Primal and Dual Methods for Estimating Random Preference Models of Risk Attitudes

by

Morten I. Lau & Hong Il Yoo†

October 2021

Abstract. We present two methods for estimating Random Preference (RP) models of risk attitudes. Our first method is intended for any class of RP models, including those addressing multidimensional risk attitudes as per non-expected utility theories. Existing methods cannot be used to estimate such models without imposing severe constraints on the model and data. We present a fully versatile alternative. Our second method is intended for models of unidimensional risk attitudes that display single-crossing properties. We show that these models are dual to seemingly non-structural regression models. One can thus use standard regression models in structural estimation of unidimensional risk attitudes.

Keywords: stochastic choice, random utility, random coefficient, risk aversion, structural estimation, maximum simulated likelihood

JEL codes: C15, C51, D81, D91

† Department of Economics, Copenhagen Business School, Copenhagen, Denmark (Lau); and Department of Economics and Finance, Durham University, Durham, United Kingdom (Lau and Yoo). E-mail addresses: mla.eco@cbs.dk and h.i.yoo@durham.ac.uk. We would like to thank the Danish Council for Independent Research | Social Sciences (project number: DFF-7015-00054) for financial support, and Jose Apesteguia, Miguel A. Ballester, Andre Hofmeyr and Glenn W. Harrison for helpful comments.

Table of Contents

1. Introduction

Recent theoretical developments in behavioral economics have generated renewed interest in the Random Preference (RP) model of choice under risk.[1] Structural models of individual choice behavior typically assume that latent preference parameters are deterministic at the individual level. That is, these models of latent preferences assume that the parameters do not vary randomly *within an individual* over different decision tasks.[2] To account for unexplained variation in individual choice behavior, the deterministic preference index function is combined with an additive error term that varies randomly between individuals and over decision tasks. By contrast, the RP model excludes the additive error term and treats the preference parameters *per se* as stochastic variables that vary randomly between individuals and over decision tasks. The descriptor "random" in the RP model is thus used in relation to the preference parameters rather than the additive error term.

From an empirical perspective, the RP model of choice under risk appears to be an unwieldy model to estimate. Structural estimation of the additive error model often proceeds by embedding a preference index function (*e.g.*, the difference in expected utility between two lotteries) in a tractable link function (*e.g.,* the logistic distribution function) to obtain analytic choice probabilities. In general, this approach does not work with the RP model since the preference index function is not additively separable from the source of stochastic variation, namely the random risk preference parameters. One may apply a general purpose simulator based on the simple frequency logic to approximate the choice probabilities with relative ease, but the simulated likelihood function is a step function which does not lend itself to gradient-based optimization.[3] Existing methods to estimate the RP model focus on special cases that enable one to derive analytic choice probabilities. These special cases require that the underlying decision model is Expected Utility Theory (EUT)

---

[1] See, for example, Gul and Pesendorfer [2006] and Apesteguia and Ballester [2018]. The RP model of risk attitudes was first presented in Becker, DeGroot and Marschak [1963; pp. 42-43].

[2] This assumption still allows the preference parameters to vary deterministically within an individual according to observed characteristics of the decision tasks; for example, Harrison, Lau and Yoo [2020] let individual risk attitudes vary with an indicator of high and low stake lotteries in the decision tasks. This assumption also allows the preference parameters to vary randomly between individuals to capture unobserved preference heterogeneity in the same manner as random coefficient models.

[3] Global search methods, such as differential evolution and particle swarm optimization algorithms (*e.g.*, Hole and Yoo [2017]), may help one maximize a step function but do not address problems with point identification of parameter values.

with a one-parameter utility function, thereby precluding non-EUT models as well as EUT with more flexible utility functions; or that each observed choice is between lotteries with a common set of three outcomes, thereby precluding modern experimental designs that specify four or more outcomes per choice task. Empirical studies of random preferences under EUT with one-parameter utility functions remain relatively small in numbers,[4] and there are virtually no studies that combine random preferences with more general decision models.

We propose two new methods to estimate the RP model. Our first method is applicable to any RP model of risky choice, regardless of the underlying theory and experimental design. This method estimates the primal representation of the RP model directly, by applying McFadden's [1989] perturbation strategy to construct an approximate likelihood function. This function can be coded with relative ease and generalized to accommodate multidimensional risk preference parameters, as well as unobserved interpersonal heterogeneity, in a tractable fashion. The main idea is to replace indicator functions in the simple frequency simulator with smoothing kernels such as logistic distribution functions. The simulated likelihood function is smooth in the parameters to be estimated and can be maximized using gradient-based optimization algorithms.

The versatility of our primal method is perhaps best illustrated with comparisons to previous efforts at estimating Quiggin's [1982] Rank-Dependent Utility (RDU), which extends EUT by adding a probability weighting function (PWF) that complements the utility function. Despite the popularity of RDU, only Loomes, Moffatt and Sugden [2002] and Wilcox [2008][2011] have estimated RP versions of this decision model. Their computational approach requires all pairwise lottery choice tasks in the data to involve the same set of three outcomes (Loomes, Moffatt and Sugden [2002]) or at most three different sets of three outcomes (Wilcox [2008][2011]), and also requires the PWF to be *deterministic* within an individual.[5] The latter constraint implies that their models are partial RP models which only apply random preferences to the utility function and not to the PWF. In contrast, our primal method can accommodate full RP models which apply random

[4] For example, Apesteguia, Ballester and Gutierrez-Daza [2020] and Jagelka [2020] estimate a RP model under EUT with a power utility function.

[5] The PWF is deterministic in relation to the RP model, *i.e.*, it does not vary randomly *within* an individual. Wilcox [2008][2011] allows the deterministic PWF to vary randomly *between* individuals to address interpersonal heterogeneity.

preferences to both sources of risk attitudes in RDU, and more generally to multiple sources of risk attitudes in any given theory.

Our second estimation method is intended for RP models with unidimensional risk preferences that satisfy the single crossing condition. This method estimates the RP model indirectly by exploiting its dual representation as a standard regression model. While the applicable class of decision models is limited, the dual likelihood function is exact rather than approximate with respect to the RP component. The dual method thus complements the primal method by producing benchmark estimates that one can use to examine the extent of approximation errors in the corresponding primal estimates empirically. Moreover, the dual method enables one to easily estimate advanced structural models that combine the RP model of *intrapersonal* heterogeneity with the random coefficient model of *interpersonal* heterogeneity. For example, let ω denote a risk preference parameter that is randomly distributed within an individual. In a binary choice task, this parameter satisfies the single crossing condition if there is a unique value w such that the individual chooses one lottery when $\omega > w$, and the other lottery when $\omega < w$. Viewing ω as a latent dependent variable and w as a known threshold allows us to recast the RP model as a linear index model with an additive error term, just like a standard discrete choice model. One can thus use logit and probit regression commands for the dual standard model to estimate the primal RP model. Discrete choice models for panel data, such as random effects logit and mixed logit, provide accessible avenues to incorporate interpersonal heterogeneity by making the distribution of the random preference parameter individual-specific. We are not aware of any existing study that applies standard regression commands in structural estimation of risk attitudes, be it in the stochastic framework of the RP model or the additive error model.

Our primal estimation method will help advance the boundaries in structural estimation of choice under risk. Wilcox [2008; §4.5], for example, concludes that the RP model lacks generalizability relative to other stochastic choice models because available econometric methods only allow one to estimate special cases of the RP model. Our primal method puts the RP model on an equal footing with other stochastic choice models in terms of generalizability. We illustrate this point with applications to EUT which take a non-parametric approach to the utility function, by specifying a separate RP parameter to represent the utility level of each outcome. This specification has been used in the theoretical formulation by Gul and Pesendorfer [2006] but cannot be estimated by existing methods unless one imposes similar restrictions on the data as Loomes, Moffatt and Sugden [2002] and Wilcox [2008][2011]. We also illustrate the method with applications to the full RP version of RDU with constant relative risk aversion (CRRA) utility, as well as non-parametric utility. These applications do not exhaust analytic opportunities that our versatile method opens up. One may also apply our method to other non-EUT models such as Prospect Theory; to more unconventional types of decision tasks with multinomial and rank-ordered choices; and to other branches of behavioral economics such as the analysis of non-exponential discounting functions with multidimensional time preference parameters.

Our dual estimation method will be useful in studies on socio-economic determinants of risk preferences that consider an individual's risk attitude as a unidimensional trait (*e.g.*, Dohmen, Falk, Huffman and Sunde [2010], Filippin and Crosetto [2016], Guiso, Sapienza and Zingales [2018] and Hryshko, Luengo-Prado and Sørensen [2011]). With our dual method, the RP model of unidimensional risk preferences (*e.g.*, EUT with CRRA utility) becomes an attractive alternative to reduced-form regression models which are widely used. Both types of models can be estimated using standard regression commands in software packages, but the RP model has a more solid theoretical foundation that allows one to distinguish interpersonal preference heterogeneity from behavioral noise.[6] Large household surveys, such as the Panel Study of Income Dynamics in the USA, Socio-Economic Panel in Germany and the UK Household Longitudinal Study, include binary choice tasks under risk for which one-parameter formulations of EUT and Yaari's [1987] Dual Theory display the single crossing property, increasing the appeal of our dual method.[7]

[6] RP models of unidimensional risk attitudes are not necessarily more restrictive than additive error models of multidimensional risk attitudes in structural Maximum Likelihood (ML) and Non-linear Least Squares (NLS) estimation. The latter models usually refer to fixed coefficient specifications that either neglect unobserved interpersonal heterogeneity or address it by applying individual-level estimations that preclude population-level inferences.

[7] Combining the RP model with the single crossing property also plays an important analytic role in Barseghyan, Molinari, and Thirkettle [2021], who develop a semi-nonparametric estimator of risk aversion in market settings where the analyst does not fully observe the decision maker's choice sets and relevant product attributes.

We stress that the RP model *does not* refer to the usual random coefficient model that has been widely used in applied microeconomics. The two models serve quite different purposes, and one can combine the RP model with the fixed coefficient model (*e.g.*, Apesteguia and Ballester [2018]) or with the random coefficient model (*e.g.*, Wilcox [2008][2011]). The RP model uses a statistical distribution to describe how preference parameters vary *within* an individual over decision tasks, whereas the random coefficient model uses a statistical distribution to describe how preference parameters vary *between* individuals.[8] Suppose that the CRRA coefficient is logistically distributed within each individual, and that the mean and log-scale of the logistic CRRA coefficient are individual-specific and follow a joint normal distribution between individuals. The RP model refers to the logistic distribution at the individual level, and the random coefficient model refers to the joint normal distribution at the population level. Our primal and dual estimation methods apply to both fixed and random coefficient versions of the RP model.

The rest of this paper is organized as follows. Section 2 summarizes the key features of the RP model and the two data sets that we use in the empirical illustrations. Section 3 presents our dual estimation method, which is intended for RP models with a unidimensional risk preference parameter that displays the single crossing property. Section 4 presents our primal method which can be applied to any RP model, regardless of the single crossing property and the dimensionality of risk preference parameters. Section 5 concludes.

2. Preliminaries

We first discuss the key characteristics of the RP model, and outline the decision theoretic components of the different RP model specifications that we estimate in sections 3 and 4. We then summarize the two data sets that we use in our empirical applications.

[8] The use of the term *random coefficient* to describe this approach to modeling interpersonal heterogeneity is well-established in all branches of applied microeconomics: See, for example, Revelt and Train [1998], Layton and Brown [2000], Knittel and Metaxoglou [2014] and Harrison, Lau and Yoo [2020].

A. Random Preference Models

Assume for now that every subject has either the same *fixed* risk preference parameter or the same "urn of *random* risk preference parameters," a metaphor to be clarified shortly.[9] Suppose that the subject makes a choice between two lotteries: A and B. Each lottery $L \in \{A, B\}$ is a probability distribution over K prizes which pays prize m_{Lk} with probability p_{Lk}, where $k = 1, 2, \cdots, K$. All prizes and probabilities in each lottery are known to the subjects when they make their decision. To facilitate the discussion of RDU, suppose that the prizes in lottery L are ordered from worst to best, $m_{L1} < m_{L2} < \cdots < m_{LK}$, and let P_{Lk} denote the cumulative probability of prizes which are at least as good as m_{Lk}, $P_{Lk} = p_{Lk} + p_{L(k+1)} + \cdots + p_{LK}$, with $P_{L(K+1)} = 0$. The cumulative probability is thus a step function that is equal to one for the worst prize and takes smaller values for better prizes.

We can maintain generality in our exposition by invoking a generic decision theory under which the subject's evaluation of lottery L is given by $V_L(\boldsymbol{\alpha})$, where the parametric form of $V_L(.)$ varies across models and $\boldsymbol{\alpha}$ is a multidimensional vector of risk preference parameters. Nevertheless, for concreteness, we start with a simple example where $V_L(.)$ is the expected utility of lottery L, denoted by $EU_L(.)$, and $\boldsymbol{\alpha}$ is a unidimensional risk preference parameter, denoted by ω. Specifically, assume that utility of prize M is given by a CRRA function

$$U(M \mid \omega) = M^{(1-\omega)}/(1-\omega), \tag{1}$$

where ω is the coefficient of relative risk aversion. The expected utility of lottery $L \in \{A, B\}$ is then

$$EU_L(\omega) = \textstyle\sum_k p_{Lk} \times U(m_{Lk} \mid \omega). \tag{2}$$

Henceforth, we use $n = 1, 2, \cdots, N$ to index subjects, and $t = 1, 2, \cdots, T$ to index decision tasks. $EU_L(\omega)$ in (2) is subsequently indexed by n and t, since prizes and probabilities in each lottery may vary across decision tasks and subjects.

Let y_{nt} indicate subject n's binary choice in task t between lottery B ($y_{nt} = 1$) or lottery A ($y_{nt} = 0$), and let $\Delta EU_{nt}(\omega)$ be the difference in expected utility between the two lotteries, $\Delta EU_{nt}(\omega) = EU_{B,nt}(\omega) - EU_{A,nt}(\omega)$. Using the indicator function $\mathbf{I}[.]$, the predicted choice under EUT can be

[9] In section 3 and 4, we will introduce random coefficient models to relax this representative agent assumption.

written as $y_{nt} = \mathbf{I}[\Delta EU_{nt}(\omega) > 0]$. However, for any given value of ω, this equality may fail to hold as some observed choices may deviate from the theoretical predictions of the decision model.

