok, Ortoleva, and Riella, 2012; Galaabaatar and Karni, 2013; Riella, 2015; Karni, 2026).

In the revealed preference tradition, Eliaz and Ok (2006) weakens WARP to a non-inferiority axiom that admits rationalization by an incomplete preference. Karni and Vierø (2023) brings the apparatus into a portfolio problem over Arrow securities—the theoretical analog of our environment—where greater incompleteness of beliefs or tastes enlarges the set of undominated positions, and they construct measures of comparative incompleteness that suggest a route to quantifying, subject by subject, what our scores detect. That both departures appear together in our data points to a gap the theory has yet to close. The multi-utility tradition relaxes completeness while keeping transitivity, and the regret and nontransitive-utility tradition relaxes transitivity while keeping completeness; a model that does both—incomplete preferences whose strict part may cycle—is coherent in principle, as Mandler (2005) and Nishimura and Ok (2016) establish, but it has neither been cast in the mixture-space, independence-based form the canonical risk and uncertainty models share nor confronted with data. Casting such a model and testing it nonparametrically is an important direction for theoretical and experimental research on choice under uncertainty.

REFERENCES

- Afriat, Sydney N. 1967. “The Construction of Utility Functions from Expenditure Data.” *International Economic Review* 8(1): 67–77.
- . 1972. “Efficiency Estimation of Production Functions.” *International Economic Review* 13(3): 568–598.
- . 1973. “On a System of Inequalities in Demand Analysis: An Extension of the Classical Method.” *International Economic Review* 14(2): 460–472.
- Ahn, David, Syngjoo Choi, Douglas Gale, and Shachar Kariv. 2014. “Estimating Ambiguity Aversion in a Portfolio Choice Experiment.” *Quantitative Economics* 5(2): 195–223.
- Ahn, David S. 2008. “Ambiguity Without a State Space.” *Review of Economic Studies* 75(1): 3–28.
- Azrieli, Yaron, Christopher P. Chambers, and Paul J. Healy. 2018. “Incentives in Experiments: A Theoretical Analysis.” *Journal of Political Economy* 126(4): 1472–1503.
- Bade, Sophie. 2015. “Randomization Devices and the Elicitation of Ambiguity-Averse Preferences.” *Journal of Economic Theory* 159(A): 221–235.
- Baillon, Aurélien and Han Bleichrodt. 2015. “Testing Ambiguity Models through the Measurement of Probabilities for Gains and Losses.” *American Economic Journal: Microeconomics* 7(2): 77–100.
- Baillon, Aurélien, Yoram Halevy, and Chen Li. 2022a. “Experimental Elicitation of Ambiguity Attitude Using the Random Incentive System.” *Experimental Economics* 25: 1002–1023.
- Baillon, Aurélien, Yoram Halevy, and Chen Li. 2022b. “Randomize at Your Own Risk: On the Observability of Ambiguity Aversion.” *Econometrica* 90(3): 1085–1107.
- Baillon, Aurélien, Zhenxing Huang, Asli Selim, and Peter P. Wakker. 2018. “Measuring Ambiguity Attitudes for All (Natural) Events.” *Econometrica* 86(5): 1839–1858.
- Bayer, Ralph-C., Subir Bose, Matthew Polisson, and Ludovic Renou. 2013. “Ambiguity Revealed.” *IFS Working Papers* W13/05.
- Bell, David E. 1982. “Regret in Decision Making under Uncertainty.” *Operations Research* 30(5): 961–981.
- Bossaerts, Peter, Paolo Ghirardato, Serena Guarnaschelli, and William R. Zame. 2010. “Ambiguity in Asset Markets: Theory and Experiment.” *Review of Financial Studies* 23(4): 1325–1359.
- Bronars, Stephen G. 1987. “The Power of Nonparametric Tests of Preference Maximization.” *Econometrica* 55(3): 693–698.
- Camerer, Colin. 1995. “Individual Decision Making.” In *Handbook of Experimental Economics*, edited by John H. Kagel and Alvin E. Roth. Princeton: Princeton University Press, 587–704.
- Camerer, Colin and Martin Weber. 1992. “Recent Developments in Modeling Preferences: Uncertainty and Ambiguity.” *Journal of Risk and Uncertainty* 5(4): 325–370.