One can account for such behavioral deviations from theory by combining the preference index function with stochastic elements to construct a latent dependent variable y_{nt}^* that determines the observed choice, $y_{nt} = \mathbf{I}[y_{nt}^* > 0]$. In an additive error model, this is specified as

$$y_{nt}^* = \Delta EU_{nt}(\omega) + \boldsymbol{\varepsilon}_{nt}, \tag{3}$$

where $\boldsymbol{\varepsilon}_{nt}$ captures zero-mean behavioral errors that are independently distributed. One can derive choice probabilities by inserting the non-linear index into a parametric distribution function for $\boldsymbol{\varepsilon}_{nt}$. For example, assume that $\boldsymbol{\varepsilon}_{nt}$ is logistically distributed with standard deviation $\sigma \times 1.81$ and density function $f(\boldsymbol{\varepsilon}_{nt} | 0, \sigma)$.[10] The probability that subject n chooses lottery B in task t is

$$L_{nt}(\omega, \sigma) = \int \mathbf{I}[\Delta EU_{nt}(\omega) + \boldsymbol{\varepsilon}_{nt} > 0] f(\boldsymbol{\varepsilon}_{nt} | 0, \sigma) d\boldsymbol{\varepsilon}_{nt} = \Lambda(\Delta EU_{nt}(\omega)/\sigma), \tag{4}$$

where $\Lambda(.) = \exp(.) / [1 + \exp(.)]$ is the standard logistic distribution function.[11]

The Random Preference (RP) model excludes the additive error term and accommodates stochastic choice behavior by treating the risk aversion parameter as a random variable, ω_{nt}, that varies across subjects and, more importantly, *over decision tasks within a subject*. The latent dependent variable is

$$y_{nt}^* = \Delta EU_{nt}(\omega_{nt}). \tag{5}$$

Suppose that the random risk aversion parameter ω_{nt} is logistically distributed with mean μ_ω and standard deviation $\sigma_\omega \times 1.81$, and let $f(\omega_{nt} | \mu_\omega, \sigma_\omega)$ denote the density function. One can interpret the density function as an "urn" that contains different values of ω_{nt} (Wilcox [2008; p.213]), and μ_ω and σ_ω as the mean and dispersion of the urn's contents. In each decision task, the subject makes a new draw from the urn, with replacement, and the outcome of that draw determines the degree of relative risk aversion in that task. The probability that subject n chooses lottery B in task t is then

[10] The factor of proportionality, σ, refers to the scale parameter of the logistic distribution, which typically is normalized to 1 to achieve identification (*e.g.,* Wooldridge [2010; §15.3]). This is the *standard* logistic distribution with a standard deviation of $\pi/3^{0.5} \approx 1.81$. The non-linear index function $\Delta EU_{nt}(\omega)$ in (4) allows one to identify both ω and σ without this normalization.

[11] This popular stochastic specification is known as the Fechner model or strong utility model. Wilcox [2008] provides an extensive review of related stochastic choice models, including moderate utility models that accommodate heteroskedasticity with respect to observed task characteristics.

$$L_{nt}(\mu_\omega, \sigma_\omega) = \int \mathbf{I}[\Delta EU_{nt}(\omega_{nt}) > 0] f(\omega_{nt} \mid \mu_\omega, \sigma_\omega) d\omega_{nt}. \tag{6}$$

One can view the random risk aversion parameter, ω_{nt}, as the sum of a core risk aversion parameter and a random shock, where the random shock is added to the risk aversion parameter instead of the expected utility difference. We will call μ_ω the risk aversion parameter and σ_ω the noise parameter in accordance with this interpretation.[12]

The lack of separability between the index function and the stochastic component in the RP model makes it difficult to analytically obtain choice probabilities. Without further assumptions the choice probability in (6) cannot be simplified to a logistic distribution function like the additive error counterpart in (4) because the index $\Delta EU_{nt}(\omega_{nt})$ is a non-linear function of the random variable, ω_{nt}. One approach is to consider Monte Carlo integration and compute a simulated analogue to (6) by

$$S_{nt}(\mu_\omega, \sigma_\omega) = (1/R) \sum_r \mathbf{I}[\Delta EU_{nt}(\mu_\omega + \sigma_\omega \times e_{nt,r}) > 0], \tag{7}$$

where r = 1, 2, ···, R refers to pseudo-random draws from the standard logistic distribution, and $e_{nt,r}$ is the value of draw r for subject n in task t. The choice probability $S_{nt}(\mu_\omega, \sigma_\omega)$ is easy to simulate once candidate values of μ_ω and σ_ω are known, but it is not amenable to maximum simulated likelihood (MSL) estimation of those two unknown parameters. Given a finite number of pseudo-random draws, the choice probability may be equal to 0, and it is a step function which implies that different candidate values of μ_ω and σ_ω may return the same value of $S_{nt}(\mu_\omega, \sigma_\omega)$. The former drawback implies that the sample log-likelihood value may be undefined, and the latter drawback precludes the use of gradient-based maximization algorithms to compute μ_ω and σ_ω.[13]

The computational problem with the RP model is compounded by multiple integrals, once the unidimensional risk aversion parameter ω_{nt} is replaced with a multidimensional vector of

[12] Some distributional assumptions on ω_{nt} do not allow this type of linear decomposition unless one adopts a contrived distribution of the random shock. For example, suppose that ω_{nt} follows a beta distribution which does not have a location parameter; then, there is no natural way to express ω_{nt} as a sum of a non-zero constant and a random variable. Whether one sees ω_{nt} as the primitive and specifies its distribution, or some core preference parameter as the primitive and defines ω_{nt} as that primitive plus a random shock, they represent two equally valid perspectives on the RP model. Apesteguia and Ballester [2018; §V] adopt the latter perspective while Wilcox [2008; §4.5][2011; Appendix B] adopts the former.

[13] Thus, one cannot use standard techniques, such as Newton-Raphson (NR) and Broyden-Fletcher-Goldfarb-Shanno (BFGS) methods, to maximize the simulated log-likelihood function. The use of gradient-free algorithms may help locate a maximum, but will not solve the failure of identification.

preference parameters. Consider, for example, an EUT model where the utility of each outcome k in lottery L is measured by a distinct parameter u_{Lk}. The expected utility index in (2) is then replaced by

$$EU_L(\mathbf{u}_L) = \sum_k p_{Lk} \times u_{Lk}, \tag{8}$$

where $\mathbf{u}_L$ is a vector of K utility parameters. The additive error model in equation (4) can incorporate this non-parametric approach to the utility function by substituting the index function and retaining the stochastic assumption. However, in the RP model with multidimensional random vectors $\mathbf{u}_{Ant}$ and $\mathbf{u}_{Bnt}$, one must use a multivariate distribution to capture within-individual heterogeneity in choice behavior. Moreover, the random variation in utility parameters is restricted by the maintained assumption of monotone preferences, which makes it difficult to derive analytic choice probabilities even for seemingly tractable pairs of lotteries (Wilcox [2008; §4.5]).

Decision theoretic alternatives to EUT often incorporate complementary sources of risk attitudes alongside utility curvature. For example, in the RDU model, the PWF generates rank-dependent decision weights which may enhance or diminish risk aversion emanating from utility curvature. Using Prelec's [1998] one-parameter PWF, one can extend the EUT model with CRRA utility in (2) to a RDU specification

$$RDU_L(\omega, \varphi) = \sum_k (\pi(P_{Lk} | \varphi) - \pi(P_{L(k+1)} | \varphi)) \times U(m_{Lk} | \omega), \tag{9}$$

where the PWF $\pi(. | \varphi)$ is given by

$$\pi(P | \varphi) = \exp\{-(-\ln(P))^{\varphi}\} \tag{10}$$

and φ is a preference parameter that determines the shape of the PWF. This function is inverse-S shaped when $0 < \varphi < 1$ and S shaped when $\varphi > 1$. The EUT model with non-parametric utility can be similarly generalized to a RDU model by replacing the outcome probabilities with rank-dependent decision weights.

Past studies by Loomes, Moffatt and Sugden [2002] and Wilcox [2008][2011] have focused on a partial RP version of RDU which assumes a non-random PWF and a random utility function. In relation to (9), this special case occurs when φ is a fixed parameter and ω is a random preference parameter ω_{nt}. Our primal estimation method enables us to consider a full RP model that has both parameters vary randomly within an individual. Our method is not restricted by a particular decision theory or functional form, and can be extended to higher dimensions by integrating more random preference parameters.

B. Data

To estimate the EUT and RDU models with CRRA utility, we use data from an artefactual field experiment reported in Andersen, Harrison, Lau and Rutström [2014]. The data set includes 413 subjects from the general adult population in Denmark, and each subject was asked to make choices from 40 distinct pairs of lotteries A and B. Each lottery pair can be written as $A = \{(m_{A1}, (1 - p_2)), (m_{A2}, p_2)\}$ and $B = \{(m_{B1}, (1 - p_2)), (m_{B2}, p_2)\}$ where $m_{B1} < m_{A1} < m_{A2} < m_{B2}$. For each set of prizes $[m_{A1}, m_{A2}, m_{B1}, m_{B2}]$, the probability p_2 varied from 0.1 to 1 in increments of 0.1. We exclude lottery pairs with dominant choices (*i.e.*, with $p_2 = 1$) which do not contribute to the identification of risk preferences in the RP model. There were four prize sets, S_1 = [1600, 2000, 100, 3850], S_2 = [750, 1125, 250, 2000], S_3 = [875, 1000, 75, 2000] and S_4 = [1000, 2250, 50, 4500], where the amounts are in Danish kroner; at the time of the experiment, the exchange rate was close to 5 kroner per US dollar. At the end of the experiment, one of the subject's 40 choices was randomly selected for payment, and each subject had a 10% chance of actually receiving the payment.

To estimate the EUT and RDU models with non-parametric utility, we use data from a lab experiment reported in Harrison and Rutström [2008; §2.6]. The experiment was conducted with a sample of 63 students at the University of Central Florida, who made choices from 60 distinct pairs of lotteries A and B. The two lotteries in each pair had the same set of three prizes $[m_1, m_2, m_3]$ but different probability distributions: $A = \{(m_1, p_{A1})), (m_2, p_{A2}), (m_3, p_{A3})\}$ and $B = \{(m_1, p_{B1})), (m_2, p_{B2}), (m_3, p_{B3})\}$. Each lottery pair can thus be seen as two points in the Marschak-Machina (MM) triangle. Each probability P_{Lk} took a value of 0, 0.13, 0.25, 0.37, 0.5, 0.62, 0.75 or 0.87, and the three probabilities summed to one in each lottery. There were four prize sets denominated in US dollars, S_1 = [5, 10, 15], S_2 = [0, 10, 15], S_3 = [0, 5, 15] and S_4 = [0, 5, 10]. At the end of the experiment, three of the subject's 60 choices were randomly selected for payment.

3. Dual Estimation Method

Before moving to a fully versatile method, we first present a dual method to estimate unidimensional RP models when the underlying preference index function satisfies the single crossing condition. Studies in economic theory often use single crossing conditions to obtain existence proofs and comparative statics results (*e.g.*, Gans and Smart [1996], Athey [2001] and Quah and Strulovici [2012]). We show that the single crossing condition enables one to specify the RP model as a standard econometric model with a linear index function. Our dual estimation method can accommodate any decision theory that represents the subject's evaluation of lottery L by an index function $V_L(\alpha)$ with a unidimensional preference parameter α. We will use EUT with CRRA utility as an example, by taking $EU_L(\omega)$ in equation (2) as $V_L(\alpha)$. The single crossing condition holds if there is a unique value w such that subjects with $\omega > w$ choose one lottery and subjects with $\omega < w$ choose the other lottery. By construction, lottery pairs in the popular multiple price list design by Holt and Laury [2002] and the MM probability triangle induce single crossing in $\Delta EU(\omega)$.[14]

The random coefficient model of interpersonal heterogeneity in choice behavior has played a central role in advancing the empirical literature in many fields, including environmental economics, industrial organization and labor economics (Train [2009]). This model enables one to estimate the between-subject distribution of preferences in the population, and study the mean of the distribution as well as the population shares of subjects fitting particular preference profiles. With our dual method, one can easily combine this modern modeling approach with the RP model.

A. Homogeneous Risk Aversion and Noise

Assume that the random risk aversion parameter ω_{nt} under EUT with CRRA utility satisfies the single crossing condition, and let w_{nt} be the indifference point where the difference in expected

---

[14] Single crossing is a joint property of theory and data since the same utility function may have one or more indifference points depending on the lottery pair in question. In the MPL, each binary choice task is deliberately designed to induce a unique indifference point under EUT with CRRA utility. In the MM triangle, EUT with CRRA utility (and any other utility function) yields linear indifference curves, and satisfies the single crossing condition. The third-order risk apportionment task in Deck and Schlesinger [2010], however, does not induce single crossing under EUT with CRRA utility; this task has two indifference points, one at $\omega = 0$ and the other at $\omega = -1$.

utility is equal to zero. The indifference point w_{nt} is a function of the prizes and probabilities in each pair of lotteries and is therefore an observed characteristic of the decision task. Suppose that $\omega_{nt} > w_{nt}$ implies $\Delta EU_{nt}(\omega_{nt}) < 0$, and $\omega_{nt} < w_{nt}$ implies $\Delta EU_{nt}(\omega_{nt}) > 0$. For example, subject n chooses lottery A (B) in task t if she is more (less) risk averse than w_{nt}. Let y_{nt} denote a choice indicator that is equal to 1 if her choice is B and equal to 0 if it is A.

The unidimensional EUT model in (6) specifies the distribution of ω_{nt} as logistic with mean μ_ω and standard deviation $\sigma_\omega \times 1.81$. We assume for now that every subject has the same "RP urn" and refer to μ_ω and σ_ω as the risk aversion parameter and the noise parameter, respectively. We can exploit the single crossing property of ω_{nt} and show that an analytic solution to the EUT model in (6) is given by

$$L_{nt}(\mu_\omega, \sigma_\omega) = \int \mathbf{I}[\Delta EU_{nt}(\omega_{nt}) > 0] f(\omega_{nt} \,|\, \mu_\omega, \sigma_\omega) d\omega_{nt} = \Lambda((w_{nt} - \mu_\omega)/\sigma_\omega), \tag{11}$$

where $\Lambda(.)$ is the standard logistic distribution function. We can also show that this specification is equivalent to the *pooled* logit model

$$L_{nt}(\beta_0, \beta_1) = \Lambda(\beta_0 + \beta_1 w_{nt}), \tag{12}$$

which can be estimated by running a logit regression of the choice indicator y_{nt} on the independent variable w_{nt} over all subjects n and decision tasks t. The steps involved in moving from the EUT model in (11) to the pooled logit model in (12) are straightforward, at least with hindsight, but the dual link has not been identified in the empirical literature on choice under risk.

The probability that subject n chooses lottery B in task t, $L_{nt}(\mu_\omega, \sigma_\omega)$, is equal to $\Pr(\Delta EU_{nt}(\omega_{nt}) > 0)$. Given the single crossing condition, $\Pr(\Delta EU_{nt}(\omega_{nt}) > 0)$ is equal to the probability that subject n in task t is less risk averse than the indifference point w_{nt}, $\Pr(\omega_{nt} < w_{nt})$. It is useful to write out the "risk aversion plus random shock" interpretation of ω_{nt} explicitly

$$\omega_{nt} = \mu_\omega + \sigma_\omega \times e_{nt}, \tag{13}$$

where e_{nt} is a standard logistic random variable. It follows that $\Pr(\omega_{nt} < w_{nt}) = \Pr(\mu_\omega + \sigma_\omega \times e_{nt} < w_{nt}) = \Pr(e_{nt} < (w_{nt} - \mu_\omega)/\sigma_\omega)$. Hence, the probability that subject n chooses lottery B in task t is the cumulative probability that the random variable e_{nt} is smaller than the standardized difference between the indifference point and risk aversion parameter $(w_{nt} - \mu_\omega)/\sigma_\omega$. One can evaluate these choice probabilities using the standard logistic distribution function $\Lambda(.)$ in (11).

The EUT model in (11) can be indirectly estimated by the pooled logit model in (12), since the latter is equivalent to the former with $\beta_0 = -\mu_\omega/\sigma_\omega$ and $\beta_1 = 1/\sigma_\omega$. One can thus obtain maximum likelihood estimates (MLEs) of β_0 and β_1 in the pooled logit model, and use the results to infer the risk aversion parameter $\mu_\omega = -\beta_0/\beta_1$ and noise parameter $\sigma_\omega = 1/\beta_1$. The invariance property of MLE implies that the transformed parameter estimates from the pooled logit model are equivalent to the direct ML estimates of μ_ω and σ_ω, which can be obtained by a user-written likelihood evaluator. As usual, standard errors of the transformed parameters can be obtained by the delta method.

B. Heterogeneous Risk Aversion and Homogeneous Noise

Since the logistic EUT model in (11) is dual to the pooled logit model in (12), the two models share the same fundamental limitations. First, neither model accounts for panel correlation across repeated choice observations on the same subject. Each choice is modeled as an independent observation, even though it forms part of a set of choices by the same subject. Second, neither model accounts for unobserved heterogeneity in choice behavior across subjects. In the RP model, this translates into the representative agent assumption that every subject has the same urn of random risk aversion parameters.

The random effects (RE) logit model for panel data addresses these two limitations of the pooled logit model, and it is dual to an EUT model that captures between-subject heterogeneity by replacing the fixed coefficient μ_ω with a random coefficient $\mu_{\omega n}$. Suppose that the random risk aversion parameter ω_{nt} is logistically distributed *within* subject n with density $f(\omega_{nt} | \mu_{\omega n}, \sigma_\omega)$. That is, if one compares the RP urns of two subjects, the contents of the two urns have different means but the same standard deviation. Suppose further that the risk aversion parameter $\mu_{\omega n}$ is normally distributed *between* different subjects, $\mu_{\omega n} \sim N(E[\mu_{\omega n}], SD[\mu_{\omega n}]^2)$, *i.e.* each subject n in the population has her own value of $\mu_{\omega n}$, and the between-subject mean and standard deviation of those values in the population are equal to $E[\mu_{\omega n}]$ and $SD[\mu_{\omega n}]$, respectively. One can interpret $N(E[\mu_{\omega n}], SD[\mu_{\omega n}]^2)$ as the population distribution of risk attitudes in this model.

The probability that subject n chooses lottery B in task t is specified as

$$L_{nt}(\mu_{\omega n}, \sigma_\omega) = \int \mathbf{I}[\Delta EU_{nt}(\omega_{nt}) > 0] f(\omega_{nt} | \mu_{\omega n}, \sigma_\omega) d\omega_{nt} \tag{14}$$
$$= \Lambda((w_{nt} - \mu_{\omega n})/\sigma_\omega) \text{ with } \mu_{\omega n} \sim N(E[\mu_{\omega n}], SD[\mu_{\omega n}]^2),$$

where $E[\mu_{\omega n}]$, $SD[\mu_{\omega n}]$ and σ_ω are parameters to be estimated.[15] This model is equivalent to the RE logit model that augments the pooled logit model in (12) with a normally distributed error component υ_n. The probability that subject n chooses lottery B in task t in the RE logit model is

$$L_{nt}(\beta_0, \beta_1, \upsilon_n) = \Lambda(\beta_0 + \beta_1 w_{nt} + \upsilon_n) \text{ with } \upsilon_n \sim N(0, \sigma_0^2), \tag{15}$$

where β_0, β_1 and σ_0 are parameters to be estimated. This model has a homogeneous slope coefficient β_1 and a heterogeneous intercept, $\alpha_n = (\beta_0 + \upsilon_n)$, which is normally distributed between subjects, $\alpha_n \sim N(\beta_0, \sigma_0^2)$. The error component υ_n captures between-subject heterogeneity around the mean intercept β_0, and the standard deviation σ_0 measures the extent of that heterogeneity.

When the RE logit model in (15) is interpreted as a random intercept model, the dual link to the EUT model in (14) with between-subject heterogeneity in risk aversion becomes apparent. The slope coefficient β_1 is equivalent to $1/\sigma_\omega$, and the random intercept $\alpha_n = (\beta_0 + \upsilon_n)$ is equivalent to $-\mu_{\omega n}/\sigma_\omega$. Thus, one can use the RE logit estimates of β_0, β_1 and σ_0 to compute $\sigma_\omega = 1/\beta_1$, $E[\mu_\omega] = -\beta_0/\beta_1$ and $SD[\mu_\omega] = \sigma_0/\beta_1$, and apply the delta method to obtain standard errors of the transformed parameters.

It is useful to clarify the meaning of "random" since the term has been used to describe two different kinds of random variations. The term "random" in the RP model refers to the random variable ω_{nt} with density $f(\omega_{nt} | \mu_{\omega n}, \sigma_\omega)$ that describes the seemingly unstable risk attitude of subject n over decision tasks. It captures *within-subject* variation that vanishes as the noise parameter σ_ω tends to zero. The term "random" in the random coefficient and RE models refers to the use of random variables $\mu_{\omega n}$ and α_n to describe interpersonal heterogeneity in the underlying population. It captures *between-subject* variation which is present even if every subject makes deterministic choices according to the non-stochastic version of EUT. The model specification in (14) displays both kinds of random variations.

---

[15] As Train [2009; §6] and Wooldridge [2010; p.613] explain, the sample log-likelihood function can be specified in terms of $E[\mu_{\omega n}]$, $SD[\mu_{\omega n}]$ and σ_μ by "integrating out" $\mu_{\omega n}$.

C. Heterogeneous Risk Aversion and Noise

The EUT model in (14) allows for between-subject heterogeneity in the risk aversion parameter, $\mu_{\omega n}$, but not in the noise parameter, σ_ω. We can capture between-subject heterogeneity in the noise parameter by introducing a second random coefficient, $\sigma_{\omega n}$, that replaces the fixed coefficient σ_ω. Assume that the random risk aversion parameter ω_{nt} is logistically distributed *within* subject n with density $f(\omega_{nt} | \mu_{\omega n}, \sigma_{\omega n})$. The subject now has an individual-specific risk aversion parameter, $\mu_{\omega n}$, *and* an individual-specific noise parameter, $\sigma_{\omega n}$. Assume further that the risk aversion and log-noise parameters, $\mu_{\omega n}$ and $\ln(\sigma_{\omega n})$, are jointly normally distributed *between* subjects, $[\mu_{\omega n}, \ln(\sigma_{\omega n})]' \sim MVN(\mathbf{b}_{EUT}, \mathbf{V}_{EUT})$. Each subject n thus has a distinct pair of values $\mu_{\omega n}$ and $\ln(\sigma_{\omega n})$, and the distribution of these bivariate parameters has between-subject mean $\mathbf{b}_{EUT}$ and covariance matrix $\mathbf{V}_{EUT}$. The probability that subject n chooses lottery B in task t is

$$\begin{aligned} L_{nt}(\mu_{\omega n}, \sigma_{\omega n}) &= \int \mathbf{I}[\Delta EU_{nt}(\omega_{nt}) > 0] f(\omega_{nt} | \mu_{\omega n}, \sigma_{\omega n}) d\omega_{nt} \\ &= \Lambda((w_{nt} - \mu_{\omega n})/\sigma_{\omega n}) \text{ with } [\mu_{\omega n}, \ln(\sigma_{\omega n})]' \sim MVN(\mathbf{b}_{EUT}, \mathbf{V}_{EUT}), \end{aligned} \tag{16}$$

where the mean vector $\mathbf{b}_{EUT}$ and covariance matrix $\mathbf{V}_{EUT}$ are parameters to be estimated.

The more general EUT model in (16) is dual to the mixed logit model in the willingness-to-pay (WTP) space (Train and Weeks [2005]). Although less known compared to pooled logit and RE logit models, the WTP space model is a standard econometric model that one can readily estimate in popular statistical packages.[16] In the WTP space model, the probability that subject n chooses lottery B in task t is

$$\begin{aligned} L_{nt}(\alpha_n, \lambda_n) &= \exp((\alpha_n + w_{nt}) \times \lambda_n) \ / \ [\exp(0) + \exp((\alpha_n + w_{nt}) \times \lambda_n)] \\ &= \Lambda((\alpha_n + w_{nt}) \times \lambda_n) \text{ with } [\alpha_n, \ln(\lambda_n)]' \sim MVN(\mathbf{b}_{WTP}, \mathbf{V}_{WTP}) \end{aligned} \tag{17}$$

where the mean vector $\mathbf{b}_{WTP}$ and the covariance matrix $\mathbf{V}_{WTP}$ are parameters to be estimated.[17] The first equality in (17) is algebraically redundant but conveys useful operating advice. Since the WTP

[16] The GMNL model of Fiebig, Keane, Louviere and Wasi [2010] is also a dual model since it algebraically nests the WTP space model (Greene and Hensher [2010; p.416]). Estimation programs for the GMNL model are also widely available in different software packages.

[17] The WTP space model originates from the non-market valuation literature. Suppose that the independent variable w_{nt} is equal to the price of good A minus that of good B. If the subject has constant marginal utility of money that is proportional to the precision parameter, λ_n, the random intercept α_n may be interpreted as subject n's WTP for acquiring good B rather than good A.

space model is primarily intended for multinomial choice applications, the available programs evaluate a separate index function for each alternative in a choice set. In our dual method, one can set the index functions for lottery A to 0 and lottery B to $(\alpha_n + w_{nt}) \times \lambda_n$.

Comparing the EUT model in (16) to the WTP space model in (17) shows that $\alpha_n = -\mu_{\omega n}$ and $\lambda_n = 1/\sigma_{\omega n}$. The latter implies that $\ln(\lambda_n) = -\ln(\sigma_{\omega n})$. Since $[\alpha_n, \ln(\lambda_n)] = [-\mu_{\omega n}, -\ln(\sigma_{\omega n})]$, it follows that $\mathbf{b}_{EUT} = -\mathbf{b}_{WTP}$ and $\mathbf{V}_{EUT} = \mathbf{V}_{WTP}$. We can thus multiply the estimates of $\mathbf{b}_{WTP}$ by -1 to compute the estimates of $\mathbf{b}_{EUT}$, and use $\mathbf{V}_{WTP}$ directly in place of $\mathbf{V}_{EUT}$.

This dual link is computationally very convenient. The use of random coefficients to capture interpersonal preference heterogeneity has become a well-established practice in environmental economics, health economics, marketing science and transportation research because standard estimation programs for the family of mixed logit models enable one to estimate the structural models of interest in those fields. By contrast, the structural models of risk preferences, seemingly, do not fit in with the mixed logit structure and the empirical analysis of those models typically relies on fixed coefficient specifications. The dual link that we have identified provides convenient means to apply recent advances in discrete choice methods to behavioral research.

D. Empirical Illustration of Dual Estimation

We illustrate dual estimation of the EUT model with CRRA utility using data from the field experiment reported in Andersen, Harrison, Lau and Rutström [2014]. Prior to the estimation, we numerically compute the indifference point, w_{nt}, that solves $\Delta EU_{nt}(w_{nt}) = 0$ for each lottery pair in the experiment. Panel A in Table 1 reports the estimation results for the EUT model in (11), which assumes that every subject carries the same urn of RP parameters with logistic density $f(\omega_{nt} | \mu_\omega, \sigma_\omega)$. We estimate the dual pooled logit model in (12) by using the *logit* command in *Stata* to regress the choice indicator y_{nt} on the indifference point w_{nt}. The estimated intercept (β_0) and coefficient on w_{nt} (β_1) are equal to -0.932 and 1.740, respectively. Unless stated otherwise, all estimated coefficients and transformed parameters have p-values < 0.001. We can use the estimated coefficients to derive the relative risk aversion parameter $\mu_\omega = -\beta_0/\beta_1 = 0.535$ and noise parameter $\sigma_\omega = 1/\beta_1 = 0.575$. We thus find that the representative agent generally is risk averse with significant variation in choice behavior.

Panel B in Table 1 reports the results for the EUT model in (14) which accommodates between-subject heterogeneity in risk aversion. The RP urn of subject n is $f(\omega_{nt} | \mu_{\omega n}, \sigma_{\omega})$ where the risk aversion parameter $\mu_{\omega n}$ is now subject-specific and assumed to be normally distributed between subjects, with population mean $E[\mu_{\omega n}]$ and standard deviation $SD[\mu_{\omega n}]$. Since this model is dual to the RE logit model, we regress y_{nt} on w_{nt} using the *xtlogit, re* command in *Stata*, which applies a Gauss-Hermite quadrature to integrate out the normal error component υ_n in (15). The random intercept is normally distributed with an estimated population mean (β_0) of −1.327 and a standard deviation (σ_0) of 1.778, while the estimated coefficient β_1 on w_{nt} is equal to 2.411. The transformed parameters $E[\mu_{\omega n}] = -\beta_0/\beta_1$, $SD[\mu_{\omega n}] = \sigma_0/\beta_1$ and $\sigma_\omega = 1/\beta_1$ are equal to 0.550, 0.737 and 0.415, respectively. Hence, the coefficient of RRA is estimated to be 0.550 on average, with significant between- and within-subject variation, and we find that 77.2% of the population are risk averse.[18]

Panel C in Table 1 reports the results for the EUT model in (16) which accommodates between-subject heterogeneity in both risk aversion and noise. The RP urn of subject n is $f(\omega_{nt} | \mu_{\omega n}, \sigma_{\omega n})$, where both risk aversion ($\mu_{\omega n}$) and noise ($\sigma_{\omega n}$) parameters are subject-specific. The model assumes that $\mu_{\omega n}$ and $\ln(\sigma_{\omega n})$ are jointly normally distributed between subjects. We estimate the dual mixed logit model in WTP space and use the *mixlogitwtp* command in *Stata* by Hole [2015], which applies simulation to integrate out the joint normal random coefficients α_n and $\ln(\lambda_n)$ in (17). Our simulated integrals are based on 100 Halton draws per subject. The estimated intercept α_n is normally distributed between subjects with a mean (β_0) of −0.519 and a standard deviation (σ_0) of 0.846. The log-precision parameter $\ln(\lambda_n)$ is normally distributed between subjects with a mean (τ) of 1.304 and a standard deviation (σ_τ) of 0.997. We can use these estimates to evaluate the between-subject distribution of the risk aversion parameter $\mu_{\omega n}$ and log-noise parameter $\ln(\sigma_{\omega n})$ in the population. The results suggest that $E[\mu_{\omega n}] = -\beta_0 = 0.519$ and $SD[\mu_{\omega n}] = \sigma_0 = 0.846$, while $E[\ln \sigma_{\omega n}]$

[18] Since the between-subject distribution of $\mu_{\omega n}$ is normal, the population share of risk averse subjects is equal to $\Pr(\mu_{\omega n} > 0) = 1-\Phi(-E[\mu_{\omega n}]/SD[\mu_{\omega n}])$, where $\Phi(.)$ is the standard normal distribution function. The estimated share is significantly greater than 0.50 or 50% (one-sided *p*-value < 0.001).