- Cappelen, Alexander W., Shachar Kariv, Erik Ø. Sørensen, and Bertil Tungodden. 2023. “The Development Gap in Economic Rationality of Future Elites.” *Games and Economic Behavior* 142: 866–878.
- Chambers, Christopher P., Federico Echenique, and Kota Saito. 2016. “Testing Theories of Financial Decision Making.” *Proceedings of the National Academy of Sciences* 113(15): 4003–4008.
- Choi, Syngjoo, Raymond Fisman, Douglas Gale, and Shachar Kariv. 2007a. “Consistency and Heterogeneity of Individual Behavior under Uncertainty.” *American Economic Review* 97(5): 1921–1938.
- Choi, Syngjoo, Raymond Fisman, Douglas M. Gale, and Shachar Kariv. 2007b. “Revealing Preferences Graphically: An Old Method Gets a New Tool Kit.” *American Economic Review: AEA Papers and Proceedings* 97(2): 153–158.
- Choi, Syngjoo, Shachar Kariv, Wieland Müller, and Dan Silverman. 2014. “Who Is (More) Rational?” *American Economic Review* 104(6): 1518–1550.
- Conte, Anna and John D. Hey. 2013. “Assessing Multiple Prior Models of Behaviour under Ambiguity.” *Journal of Risk and Uncertainty* 46(113-132).
- de Clippel, Geoffroy and Kareen Rozen. 2023. “Relaxed Optimization: How Close Is a Consumer to Satisfying First-Order Conditions?” *Review of Economics and Statistics* 104(4): 883–898.
- Dembo, Aluma, Shachar Kariv, Matthew Polisson, and John K.-H. Quah. 2026. “Ever since Allais.” *Journal of Political Economy* 134(6): 1846–1889.
- Diecidue, Enrico and Jeeva Somasundaram. 2017. “Regret Theory: A New Foundation.” *Journal of Economic Theory* 172(88-119).
- Diewert, W. Erwin. 2012. “Afriat’s Theorem and Some Extensions to Choice under Uncertainty.” *Economic Journal* 122(560): 305–331.
- Dimmock, Stephen G., Roy Kouwenberg, and Peter P. Wakker. 2016. “Ambiguity Attitudes in a Large Representative Sample.” *Management Science* 62(5): 1363–1380.
- Dubra, Juan, Fabio Maccheroni, and Efe A. Ok. 2004. “Expected Utility Theory without the Completeness Axiom.” *Journal of Economic Theory* 115(1): 118–133.
- Echenique, Federico, Taisuke Imai, and Kota Saito. 2021. “Decision Making under Uncertainty: An Experimental Study in Market Settings.” Working paper.
- . 2023. “Approximate Expected Utility Rationalization.” *Journal of the European Economic Association* 21(5): 1821–1864.
- Echenique, Federico and Kota Saito. 2015. “Savage in the Market.” *Econometrica* 83(4): 1467–1495.
- Eliasz, Kfir and Efe A. Ok. 2006. “Indifference or Indecisiveness? Choice-Theoretic Foundations of Incomplete Preferences.” *Games and Economic Behavior* 56(1): 61–86.
- Ellsberg, Daniel. 1961. “Risk, Ambiguity, and the Savage Axioms.” *Quarterly Journal of Economics* 75(4): 643–669.