$= -\tau = -1.304$ and SD[ln $\sigma_{\omega n}$] $= \sigma_0 = 0.997$.[19] Hence, the estimated population mean of relative risk aversion is equal to 0.519, and we find that 73% of the population are risk averse.[20]

We next estimate the EUT model with non-parametric utility using data from the lab experiment reported in Harrison and Rutström [2008; §2.6]. When all lottery pairs are defined over the same set of three prizes, this model has a unidimensional parameter, which is required for dual estimation: Since the utility function is unique up to an affine transformation, we can normalize utility u_{1nt} of the lowest prize to zero, set u_{3nt} of the highest prize to one, and retain u_{2nt} of the medium prize as a free parameter. The indifference point $\bar{u}_{2nt}$ then has a closed-form solution, $\bar{u}_{2nt} = (p_{B3nt} - p_{A3nt}) / [(p_{B3nt} - p_{A3nt}) + (p_{B1nt} - p_{A1nt})]$. We assume that the RP urn for u_{2n} is a logit-logistic distribution, which can capture a variety of shapes: $u_{2nt} = \Lambda(\zeta_{nt})$, where the inverted utility ζ_{nt} is logistically distributed within subjects with density $f(\zeta_{nt} | \mu_\zeta, \sigma_\zeta)$.[21] Since the data set includes four sets of three prizes, we can fit a separate model to each prize set. We focus here on the prize set with {\$0, \$5, \$10}. The constant distance between outcomes in the prize set allows one to associate positive values of μ_ζ with diminishing marginal utility, hence risk aversion. A positive value of μ_ζ implies $u_{2nt} = \Lambda(\mu_\zeta) > 0.5$, which in turn implies $(u_{3nt} - u_{2nt}) < (u_{2nt} - u_{1nt})$ since $u_{1nt} = 0$ and $u_{3nt} = 1$.[22]

Figure 1 displays the representative agent's RP urn that is derived from the dual pooled logit model. The top panel shows that the representative agent is risk averse: the estimated risk aversion (μ_ζ) and noise (σ_ζ) parameters in the logistic distribution of inverted utility ζ_{nt} are equal to 1.066 and 0.693, respectively, which implies that the estimated utility of \$5 is $\Lambda(1.066) = 0.744$ in the absence of behavioral noise. The bottom panel shows that the μ_ζ and σ_ζ values leads to a logit-logistic distribution of utility $u_{2nt} = \Lambda(\zeta_{nt})$ that is visibly skewed to the left. To capture between-subject

[19] Using the analytic properties of the log-normal distribution, we can also evaluate the population moments of $\sigma_{\omega n}$. We find that E[$\sigma_{\omega n}$] = 0.446, SD[$\sigma_{\omega n}$] = 0.582, and $\varrho(\mu_{\omega n}, \sigma_{\omega n}) = -0.042$ (*p*-value = 0.570).

[20] The between-subject correlation coefficient for α_n and ln(λ_n), $\varrho_{\tau 0}$, is equal to -0.055 with *p*-value = 0.564. Since the correlation coefficient for $\mu_{\omega n}$ and ln($\sigma_{\omega n}$), $\varrho[\mu_{\omega n}, \ln \sigma_{\omega n}]$, is identical to $\varrho_{\tau 0}$, we do not find that more risk averse subjects have more or less variation in risk preferences.

[21] As Smithson and Shou [2017; p.423] graphically show, a logit-logistic distribution can approximate a wide range of shapes (*e.g.*, uniformity, unimodality, bimodality, and left and right skewness) without imposing any particular shape restriction *a priori*.

[22] The detailed results are reported in Table B1 in Appendix B.

heterogeneity, the fixed coefficients μ_ζ and σ_ζ can be replaced with subject-specific random coefficients $\mu_{\zeta n}$ and $\sigma_{\zeta n}$. The results from the dual RE logit and WTP space models are reported in Appendix B, and we find again that the average subject is risk averse, and that approximately 90% of the subjects are risk averse.

One can easily control for observed heterogeneity in risk preferences by adding explanatory variables to the dual econometric models. One can also apply the dual estimation method to one-parameter formulations of Yaari's [1987] dual theory of choice under risk, and with the latent class conditional logit model which is perhaps the most widely used alternative to the mixed logit model. We discuss these applications in Appendix B.

4. Primal Estimation Method

We now turn to a versatile approach to estimating the primal representation of the RP model. The estimation method can accommodate any theory that represents the subject's evaluation of lottery L by an index function $V_L(\boldsymbol{\alpha})$, where the argument $\boldsymbol{\alpha}$ is a multidimensional vector of preference parameters. We do not impose any restrictions on the dimensionality and auxiliary properties of the preference parameters, including the single crossing condition. The method can be applied to RP versions of EUT with multiple utility parameters, such as the model specification in (8) that takes a non-parametric approach to the utility function, and extensions of EUT that account for multiple sources of risk attitudes, such as the RDU specification in (9) and (10).

The main challenge in estimating RP models is that a general purpose simulator like equation (7) generates a step function which does not lend itself to numerical maximization.[23] We solve the problem by applying the kernel-smoothed frequency simulator that McFadden [1989] developed to address the same issue for multinomial probit models. Our primal method requires a user-written likelihood evaluator and does not share the remarkable accessibility of the dual method in section 3. However, the likelihood evaluator of the RP model with fixed coefficients and kernel smoothing is

---

[23] Simulated likelihood functions are also required when additive error models are combined with random coefficients. As Revelt and Train [1998] explain, however, the general purpose simulator is not a step function in those cases: When the additive error term is assumed to follow a parametric distribution which produces a smooth link function (*e.g.*, logit), the simulated likelihood function is a smooth function of the parameters to be estimated.

not harder to program than that of the additive error model with random coefficients (*e.g.*, Harrison, Lau and Yoo [2020]), as the two classes of evaluators share similar algebraic structures. This similarity also helps one combine the RP and random coefficient models, by adding the "random coefficient" integral at the subject level to complement the "RP" integral at the task level.

A. Kernel-Smoothed RP Choice Probabilities

Consider first the simple EUT model with CRRA utility in (5) and (6). This model assumes that every subject carries the same RP urn, which follows a logistic distribution with density $f(\omega_{nt} | \mu_\omega, \sigma_\omega)$. When the single crossing condition is violated, one cannot apply the dual estimation method and must instead rely on user-written likelihood evaluators. It is difficult, however, to obtain analytic solutions to RP choice probabilities in general, and a general purpose simulator of choice probabilities based on the simple frequency method generates a likelihood function that is not amenable to numerical maximization.

We use a perturbation strategy by McFadden [1989; p.1001] to construct a kernel-smoothed frequency simulator that more easily lends itself to numerical maximization. Let the latent dependent variable representation of the RP model in (5), $y_{nt}^* = \Delta EU_{nt}(\omega_{nt})$, be perturbed by a *contaminating* disturbance term that is logistically distributed with zero mean and standard deviation $\varkappa \times 1.81$. The model is written as

$$y_{nt}^* = \Delta EU_{nt}(\omega_{nt}) + \varkappa \times v_{nt}, \tag{18}$$

where v_{nt} is a standard logistic random variable and $\varkappa$ is a positive constant that is selected prior to estimation. The disturbance $\varkappa \times v_{nt}$ is *contaminating* in the sense that it is not part of the assumed stochastic choice process. It represents an intentional specification error which is added to obtain a perturbed model that is easier to simulate than the target model.[24] By selecting a suitably small value of $\varkappa$, one can approximate the target model to a desired degree of accuracy. McFadden and Train [2000; p.451] provide an approximation theorem which formally supports this intuition.

[24] The perturbed RP model in (18) is algebraically equivalent to a hybrid stochastic choice model that combines the RP model with the additive behavioral error model. The hybrid model would have $\varkappa$ as a behavioral noise parameter to be estimated, whereas the RP model has it as a measure of approximation error to be selected prior to estimation.

The logistic distribution of the contaminating disturbance term implies that the expected value of $\mathbf{I}[\Delta EU_{nt}(\omega_{nt}) + \varkappa \times v_{nt} > 0]$ conditional on ω_{nt} takes the same functional form as the additive error model in (4). Using the law of iterative expectations, the probability that subject n chooses lottery B in task t is then equal to

$$L_{nt}(\mu_\omega, \sigma_\omega) = \int \Lambda(\Delta EU_{nt}(\omega_{nt})/\varkappa) f(\omega_{nt} \mid \mu_\omega, \sigma_\omega) d\omega_{nt}, \tag{19}$$

which can be simulated as

$$S_{nt}(\mu_\omega, \sigma_\omega) = (1/R)\textstyle\sum_r \Lambda(\Delta EU_{nt}(\mu_\omega + \sigma_\omega \times e_{nt,r})/\varkappa), \tag{20}$$

where r = 1, 2, ···, R are pseudo-random draws from the standard logistic distribution, and $e_{nt,r}$ is the value of draw r. The perturbed choice probability in (19) and its simulated analogue in (20) are identical to the RP choice probability in (6) and the general purpose simulator in (7), except that the standard logistic distribution function, $\Lambda(\Delta EU_{nt}(.)/\varkappa)$, has replaced the indicator function, $\mathbf{I}(\Delta EU_{nt}(.) > 0)$. From a computational perspective, one can consider $\Lambda(\Delta EU_{nt}(.)/\varkappa)$ as the logit kernel with smoothing factor $\varkappa$ and directly motivate (20) as a kernel-smoothed version of (7), instead of seeing it as a simulator of the perturbed choice probability.

To estimate the EUT model with homogeneous risk aversion μ_ω and noise σ_ω, we can simulate the sample likelihood function by taking the likelihood function of the pooled logit model (*e.g.*, Wooldridge [2010; §15.8.1]) as a template and use the kernel-smoothed simulator in (20) to replace the standard logit probability.[25] Since the kernel-smoothed simulator is a finite sum of logistic distribution functions, it returns algebraically positive probabilities and is twice continuously differentiable in μ_ω and σ_ω. We can thus use conventional maximization algorithms to compute MSL estimates of the two fixed coefficients.

To account for between-subject heterogeneity, we can replace μ_ω and σ_ω with subject-specific random coefficients, as in (16), which assumes $[\mu_{\omega n}, \ln(\sigma_{\omega n})]' \sim MVN(\mathbf{b}_{EUT}, \mathbf{V}_{EUT})$. We can simulate the likelihood function of this random coefficient model by adding one more layer of simulation that integrates the pooled RP likelihood function over the joint distribution of $\mu_{\omega n}$ and $\ln(\sigma_{\omega n})$. One

[25] Since the perturbation strategy originates from the multinomial probit literature, we can adapt our primal method to RP models for multinomial or rank-ordered choice data. We only need to replace the binomial logit link with the multinomial logit or rank-ordered logit link, and divide each of the multiple index functions in the model by the same smoothing factor $\varkappa$.

can use the mixed logit likelihood for panel data (*e.g.*, Train [2009; §6.7]) and replace the pooled multinomial logit likelihood therein with the pooled RP likelihood. The simulated likelihood functions for both fixed and random coefficient versions of the RP-EUT model with CRRA utility are specified in Appendix A.

B. Multidimensional RP Models

Let $\Delta V_{nt}(\boldsymbol{\alpha}_{nt})$ denote an index that represents the subject's relative valuation of two lotteries as a function of a multidimensional vector of RP parameters $\boldsymbol{\alpha}_{nt}$, where the exact form of the index may vary across theory and data. Suppose that $\Delta V_{nt}(\boldsymbol{\alpha}_{nt}) > 0$ corresponds to the choice of lottery B, and let $f(\boldsymbol{\alpha}_{nt} | \boldsymbol{\theta})$ denote the joint density of $\boldsymbol{\alpha}_{nt}$ which is characterized by the distributional parameters in $\boldsymbol{\theta}$. The probability that subject n chooses lottery B in task t is simulated by

$$S_{nt}(\boldsymbol{\theta}) = (1/R)\textstyle\sum_r \Lambda(\Delta V_{nt}(\boldsymbol{\alpha}_{nt,r})/\varkappa), \tag{21}$$

where $\boldsymbol{\alpha}_{nt,r}$ refers to draw r of $\boldsymbol{\alpha}_{nt}$ from $f(\boldsymbol{\alpha}_{nt} | \boldsymbol{\theta})$. The likelihood function of this generic RP model can be simulated in the same way as the EUT model in (20) by using the pooled logit likelihood function as a template. The likelihood function for the extended model with random coefficients, which accounts for between-subject heterogeneity in $\boldsymbol{\theta}$, can be simulated in similar fashion by using the mixed logit likelihood function for panel data as a template. We describe this procedure in Appendix A using the RDU model with CRRA utility as an example. Our primal estimation method can be extended to other classes of RP models and distributional assumptions, and we are not aware of other estimation methods which are equally versatile.[26]

---

[26] Our notation of $\Delta V_{nt}(\boldsymbol{\alpha}_{nt})$ also allows the decision theory of interest to vary across decision tasks t. Thus, one can apply the kernel-smoothed simulator in joint estimation of risk and time preferences, by specifying a RP model of multidimensional risk parameters for the set of lottery choice tasks under risk and a RP model of multidimensional discounting parameters for the set of intertemporal choice tasks, with two different subsets of $\boldsymbol{\alpha}_{nt}$ becoming relevant to the two different sets of decision tasks.

Apesteguia, Ballester and Gutierrez-Daza [2020] estimate a two-dimensional RP model, which comprises EUT with CRRA utility for choice under risk and exponential discounting (EXP) for intertemporal choice behavior. Their likelihood evaluator replaces a continuous density function of the RP parameters with purpose-built discrete probability masses, making it difficult to interpret the estimates in relation to the continuous density function of interest. Jagelka [2020] estimates the same type of two-dimensional RP model. He constructs a specialized likelihood evaluator which relies on the single crossing condition, just like earlier RP applications in standalone estimation of EUT and EXP specifications (*e.g.*, Loomes, Moffatt and Sugden [2002] and Apesteguia and Ballester [2018]).

Given a non-parametric approach to the utility function like (8), the multidimensional RP parameters in $\boldsymbol{\alpha}_{nt}$ include utility levels of alternative prizes. This type of utility specification poses a particular challenge. Consider, for example, the EUT model with four utility levels, such that one retains two RP parameters after normalizing the utility of the worst and best prizes. The specification of RP urns in this case is complicated by the requirement that every realization of the random vector $\boldsymbol{\alpha}_{nt}$ must satisfy monotone preferences, $u_{1nt} = 0 < u_{2nt} < u_{3nt} < u_{4nt} = 1$. Wilcox [2008; §4.5] provides an analytic example which illustrates the near impossibility of finding a tractable statistical distribution that satisfies such inter-parameter inequality constraints.

We can address this challenge by adopting a multinomial distribution. Let $\Delta u_{knt} = u_{knt} - u_{(k-1)nt}$ denote the marginal utility between two prizes, k and k−1. We choose the within-subject distribution of u_{2nt} and u_{3nt} indirectly by specifying the primitive joint distribution of Δu_{2nt} and Δu_{3nt}. Assume that the marginal utility of prize k follows a logit-logistic distribution

$$\Delta u_{knt} = \exp(\zeta_{knt}) \ / \ [1 + \exp(\zeta_{2nt}) + \exp(\zeta_{3nt})], \tag{22}$$

where ζ_{knt} is a logistic random variable with density $f(\zeta_{knt} \mid \mu_{\zeta k}, \sigma_{\zeta k})$ for $k \in \{2, 3\}$ and $\zeta_{4nt} = 0$. The utility of prizes $k \in \{2, 3\}$ can be then computed as a sum of marginal utilities

$$u_{2nt} = \Delta u_{2nt} \text{ and } u_{3nt} = \Delta u_{2nt} + \Delta u_{3nt}, \tag{23}$$

so that $u_{1nt} = 0 < u_{2nt} < u_{3nt} < u_{4nt} = 1$, by construction. This RP model with non-parametric utility is attractive because it uses logit-logistic distributions which are known for their ability to display a variety of shapes, and also because it can be easily extended to higher dimensions by expanding the denominator in the multinomial logit formula in (22). We can also combine the RP model with between-subject heterogeneity, by specifying the mean and scale parameters in each density $f(. \mid \mu_{\zeta k}, \sigma_{\zeta k})$ as random coefficients.

Selecting an appropriate smoothing factor $\varkappa$ is not straightforward (Train [2009; §5.6.2]).[27]

[27] A smaller value of $\varkappa$ (*i.e.,* less smoothing) makes the value of the logit kernel, $\Lambda(\Delta V_{nt}(\boldsymbol{\alpha}_{nt,r})/\varkappa)$, more similar to the value of the indicator function, $\mathbf{I}(\Delta V_{nt}(\boldsymbol{\alpha}_{nt,r}) > 0)$. Once they become too similar, $\Lambda(\Delta V_{nt}(\boldsymbol{\alpha}_{nt,r})/\varkappa)$ will induce the same numerical problems as $\mathbf{I}(\Delta V_{nt}(\boldsymbol{\alpha}_{nt,r}) > 0)$. While gradually reducing $\varkappa$ until the numerical problems occurs may sound appealing, this mistakes the general purpose simulator in (7) for the target of approximation; the kernel-smoothed simulator aims to approximate the RP choice probability in (6). When the number of pseudo-random draws R is arbitrarily large, the general purpose simulator returns the RP choice probability and selecting a smaller value of $\varkappa$ thus improves the quality of approximation. However, at a fixed value of R, the general purpose simulator may deviate from the RP choice probability due to

Since $\varkappa$ enters the smoothing kernel as a ratio, $\Delta V_{nt}(\boldsymbol{\alpha}_{nt,r})/\varkappa$, the effect of $\varkappa$ is sensitive to the scale of $\Delta V_{nt}(.)$. We set this scale by normalizing the utility of the best prize in the sample to unity. In the literature on semi-parametric estimation of discrete choice models, a factor of $\varkappa = 1/(N \times T)^{0.2}$ is often used to smooth the maximum score estimator (*e.g.,* Horowitz [1992] and Yan and Yoo [2019]), where $N \times T$ is the total number of observations. We have collated results from multiple estimations with $\varkappa = \# /(N \times T)^{0.2}$, using multiplicative factors $\# \in \{0.01, 0.05, 0.1, 0.25, 0.5, 1, 2\}$, and find that values of $\varkappa$ between 0.01 and 0.02 work best in terms of convergence and, where applicable, approximation of the dual estimates in section 3. While the latter results are only available for models of unidimensional risk preferences, they provide useful benchmarks because the dual likelihood functions do not include the contaminating disturbance term.

C. Empirical Illustration: Primal Estimation of RDU with CRRA Utility

We use our primal method to estimate the RDU model where each subject has two RP urns: one urn with values of the random utility parameter ω_{nt} and another urn with values of the random probability weighting parameter φ_{nt}. We first consider a representative agent model where every subject has the same pair of RP urns. The utility urn has logistic density $f(\omega_{nt} | \mu_\omega, \sigma_\omega)$ and the PWF urn has normal density $g(\ln(\varphi_{nt}) | m_\varphi, s_\varphi)$, where m_φ and s_φ is the mean and standard deviation of $\ln(\varphi_{nt})$.[28] The upper panel in Figure 2 displays the estimated RP urns for the representative agent. When $\varkappa = 0.015$, the logistic density of the utility urn $f(\omega_{nt} | \mu_\omega, \sigma_\omega)$, is generated by $\mu_\omega = 0.635$ and $\sigma_\omega = 0.643$. We thus observe more concavity in the utility function (as measured by μ_ω) relative to the EUT model. The normal density of the probability weighting urn $g(\ln(\varphi_{nt}) | m_\varphi, s_\varphi)$ is generated by $m_\varphi = 0.224$ (*p*-value = 0.004) and $s_\varphi = 1.017$. Since ln(1) is equal to 0, the estimated value of m_φ indicates that the representative agent's PWF is S-shaped, with the shape parameter equal to $\exp(m_\varphi)$

---

simulation noise and minimal smoothing is not necessarily desirable. Some smoothing may reduce simulation noise and improve the quality of approximation, just as a kernel density may approximate a true density better than a histogram does.

[28] We use 100 Halton draws per choice observation to simulate the RP component of each model and, where applicable, another set of 100 Halton draws per subject to simulate the random coefficient component. Bhat [2001] finds that 100 Halton draws provide about the same level of accuracy as 2,000 pseudo-random draws in a Monte Carlo study of discrete choice models. The difference is attributed to the deliberate construction of Halton draws to provide good coverage of the parametric space.

= 1.251. The two preference parameters (μ_ω and m_φ) are significantly different from zero, and we reject the null hypotheses of linear utility and linear PWF at the 1% significance level. The two noise parameters (σ_ω and s_φ) are also significantly different from zero. We obtain qualitatively similar findings for the two adjacent values of $\varkappa$ (0.037 and 0.007).[29]

In the lower panel of Figure 2, we extend the RDU model to capture between-subject heterogeneity in both RP urns. We allow each subject to have unique RP urns $f(\omega_{nt} | \mu_{\omega n}, \sigma_{\omega n})$ and $g(\ln(\varphi_{nt}) | m_{\varphi n}, s_{\varphi n})$ and specify the four subject-specific parameters, $\mu_{\omega n}$, $\sigma_{\omega n}$, $m_{\varphi n}$, and $s_{\varphi n}$ as independent random coefficients.[30] We focus on $\varkappa = 0.015$, as the two adjacent values yield similar results. The utility curvature parameter, $\mu_{\omega n}$, is normally distributed with population mean $E[\mu_{\omega n}] = 0.558$ and standard deviation $SD[\mu_{\omega n}] = 0.697$. The results imply that the average subject, along with 78.8% of the decision makers, has a concave utility function. The log shape parameter, $m_{\varphi n}$, is normally distributed with $E[m_{\varphi n}] = 0.066$ (*p*-value = 0.469) and $SD[m_{\varphi n}] = 1.174$, and we cannot reject the hypothesis that the population is equally divided between those with inverse-S and S shaped PWFs. The estimated between-subject distribution of $m_{\varphi n}$ implies that the population median and mean of the shape parameter φ_{nt} are equal to 1.068 (*p*-value = 0.484) and 2.163 (*p*-value < 0.001), respectively, where each *p*-value refers to the hypothesized value of unity that simplifies RDU to EUT.[31] Thus, the S-shaped weighting function in the representative agent model qualitatively reproduces the shape of the average weighting function in the heterogenous population rather than that of the median weighting function.

[29] The detailed results are reported in Appendix C. We also use our primal method to estimate the unidimensional EUT models in section 3, and the estimated distributions of the risk aversion parameter are displayed in Figure C1. The primal estimates of risk aversion and between-subject heterogeneity are similar to the dual estimates in section 3. We observe some variation in the estimated noise parameter for different values of $\varkappa$ in the simple EUT model with no between-subject heterogeneity, but the divergence in the noise parameter between the primal and dual models disappears at $\varkappa = 0.037$.

[30] The numerical optimizer failed to obtain a solution to the full random coefficient model with a 4×4 covariance matrix. While it may be possible to estimate intermediate models that retain a subset of the covariance parameters (*e.g.*, Layton and Brown [2000]), we do not pursue this route.

[31] The median is equal to $\exp(E[m_{\varphi n}])$ and the mean is equal to $\exp(E[m_{\varphi n}] + SD[m_{\varphi n}]^2/2)$.

D. Empirical Illustration: Primal Estimation with Non-parametric Utility

We next consider the EUT model with non-parametric utility and estimate a joint model for all available prize sets, {\$5, \$10, \$15}, {\$0, \$10, \$15}, {\$0, \$5, \$15} and {\$0, \$5, \$10}. Despite the prominence of RP-EUT models with non-parametric utility in theoretic formulations such as Gul and Pesendorfer [2006], there is no available method to estimate this structural model in a general data environment.[32] Our primal method provides the missing tool.

Figure 3 reports the representative agent's RP urns for marginal utilities of \$5. In relation to (22), the inverted marginal utility from \$0 to \$5 (\$5 to \$10) in the top-left (bottom-left) panel refers to the logistic random variable ζ_{2nt} (ζ_{3nt}), and the corresponding marginal utility in the top-right (bottom-right) panel refers to the logit-logistic random variable Δu_{2nt} (Δu_{3nt}).[33] When $\varkappa = 0.010$, the estimated means of ζ_{2nt} and ζ_{3nt} in the left column of Figure 3 are equal to 1.092 and −0.143 (p-value = 0.082). This implies that marginal utility from \$0 to \$5 is $\Delta u_{2nt} = 0.615$; from \$5 to \$10 is $\Delta u_{3nt} = 0.179$; and from \$10 to \$15 is $\Delta u_{4nt} = 0.206$. We cannot reject the hypothesis that $\Delta u_{3nt} = \Delta u_{4nt}$ at the 5% level (p-value = 0.098). Hence, the representative EUT decision maker is risk averse in the prize set with {\$0, \$5, \$10} and risk neutral when the prize set is {\$5, \$10, \$15}. The EUT model with between-subject heterogeneity in the two RP urns reinforces this finding.[34] The estimated population means of Δu_{2nt}, Δu_{3nt} and Δu_{4nt} are equal to 0.580, 0.205 and 0.215, respectively, and we find that 86.3% of decision makers in the population exhibit diminishing marginal utility when comparisons are made in relation to {\$0, \$5, \$10}, but only 41.3% do so in relation to {\$5, \$10, \$15}. We fail to reject the hypothesis that the latter population share is different from 0.5 at the 5% significance level (p-value = 0.053).

[32] We stress that three prizes per set is a feature of the data, and *it is not a requirement by our primal estimation method*. In principle, we can also fit the EUT model with non-parametric utility to the Andersen, Harrison, Lau and Rutström [2014] data, where the lottery pairs are defined over four sets of four prizes. We do not consider this application because it would require us to report and discuss the results for 14 different RP urns.

[33] Of course, we can also restrict the EUT model to each prize set and obtain results that can be directly compared to the dual estimates. Figure C2 and C3 in Appendix C report these results. We find that the primal estimates are indistinguishable from the dual estimates.

[34] The results are reported in the last panel of Table C3 and displayed in Figure C4.

Finally, we estimate the RDU model with non-parametric utility. The model combines a logistic RP urn for the probability weighting parameter and two multinomial logit-logistic RP urns for the utility parameters. Figure 4 displays the representative agent's RP urns for the utility and probability weighting parameters. When $\varkappa = 0.010$, we find that all three RP urns play non-redundant roles in explaining stochastic choice behavior, in the sense that the noise parameter of each urn is significantly greater than zero with a p-value < 0.001. We draw similar conclusions as before with respect to diminishing marginal utility, which suggests that probability weighting has limited influence on the representative agent's risk preferences. Indeed, the estimated mean of the log-shape parameter, m_{φ}, in the bottom-left panel is -0.087 (p-value = 0.059) and we fail to reject the hypothesis that the PWF is linear at the 5% significance level. The results thus suggest that the representative agent is an EUT decision maker, albeit random fluctuations in the PWF may induce choice behavior that deviates from this characterization.

Figure 5 displays the estimated distributions from the random coefficient RDU model that accounts for between-subject heterogeneity in all three RP urns.[35] Between-subject heterogeneity in marginal utility, reported in the top and middle panels, is similar to what we find under EUT. However, between-subject heterogeneity in probability weighting, reported in the bottom panel, does not support the verdict in favor of EUT. When $\varkappa = 0.010$, the bottom-left panel shows that the within-subject mean of the log-shape parameter, $m_{\varphi n}$, is normally distributed between subjects with mean $E[m_{\varphi n}] = 0.013$ (p-value = 0.710) and standard deviation $SD[m_{\varphi n}] = 0.482$. The standard deviation is significantly greater than zero (p-value < 0.001), which suggests that a substantial share of decision makers in the population deviate from EUT. The bottom-right panel shows the between-subject distribution of the shape parameter, defined as $\exp(m_{\varphi n})$. The population mean of this distribution is equal to 1.477, which is significantly greater than 1 (p-value = 0.002). The average shape of the PWF in the population is thus S-shaped rather than linear. The results for $\varkappa = 0.020$ lead to qualitatively the same conclusions.

[35] The results for $\varkappa = 0.002$ are omitted as we could not obtain numerical solutions to the model.

5. Conclusion

The RP model provides an integral framework for modeling within-individual heterogeneity in choice behavior, by attributing such heterogeneity to preference parameters in the underlying theory of risk attitudes rather than an additive error term that is external to the theory under consideration. However, the RP likelihood function is computationally unattractive, and most studies use additive error specifications in structural estimation of risk attitudes. We propose two estimation methods to facilitate empirical applications of RP models. Our primal method illustrates that the RP model does not fall behind other stochastic choice models in terms of generalizability. By applying a kernel smoothing method, one can construct a versatile likelihood evaluator of the RP model that can accommodate any decision theoretic structure and parametric distribution of unobserved heterogeneity. Our dual method illustrates that the single crossing condition of preferences can serve as a simple but powerful tool in empirical economics, as it has been in economic theory. This property enables one to recast the RP model as a standard econometric model and apply a standard regression command, making structural estimation of unidimensional RP models immediately accessible to a broad audience.