- Epstein, Larry G. 2000. "Are Probabilities Used in Markets?" *Journal of Economic Theory* 91(1): 86–90.
- Ergin, Haluk and Faruk Gul. 2009. "A Theory of Subjective Compound Lotteries." *Journal of Economic Theory* 144(3): 899–929.
- Etner, Johanna, Meglena Jeleva, and Jean-Marc Tallon. 2012. "Decision Theory under Ambiguity." *Journal of Economic Surveys* 26(2): 177–372.
- Fishburn, Peter C. 1982. "Nontransitive Measurable Utility." *Journal of Mathematical Psychology* 26(1): 31–67.
- . 1984. "SSB Utility Theory and Decision-Making under Uncertainty." *Mathematical Social Sciences* 8(3): 253–285.
- . 1989. "Non-transitive Measurable Utility for Decision under Uncertainty." *Journal of Mathematical Economics* 18(2): 187–207.
- Fisman, Raymond, Pamela Jakiela, and Shachar Kariv. 2015. "How Did the Great Recession Impact Social Preferences?" *Journal of Public Economics* 128: 84–95.
- . 2017. "Distributional Preferences and Political Behavior." *Journal of Public Economics* 155: 1–10.
- Fisman, Raymond, Pamela Jakiela, Shachar Kariv, and Daniel Markovits. 2015. "The Distributional Preferences of an Elite." *Science* 349(6254): 1300.
- Fisman, Raymond, Pamela Jakiela, Shachar Kariv, and Silvia Vannutelli. 2023. "The Distributional Preferences of Americans, 2013–2016." *Experimental Economics* 26: 727–748.
- Fisman, Raymond, Shachar Kariv, and Daniel Markovits. 2007. "Individual Preferences for Giving." *American Economic Review* 97(5): 1858–1876.
- Gajdos, Thibault, Takashi Hayashi, Jean-Marc Tallon, and Jean-Christophe Vergnaud. 2008. "Attitude Toward Imprecise Information." *Journal of Economic Theory* 140(1): 27–65.
- Galaabaatar, Tsogbadral and Edi Karni. 2013. "Subjective Expected Utility with Incomplete Preferences." *Econometrica* 81(1): 255–284.
- Ghirardato, Paolo, Fabio Maccheroni, and Massimo Marinacci. 2004. "Differentiating Ambiguity and Ambiguity Attitude." *Journal of Economic Theory* 118(2): 133–173.
- Gilboa, Itzhak and David Schmeidler. 1989. "Maxmin Expected Utility with Non-Unique Prior." *Journal of Mathematical Economics* 18(2): 141–153.
- Giraud, Raphaël. 2005. "Objective Imprecise Probabilistic Information, Second Order Beliefs and Ambiguity Aversion: An Axiomatization." In *4th International Symposium on Imprecise Probabilities and Their Applications*.
- Halevy, Yoram. 2007. "Ellsberg Revisited: An Experimental Study." *Econometrica* 75(2): 503–536.
- Halevy, Yoram and Vincent Feltkamp. 2005. "A Bayesian Approach to Uncertainty Aversion." *Review of Economic Studies* 72(2): 449–466.

- Halevy, Yoram, Dotan Persitz, and Lanny Zrill. 2018. "Parametric Recoverability of Preferences." *Journal of Political Economy* 126(4): 1558–1593.
- Hey, John D., Gianna Lotito, and Anna Maffioletti. 2010. "The Descriptive and Predictive Adequacy of Theories of Decision Making under Uncertainty/Ambiguity." *Journal of Risk and Uncertainty* 41: 81–111.
- Karni, Edi. 2026. "Incomplete Preferences: Modeling, Representations, and Behavioral Implications." Working paper.
- Karni, Edi and Marie-Louise Vierø. 2023. "Comparative Incompleteness: Measurement, Behavioral Manifestations and Elicitation." *Journal of Economic Behavior and Organization* 205: 423–442.
- Ke, Shaowei and Qi Zhang. 2020. "Randomization and Ambiguity Aversion." *Econometrica* 88(3): 1159–1195.
- Klibanoff, Peter, Massimo Marinacci, and Sujoy Mukerji. 2005. "A Smooth Model of Decision Making under Ambiguity." *Econometrica* 73(6): 1849–1892.
- Knight, Frank H. 1921. *Risk, Uncertainty, and Profit*. Boston and New York: Houghton Mifflin Company.
- Kuzmics, Christoph. 2017. "Abraham Wald's complete class theorem and Knightian uncertainty." *Games and Economic Behavior* 104: 666–673.
- Li, Jing, Lawrence P. Casalino, Raymond Fisman, Shachar Kariv, and Daniel Markovits. 2022. "Experimental Evidence of Physician Social Preferences." *Proceedings of the National Academy of Sciences* 119(28): 1–11.
- Li, Jing, William H. Dow, and Shachar Kariv. 2017. "Social Preferences of Future Physicians." *Proceedings of the National Academy of Sciences* 114(48): 10291–10300.
- Loomes, Graham and Robert Sugden. 1982. "Regret Theory: An Alternative Theory of Rational Choice under Uncertainty." *Economic Journal* 92(368): 805–824.
- Machina, Mark J. and D. Schmeidler. 1992. "A More Robust Definition of Subjective Probability." *Econometrica* 60(4): 745–780.
- Machina, Mark J. and Marciano Siniscalchi. 2014. "Ambiguity and Ambiguity Aversion." In *Handbook of the Economics of Risk and Uncertainty*, edited by Mark J. Machina and W. Kip Viscusi. Oxford: Elsevier, 729–807.
- Mandler, Michael. 2005. "Incomplete Preferences and Rational Intransitivity of Choice." *Games and Economic Behavior* 50(2): 255–277.
- Nakamura, Yutaka. 1998. "Skew-Symmetric Additive Representations of Preferences." *Journal of Mathematical Economics* 30(3): 367–387.
- Nau, Robert F. 2006. "Uncertainty Aversion with Second-Order Utilities and Probabilities." *Management Science* 52(1): 136–145.
- Nishimura, Hiroki and Efe A. Ok. 2016. "Utility Representation of an Incomplete and Nontransitive Preference Relation." *Journal of Economic Theory* 166: 164–185.