Table 1: Dual EUT Models with CRRA Utility

Parameter	Estimate	Standard Error	*p*-value	95% Confidence Interval	
			A. Pooled Logit (*Log-likelihood* = −7501.476)		
β_0	−0.932	0.069	<0.001	−1.068	−0.795
β_1	1.740	0.088	<0.001	1.567	1.913
$\mu_\omega = -\beta_0/\beta_1$	0.535	0.033	<0.001	0.470	0.601
$\sigma_\omega = 1/\beta_1$	0.575	0.029	<0.001	0.518	0.632
			B. Random Effects Logit (*Log-likelihood* = −6143.673)		
β_0	−1.327	0.114	<0.001	−1.550	−1.104
β_1	2.411	0.133	<0.001	2.150	2.672
σ_0	1.778	0.143	<0.001	1.498	2.058
$E[\mu_{\omega n}] = -\beta_0/\beta_1$	0.550	0.038	<0.001	0.475	0.625
$SD[\mu_{\omega n}] = \sigma_0/\beta_1$	0.737	0.056	<0.001	0.628	0.847
$\sigma_\omega = 1/\beta_1$	0.415	0.023	<0.001	0.370	0.460
			C. Mixed Logit in WTP Space (*Log-likelihood* = −5374.130)		
β_0	−0.519	0.022	<0.001	−0.562	−0.475
τ	1.304	0.074	<0.001	1.158	1.450
σ_0	0.846	0.026	<0.001	0.794	0.898
σ_τ	0.997	0.070	<0.001	0.859	1.135
$\varrho_{0\tau}$	−0.055	0.095	0.564	−0.242	0.132
$E[\mu_{\omega n}] = -\beta_0$	0.519	0.022	<0.001	0.475	0.562
$E[\ln \sigma_{\omega n}] = -\tau$	−1.304	0.074	<0.001	−1.450	−1.158
$SD[\mu_{\omega n}] = \sigma_0$	0.846	0.026	<0.001	0.794	0.898
$SD[\ln \sigma_{\omega n}] = \sigma_\tau$	0.997	0.070	<0.001	0.859	1.135
$\varrho[\mu_{\omega n}, \ln \sigma_{\omega n}] = \varrho_{0\tau}$	−0.055	0.095	0.564	−0.242	0.132

Notes: All models have been estimated using the Andersen, Harrison, Lau and Rutström [2014] data set. Standard errors have been adjusted for clustering at the subject level, except in panel C. The *mixlogitwtp* (version 1.1.0) command in *Stata* does not support clustered standard errors. The coefficient $\varrho_{0\tau}$ is the correlation coefficient between the random intercept and the random log-precision parameter.

Figure 1: Dual EUT Model with Non-parametric Utility

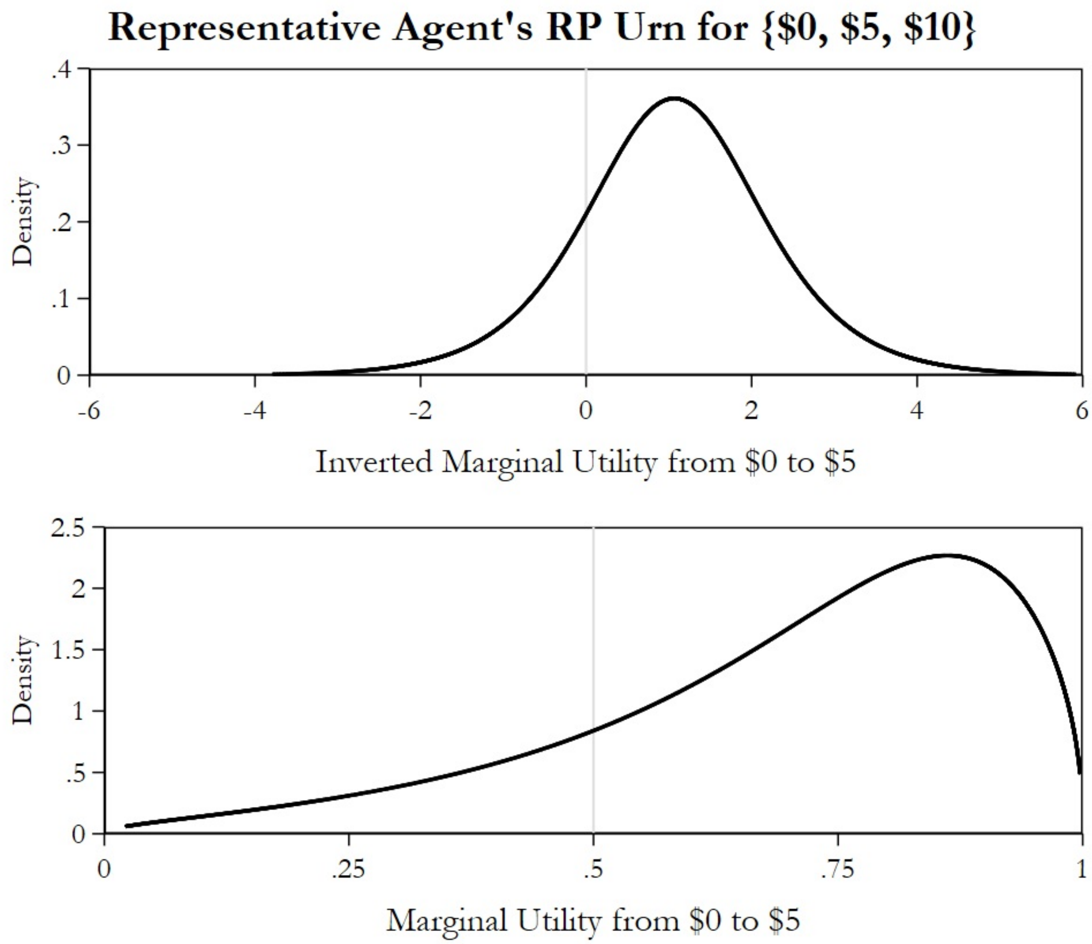


Figure 2: Kernel-Smoothed RDU Model with CRRA Utility

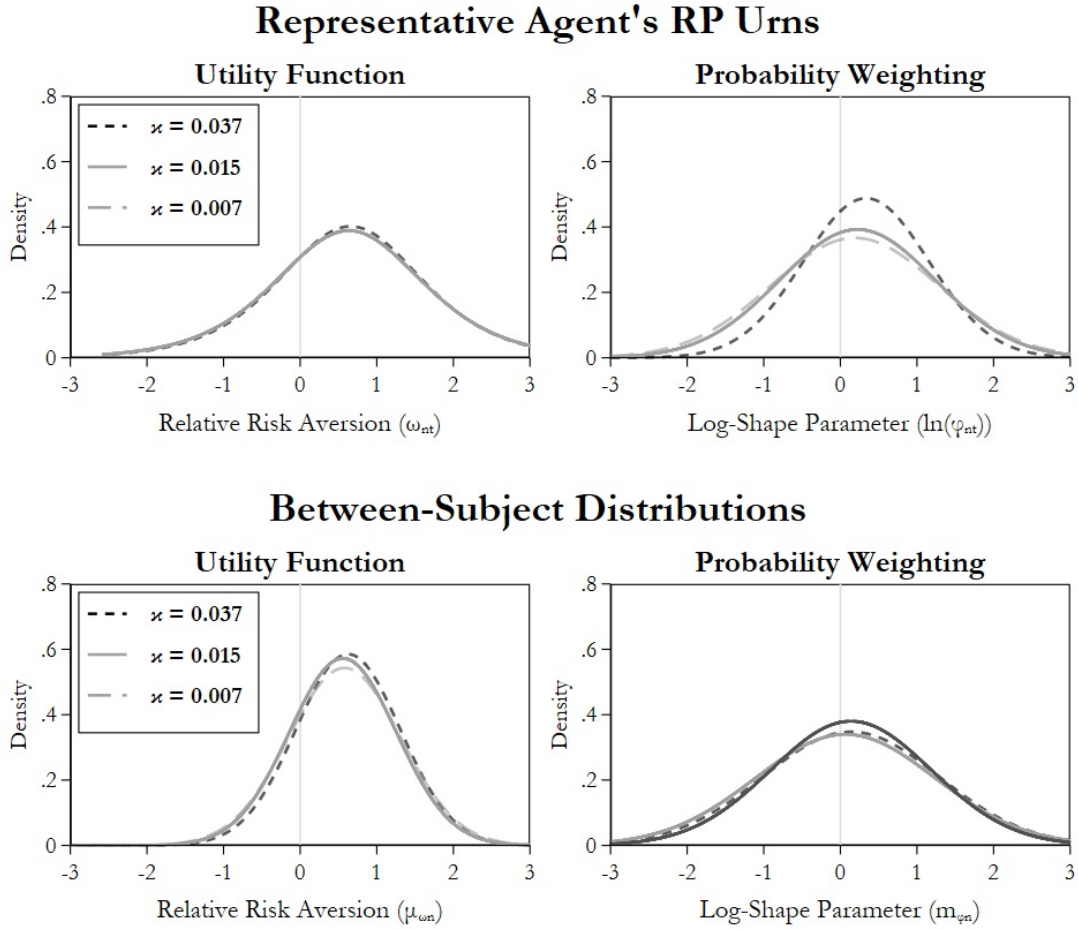


Figure 3: Kernel-Smoothed EUT Model with Non-parametric Utility

Representative Agent's RP Urns for {$0, $5, $10, $15}

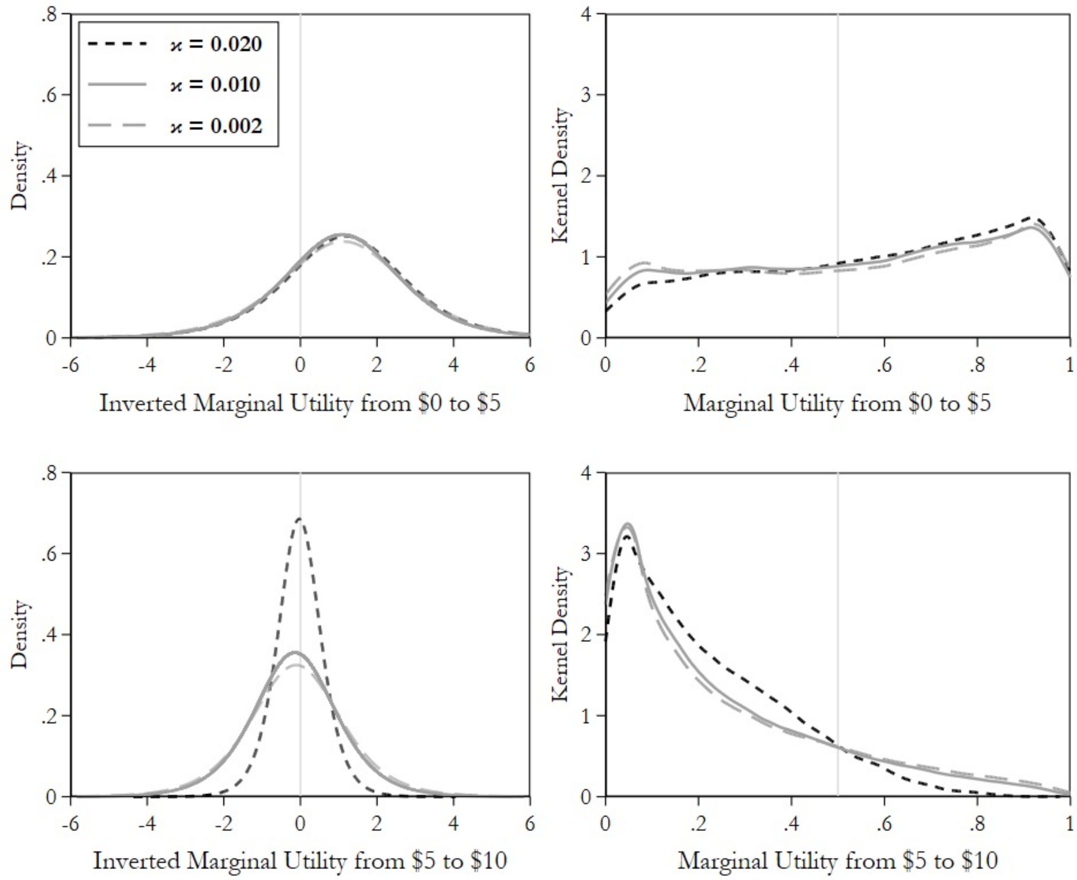


Figure 4: Kernel-Smoothed RDU Model with Non-parametric Utility

Representative Agent's RP Urns for {$0, $5, $10, $15}

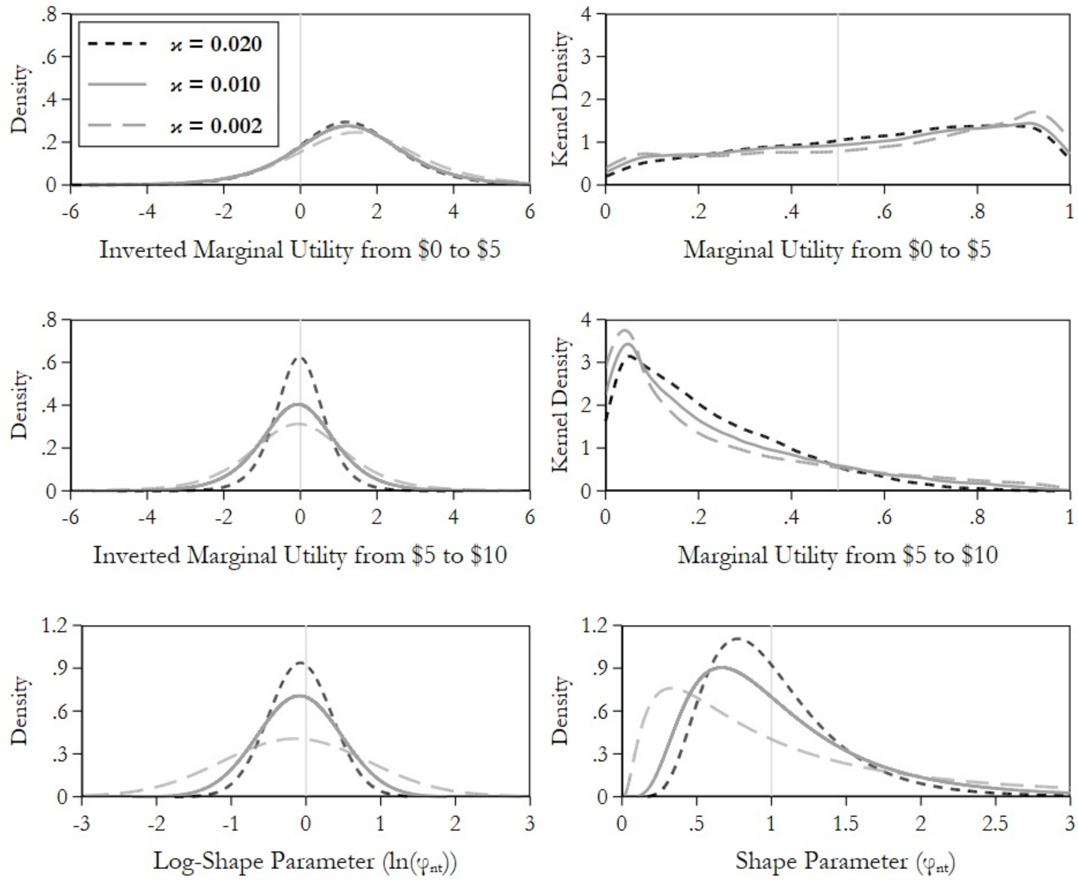
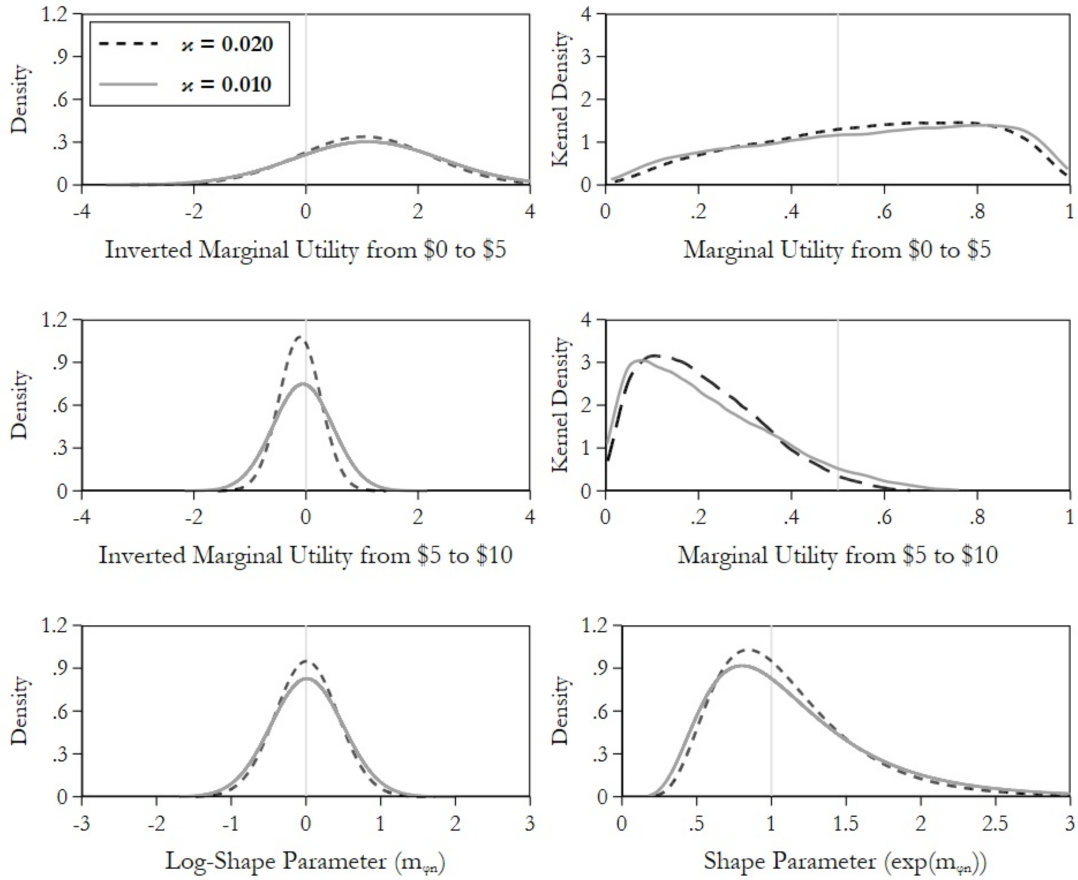


Figure 5: Kernel-Smoothed RDU Model with Non-parametric Utility

Between-Subject Distributions for {$0, $5, $10, $15}



References

Andersen, Steffen, Harrison, Glenn W.; Lau, Morten I., and Rutström, E. Elisabet, "Discounting Behavior: A Reconsideration," *European Economic Review*, 71(1), 2014, 15-33.

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

Apesteguia, Jose; Ballester, Miguel A, and Gutierrez-Daza, Angelo, "Random Models for the Joint Treatment of Risk and Time Preferences," *mimeo*, 2020.

Athey, Susan, "Single Crossing Properties and the Existent of Pure Strategy Equilibria in Games of Incomplete Information," *Econometrica*, 69, 2001, 861-889.

Barseghyan, Levon; Molinari, Francesca, and Thirkettle, Matthew, "Discrete Choice under Risk with Limited Consideration," *American Economic Review*, 111, 2021, 1972-2006.

Becker, Gordon M.; DeGroot, Morris H., and Marschak, Jacob, "Stochastic Models of Choice Behavior," *Behavioral Science*, 8, 1963, 41-55.

Bhat, Chandra R., "Quasi-random simulated likelihood estimation of the mixed multinomial logit model," *Transporation Research Part B*, 35, 2001, 677-693.

Deck, Cary, and Schlesinger, Harris, "Exploring Higher Order Risk Effects," *Review of Economic Studies*, 77, 2010, 1403-1420.

Dohmen, Thomas; Falk, Armin; Huffman, David, and Sunde, Uwe, "Are Risk Aversion and Impatience Related to Cognitive Ability," *American Economic Review*, 100, 2010, 1238-1260.

Fiebig, Denzil G.; Keane, Michael P.; Louviere, Jordan, and Wasi, Nada, "The Generalized Multinomial Logit Model: Accounting for Scale and Coefficient Heterogeneity," *Marketing Science*, 29, 2010, 393-584.

Filippin, Antonio, and Crosetto, Paolo, "A Reconsideration of Gender Differences in Risk Attitudes," *Management Science*, 62, 2016, 3138-3160.

Gans, Joshua S., and Smart, Michael, "Majority voting with single-crossing preferences," *Journal of Public Economics*, 59, 1996, 219-237.

Greene, William H., and Hensher, David A., "Does scale heterogeneity across individuals matter? An empirical assessment of alternative logit models," *Transporation*, 37, 2010, 413-428.

Guiso, Luigi; Sapienza, Paola, and Zingales, Luigi, "Time varying risk aversion," *Journal of Financial Economics*, 128, 2018, 403-421.

Gul, Faruk, and Pesendorfer, Wolfgang, "Random Expected Utility," *Econometrica*, 74, 2006, 121-146.

Harrison, Glenn W.; Lau, Morten I., and Yoo, Hong Il, "Risk Attitudes, Sample Selection and

Attrition in a Longitudinal Field Experiment," *Review of Economics and Statistics*, 102, 2020, 552-568.

Harrison, Glenn W., and Rutström, E. Elisabet, "Risk Aversion in the Laboratory," in J.C. Cox and G.W. Harrison (eds.), *Risk Aversion in Experiments* (Bingley, UK: Emerald, Research in Experimental Economics, Volume 12, 2008).

Hole, Arne Risa, "MIXLOGITWTP: Stata module to estimate mixed logit models in WTP space," *Statistical Software Components*, S458037, 2015 (Boston College Department of Economics, revised 20 Apr 2016).

Hole, Arne Risa, and Yoo, Hong Il, "The Use of Heuristic Optimization Algorithms to Facilitate the Maximum Simulated Likelihood Estimation of Random Parameter Logit Models," *Journal of the Royal Statistical Society: Series C*, 66, 2017, 997-1013.

Holt, Charles A., and Laury, Susan K., "Risk Aversion and Incentive Effects," *American Economic Review*, 92, 2002, 1644-1655.

Horowitz, Joel L., "A Smoothed Maximum Score Estimator for the Binary Response Model," *Econometrica*, 60, 1992, 505-531.

Hryshko, Dmytro; Luengo-Prado, María José Luengo, and Sørensen, Bent E., "Childhood determinants of risk aversion: The long shadow of compulsory education," *Quantitative Economics*, 2, 2011, 37-72

Jagelka, Tomáš, "Are Economists' Preferences Psychologists' Personality Traits? A Structural Approach," *IZA Discussion Paper Series*, No. 13303, 2020.

Knittel, Christopher R., and Metaxoglou, Konstantinos, "Estimation of Random-Coefficient Demand Models: Two Empirical Economists' Perspective," *Review of Economics and Statistics*, 96, 2014, 34-59.

Layton, David F., and Brown, Gardner, "Heterogeneous Preferences Regarding Global Climate Change," *Review of Economics and Statistics*, 82, 2000, 616-624.

Loomes, Graham; Moffatt, Peter G., and Sugden, Robert, "A Microeconometric Test of Alternative Stochastic Theories of Risky Choice," *Journal of Risk and Uncertainty*, 24, 2002, 103-130.

McFadden, Daniel, "Conditional logit analysis of qualitative choice behavior," in P. Zarembka (ed.), *Frontiers in Econometrics* (New York, USA: Academic Press, 1974).

McFadden, Daniel, "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica*, 57(5), 1989, 995-1026.

McFadden, Daniel, and Train, Kenneth, "Mixed MNL models for discrete response," *Journal of Applied Econometrics*, 15, 2000, 447-470.

Prelec, Drazen, "The Probability Weighting Function," *Econometrica*, 66, 1998, 497-527.

Quah, John K.-H., and Strulovici, Bruno, “Aggregating the Single Crossing Property,” *Econometrica*, 80, 2012, 2333-2348.

Quiggin, John, “A Theory of Anticipated Utility,” *Journal of Economic Behavior & Organization*, 3(4), 1982, 323-343.

Revelt, David, and Train, Kenneth, “Mixed Logit with Repeated Choices: Households’ Choices of Appliance Efficiency Level,” *Review of Economics and Statistics*, 80, 1998, 647-657.

Smithson, Michael, and Shou, Yiyun, “CDF-quantile distributions for modelling random variables on the unit interval,” *British Journal of Mathematical and Statistical Psychology*, 70, 412-438.

Train, Kenneth, *Discrete Choice Models with Simulation* (Cambridge, UK: Cambridge University Press, Second Edition, 2009).

Train, Kenneth, and Weeks, Melvyn, “Discrete Choice Models in Preference Space and Willingness-to-Pay Space,” in A. Alberini and R. Scarpa (eds.), *Applications of Simulation Methods in Environmental and Resource Economics* (Dordrecht, The Netherlands: Springer, The Economics of Non-Market Goods and Resources, Volume 6, 2005).

Wilcox, Nathaniel T, “Stochastic models for binary discrete choice under risk: A critical stochastic modeling primer and econometric comparison,” in J.C. Cox and G.W. Harrison (eds.), *Risk Aversion in Experiments* (Bingley, UK: Emerald, Research in Experimental Economics, Volume 12, 2008).

Wilcox, Nathaniel T. “‘Stochastically more risk averse:’ A contextual theory of stochastic discrete choice under risk,” *Journal of Econometrics*, 162, 2011, 89-104.

Wooldridge, Jeffrey, *Econometric Analysis of Cross Section and Panel Data* (Cambridge, USA: MIT Press, Second Edition, 2010).

Yan, Jin, and Yoo, Hong Il, “Semiparametric Estimation of the Random Utility Models with Rank-Ordered Choice Data," *Journal of Econometrics*, 211, 2019, 414-438.

Yaari, Menahem E., “The Dual Theory of Choice under Risk,” *Econometrica*, 55, 1987, 95-115.

Online Appendix A: Simulated Likelihood Functions

We first present simulated likelihood functions of the RP-EUT models with CRRA utility which are discussed in section 4 of the main text. We then generalize the example to RDU models with CRRA utility, and finally to a generic preference index function $\Delta V_{nt}(\boldsymbol{\alpha}_{nt})$ of a random preference vector $\boldsymbol{\alpha}_{nt}$. Since the appendix focuses exclusively on simulated choice probabilities and likelihood functions, we omit the descriptor "simulated" henceforth.

A1. Expected Utility Theory with CRRA Utility

We first consider the EUT model with CRRA utility where each subject has the same RP urn with logistic density $f(\omega_{nt} | \mu_\omega, \sigma_\omega)$. The likelihood of subject n's choice in task t is

$$h_{nt}(\mu_\omega, \sigma_\omega) = S_{nt}(\mu_\omega, \sigma_\omega)^{ynt} \times [1 - S_{nt}(\mu_\omega, \sigma_\omega)]^{(1 - ynt)}, \tag{A1}$$

where $S_{nt}(\mu_\omega, \sigma_\omega)$ refers to the choice probability in equation (20), and y_{nt} is a binary indicator that is equal to 1 if the observed choice is lottery B and equal to 0 if it is lottery A. The joint likelihood of all T choices by subject n is

$$H_n(\mu_\omega, \sigma_\omega) = \prod_t h_{nt}(\mu_\omega, \sigma_\omega), \tag{A2}$$

where t = 1, 2, ···, T. The MSL estimates of μ_ω and σ_ω can be computed by maximizing the sample log-likelihood function, $H(\mu_\omega, \sigma_\omega) = \sum_n \ln(H_n(\mu_\omega, \sigma_\omega))$, where n = 1, 2, ···, N. Evaluating $H(\mu_\omega, \sigma_\omega)$ requires a total of N × T × R pseudo-random number draws. As specified in (20), for each tuple of subject n and task t, one should generate R draws of ω_{nt} from $f(.|\mu_\omega, \sigma_\omega)$.[36]

Next, suppose that each subject has her own RP urn with logistic density $f(\omega_{nt} | \mu_{\omega n}, \sigma_{\omega n})$. Assume that the between-subject distribution of $\mu_{\omega n}$ and $\ln(\sigma_{\omega n})$ in the population is multivariate normal with density $\mathscr{F}([\mu_{\omega n}, \ln(\sigma_{\omega n})]' | \mathbf{b}_{EUT}, \mathbf{V}_{EUT})$, where $\mathbf{b}_{EUT}$ and $\mathbf{V}_{EUT}$ are the mean vector and covariance matrix, respectively. The joint likelihood of all T choices by subject n is then

$$J_n(\mathbf{b}_{EUT}, \mathbf{V}_{EUT}) = (1/Q)\sum_q H_n(\mu_{\omega n,q}, \sigma_{\omega n,q}), \tag{A3}$$

where $\mu_{\omega n,q}$ and $\sigma_{\omega n,q} \equiv \exp\{\ln(\sigma_{\omega n})\}$ are the q^{th} draw from the multivariate normal density $\mathscr{F}(.|\mathbf{b}_{EUT}, \mathbf{V}_{EUT})$ and q = 1, 2, ···, Q. The MSL estimates of $\mathbf{b}_{EUT}$ and $\mathbf{V}_{EUT}$ can be computed by maximizing the sample log-likelihood function, $J(\mathbf{b}_{EUT}, \mathbf{V}_{EUT}) = \sum_n \ln(J_n(\mathbf{b}_{EUT}, \mathbf{V}_{EUT}))$. Evaluating $J(\mathbf{b}_{EUT}, \mathbf{V}_{EUT})$ requires a total of N × [Q + (Q × T × R)] sets of pseudo-random number draws. For each of the N subjects, one should make Q sets of draws for $\mu_{\omega n}$ and $\sigma_{\omega n}$ from $\mathscr{F}(.|\mathbf{b}_{EUT}, \mathbf{V}_{EUT})$; conditional on each of those Q subject-specific sets of draws, for each of T decision tasks, one should make R draws of ω_{nt} from $f(.|\mu_{\omega n,q}, \sigma_{\omega n,q})$.[37]

---

[36] We find it useful to truncate the primitive variate $e_{nt,r}$ in (20) at the 0.5th and 99.5th percentiles of the standard logistic distribution. The truncation precludes rare draws of outliers that can cause numerical problems. At R = 100, we make almost 1.5 million draws of ω_{nt} since our sample includes N × T = 413 × 36 = 14,868 choice observations. The sample log-likelihood function may display numerical problems if any of those draws is extremely large in magnitude.