- Nishimura, Hiroki, Efe A. Ok, and John K.-H. Quah. 2017. "A Comprehensive Approach to Revealed Preference Theory." *American Economic Review* 107(4): 1239–1263.
- Oechssler, Jörg and Alex Roomets. 2014. "Unintended Hedging in Ambiguity Experiments." *Economics Letters* 122(2): 243–246.
- Ok, Efe A., Pietro Ortoleva, and Gil Riella. 2012. "Incomplete Preferences under Uncertainty: Indecisiveness in Beliefs versus Tastes." *Econometrica* 80(4): 1791–1808.
- Olszewski, Wojciech. 2007. "Preferences Over Sets of Lotteries." *Review of Economic Studies* 74(2): 567–595.
- Polisson, Matthew, John K.-H. Quah, and Ludovic Renou. 2020. "Revealed Preferences over Risk and Uncertainty." *American Economic Review* 110(6): 1782–1820.
- Quiggin, John. 1994. "Regret Theory with General Choice Sets." *Journal of Risk and Uncertainty* 8(2): 153–165.
- Riella, Gil. 2015. "On the Representation of Incomplete Preferences under Uncertainty with Indecisiveness in Tastes and Beliefs." *Economic Theory* 58: 571–600.
- Saito, Kota. 2015. "Preferences for Flexibility and Randomization under Uncertainty." *American Economic Review* 105(3): 1246–1271.
- Savage, Leonard J. 1954. *The Foundations of Statistics*. New York: John Wiley and Sons.
- Schmeidler, David. 1989. "Subjective Probability and Expected Utility Without Additivity." *Econometrica* 57(3): 571–587.
- Segal, Uzi. 1987. "The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach." *International Economic Review* 28(1): 175–202.
- . 1990. "Two-Stage Lotteries without the Reduction Axiom." *Econometrica* 58(2): 349–377.
- Seo, Kyoungwon. 2009. "Ambiguity and Second-Order Belief." *Econometrica* 77(5): 1575–1605.
- Sugden, Robert. 1993. "An Axiomatic Foundation for Regret Theory." *Journal of Economic Theory* 60(1): 159–180.
- Trautmann, Stefan T. and Gijs van de Kuilen. 2015. "Ambiguity Attitudes." In *The Wiley Blackwell Handbook of Judgment and Decision Making*, edited by Gideon Keren and George Wu. Chichester: John Wiley and Sons, 89–116.
- Wakker, Peter P. 2012. *Prospect Theory: For Risk and Ambiguity*. Cambridge University Press.
- Zame, William R., Bertil Tungodden, Erik Ø. Sørensen, Shachar Kariv, and Alexander W. Capelen. 2026. "Linking Social and Personal Preferences: Theory and Experiment." *Journal of Political Economy* 134(6): 1890–1930.