[37] One can free up a substantial amount of system memory by using N × [Q + (T × R)] sets of primitive draws to generate N × [Q + (Q × T × R)] sets of final draws. For each subject n and task t, one can reuse the same set of primitive standard logistic variates $e_{nt,r}$, instead of using Q different sets of $e_{nt,r}$.

A2. Rank-Dependent Utility with CRRA Utility

Consider the simple RDU model with CRRA utility where each subject has the same pair of RP urns, and let the utility urn follow a logistic density $f(\omega_{nt}|\mu_\omega, \sigma_\omega)$ and the PWF urn follow a normal density $g(\ln(\varphi_{nt})|m_\varphi, s_\varphi)$. The probability that subject n chooses lottery B in task t is

$$S_{nt}(\mu_\omega, \sigma_\omega, m_\varphi, s_\varphi) = (1/R)\sum_r \Lambda(\Delta RDU_{nt}(\omega_{nt,r}, \varphi_{nt,r})/\varkappa), \tag{A4}$$

where $\omega_{nt,r}$ and $\varphi_{nt,r}$ refer to the r^{th} draws of ω_{nt} and $\varphi_{nt} \equiv \exp\{\ln(\varphi_{nt})\}$ from $f(.|\mu_\omega, \sigma_\omega)$ and $g(.|m_\varphi, s_\varphi)$, respectively. The joint likelihood of all T choices by subject n is identical to (A2), except that $S_{nt}(\mu_\omega, \sigma_\omega)$ is replaced with $S_{nt}(\mu_\omega, \sigma_\omega, m_\varphi, s_\varphi)$. Let $H_n(\mu_\omega, \sigma_\omega, m_\varphi, s_\varphi)$ denote subject n's joint likelihood function. The MSL estimates of μ_ω, σ_ω, m_φ, and s_φ can be computed by maximizing the sample log-likelihood function, $\sum_n \ln(H_n(\mu_\omega, \sigma_\omega, m_\varphi, s_\varphi))$, which again entails $N \times T \times R$ sets of draws.

Next, suppose that each subject has her own pair of RP urns, $f(\omega_{nt}|\mu_{\omega n}, \sigma_{\omega n})$ and $g(\ln(\varphi_{nt})|m_{\varphi n}, s_{\varphi n})$. Let $\sigma^{UN}_{\omega n}$ and $s^{UN}_{\varphi n}$ denote primitive coefficients that satisfy $|\sigma^{UN}_{\omega n}| = \sigma_{\omega n}$ and $|s^{UN}_{\varphi n}| = s_{\varphi n}$. We assume that the between-subject distribution of $\boldsymbol{\theta}_n = [\mu_{\omega n}, \sigma^{UN}_{\omega n}, m_{\varphi n}, s^{UN}_{\varphi n}]$ in the population is multivariate normal with density $\mathscr{F}(\boldsymbol{\theta}_n|\mathbf{b}_{RDU}, \mathbf{V}_{RDU})$.[38] The joint likelihood of all T choices by subject n is then

$$J_n(\mathbf{b}_{RDU}, \mathbf{V}_{RDU}) = (1/Q)\sum_q H_n(\mu_{\omega n,q}, \sigma_{\omega n,q}, m_{\varphi n,q}, s_{\varphi n,q}), \tag{A5}$$

where $\mu_{\omega n,q}$, $\sigma_{\omega n,q}$, $m_{\varphi n,q}$ and $s_{\varphi n,q}$ refer to the q^{th} draw of $\boldsymbol{\theta}_n$ from $\mathscr{F}(.|\mathbf{b}_{RDU}, \mathbf{V}_{RDU})$. The MSL estimates of $\mathbf{b}_{RDU}$ and $\mathbf{V}_{RDU}$ can be computed by maximizing the sample log-likelihood function, $\sum_n \ln(J_n(\mathbf{b}_{RDU}, \mathbf{V}_{RDU}))$, which entails $N \times [Q + (Q \times T \times R)]$ sets of draws.

A3. Generalization

The EUT and RDU examples can be generalized to a generic preference index function $\Delta V_{nt}(\boldsymbol{\alpha}_{nt})$ of a random preference vector $\boldsymbol{\alpha}_{nt}$. Let $f(\boldsymbol{\alpha}_{nt}|\boldsymbol{\varsigma})$ denote the assumed joint density of $\boldsymbol{\alpha}_{nt}$ as a function of parameters $\boldsymbol{\varsigma}$. The probability that subject n chooses lottery B in task t is then $S_{nt}(\boldsymbol{\varsigma}) = (1/R)\sum_r \Lambda(\Delta(\boldsymbol{\alpha}_{nt,r})/\varkappa)$, where $\boldsymbol{\alpha}_{nt,r}$ is the r^{th} draw of $\boldsymbol{\alpha}_{nt}$ from $f(.|\boldsymbol{\varsigma})$. Subject n's joint likelihood function, $H_n(\boldsymbol{\varsigma})$, can be constructed by replacing $S_{nt}(\mu_\omega, \sigma_\omega)$ with $S_{nt}(\boldsymbol{\varsigma})$ in (A2), and used subsequently in estimation. For instance, consider a random coefficient model that replaces $\boldsymbol{\varsigma}$ with a vector of random coefficients $\boldsymbol{\varsigma}_n$, which has a joint density function $\mathscr{F}(\boldsymbol{\varsigma}_n|\boldsymbol{\Psi})$ of hyper-parameters $\boldsymbol{\Psi}$ that characterize its between-subject distribution. We use $\mathscr{F}(.|\boldsymbol{\Psi})$ to denote a generic joint density function rather than a multivariate normal density function. Subject n's joint likelihood in this case can be evaluated as

$$J_n(\boldsymbol{\Psi}) = (1/Q)\sum_q H_n(\boldsymbol{\varsigma}_{n,q}), \tag{A6}$$

where $\boldsymbol{\varsigma}_{n,q}$ is the q^{th} draw of $\boldsymbol{\varsigma}_n$ from $\mathscr{F}(.|\boldsymbol{\Psi})$.

[38] The between-subject distributions of $\sigma_{\omega n}$ is now folded-normal instead of log-normal. We could not obtain a solution to the RDU specification with log-normal $\sigma_{\omega n}$.

Online Appendix B: Supporting Results for Section 3

We report additional results to support our discussion in section 3 of the main paper. Tables B1 and B2 report the dual estimates of the EUT models with non-parametric utility. Figure 1 in the main text displays the pooled logit estimates reported in Table B1. Tables B3, B4 and B5 report the dual estimation results for other types of model specifications, which we will explain in the remainder of this appendix.

B1. Variations on Model Specification

We can allow the risk aversion parameter, μ_ω or $E[\mu_{\omega n}]$, to vary with observed characteristics by allowing the intercept β_0 in the dual logit model to vary with the same characteristics. In practice, this simply entails adding independent variables to the logit model as usual. For example, adding the subject's age as an independent variable to the logit model is equivalent to replacing β_0 with a linear function ($b_0 + b_1 \times \mathit{age}_n$), where b_0 and b_1 are unknown parameters to be estimated. These parameters can be transformed in the same manner as β_0 itself to derive the overall intercept and slope coefficient for the risk aversion parameter: $-b_0/\beta_1$ and $-b_1/\beta_1$ in the pooled and RE logit models, and $-b_0$ and $-b_1$ in the mixed logit model in the WTP space.

We can also allow the log-noise parameter, $\ln(\sigma_\omega)$ or $E[\ln(\sigma_{\omega n})]$, to vary with observed characteristics. In the WTP space model, this is done by replacing the intercept τ in its log-precision parameter, $\ln(\lambda_n) \sim N(\tau, \sigma_\tau^2)$, with ($c_0 + c_1 \times \mathit{age}_n$) where c_0 and c_1 are parameters to be estimated. Just as multiplying τ by -1 results in the log-noise parameter, multiplying c_0 and c_1 by -1 results in the intercept and slope coefficient for the log-noise parameter. The pooled and RE logit models with observed heterogeneity in the log-noise parameter can be estimated as a special case of this WTP space model, by specifying both the overall intercept and the log-precision parameter as fixed coefficients (pooled) or only the log-precision parameter as a fixed coefficient (RE).

Our examples for the RE logit and WTP space models use normal distributions to accommodate between-subject heterogeneity in the dual RP urns. This distributional assumption can be changed by using standard models for panel data which are typically estimated as alternatives to those models. For example, most software packages support the estimation of mixed effects models that allow the random intercept in the RE logit model to follow a non-normal parametric distribution, and latent class logit models that use categorical distributions as non-parametric approximation to an unknown between-subject distribution of the mixed logit coefficients.

If one is willing to impose *a priori* bounds for the random risk aversion parameter ω_{nt}, one can exploit that restriction to apply a very flexible parametric distribution to model the RP urn for ω_{nt}, retaining at the same time the logit link which gives access to the mixed logit commands.[39] Suppose that ω_{nt} lies in a bounded interval (w_{MIN}, w_{MAX}), where w_{MIN} is any number smaller than the sample minimum of w_{nt} and w_{MAX} is any number greater than the sample maximum of w_{nt}. Assume

[39] This is not to say that the use of the logit link is essential to the dual estimation. All major software packages support the estimation of pooled and RE binomial choice models with alternative link functions, such as probit, cauchit and complementary log-log; they can be applied to the dual estimation in the same manner as we have described, with suitable re-labeling of the location/mean and scale parameters. Retaining the logit link, however, gives a distinctive advantage when it comes to modeling between-subject heterogeneity in both risk and noise parameters: Non-logit alternatives to the mixed logit in the WTP space or GMNL-II are not available as standard estimation commands, as they are rarely used in empirical research.

further that the RP urn for ω_{nt} is a logit-logistic random variable: This stochastic assumption may be represented as either $\omega_{nt} = w_{MIN} + (w_{MAX} - w_{MIN}) \times \Lambda(\zeta_{nt})$ or $\zeta_{nt} = \Lambda^{-1}((\omega_{nt} - w_{MIN}) / (w_{MAX} - w_{MIN}))$, where $\Lambda^{-1}(.)$ is the inverse of the standard logistic distribution function $\Lambda(.)$ and ζ_{nt} is a logistic random variable with a mean of μ_ζ and a standard deviation of $\sigma_\zeta \times 1.81$. Denote by $z_{nt} = \Lambda^{-1}((w_{nt} - w_{MIN}) / (w_{MAX} - w_{MIN}))$ the corresponding inversion that converts the indifference point w_{nt} into the same units as ζ_{nt}.

The within-subject distribution of the logit-logistic ω_{nt} resembles a beta distribution. It can display a variety of shapes, such as bimodality and non-zero skewness, depending on the mean μ_ζ and scale σ_ζ of the underlying logistic variable. Smithson and Shou [2017; p.423] graphically illustrate this flexibility. The probability that subject n chooses lottery B in task t is given by $\Pr(\omega_{nt} < w_{nt}) = \Pr(\zeta_{nt} < z_{nt})$, and has an analytic solution identical to equation (14), once we let μ_ζ, σ_ζ and z_{nt} take the place of μ_ω, σ_ω and w_{nt}, respectively. Thus, in the representative agent case, we can use the pooled logit model to recover the mean and scale of the logistic component that underlies the logit-logistic ω_{nt}. To account for between-subject heterogeneity in the mean (and scale), we can continue using the RE logit model (or the WTP space model) as described in the preceding sections.

Our dual estimation method can be directly applied to EUT with constant absolute risk aversion (CARA) utility. Algebraically, we simply need to replace the CRRA utility function with the CARA utility function, prior to finding the root of the equation $\Delta EU_{nt}(w_{nt}) = 0$. Of course, in this case, both random risk aversion parameter ω_{nt} and indifference point w_{nt} must be interpreted in relation to the coefficient of ARA rather than the coefficient of RRA.

When all decision tasks in the sample involve probability distributions over the same set of three prizes (*e.g.*, Loomes, Moffatt and Sugden [2002]), our dual method can be adapted to estimate a doubly flexible RP model that combines a non-parametric approach to the utility function in equation (8) with a logit-logistic distribution of the RP component.[40] As with (8), let p_{Lk} denote the probability of prize m_k in lottery $L \in \{A, B\}$ and u_k denote the utility of that prize, where $u_1 < u_2 < u_3$; we have omitted the nt subscript from each of these variables and parameters to reduce notational cluttering. Suppose that B is risky relative to A in the sense that the former offers higher chances of the best and worst outcomes, *i.e.*, $p_{A1} < p_{B1}$ and $p_{A3} < p_{B3}$. Since the utility function under EUT is unique up to a positive affine transformation, we can normalize $u_1 = 0$ and $u_3 = 1$ without the loss of generality, retaining u_2 as a free parameter that lies in the unit interval. Then, the expected utility difference can be simplified as $\Delta EU(u_2) = (p_{B3} - p_{A3}) + (p_{B2} - p_{A2})u_2$, and the indifference point $\bar{u}_2$ that solves $\Delta EU(\bar{u}_2) = 0$ has a closed-form solution of $\bar{u}_2 = (p_{B3} - p_{A3}) / [(p_{B3} - p_{A3}) + (p_{B1} - p_{A1})]$, which also lies in the unit interval. The utility parameter u_2 is a measure of risk aversion that satisfies the single crossing property: Anyone with $u_2 < \bar{u}_2$ chooses the risky lottery B, *i.e.*, $\Delta EU(u_2) > 0$, and anyone with $u_2 > \bar{u}_2$ chooses the safe lottery A, *i.e.*, $\Delta EU(u_2) < 0$.

[40] Our dual estimation method cannot be applied to EUT with non-parametric utility when some decision tasks involve probability distributions over one set of three prizes and other decision tasks involve probability distributions over another set(s) of three prizes (*e.g.*, Harrison and Rutström [2008; §2.6]). In such cases, the model has a multidimensional vector of preference parameters because even after normalizing the sample maximum and minimum of the utility levels, one retains two or more utility levels as free parameters. Therefore, to proceed with the estimation of EUT with non-parametric utility, one must apply a more general approach that we present in section 4; alternatively, one may split the sample into different sub-samples based on prize sets, and analyze each sub-sample separately using the dual estimation method.

To introduce the RP component, we specify u_2 as a random risk aversion parameter u_{2nt} that varies randomly across subjects and over choice tasks. Assume further that the RP urn for u_{2nt} is a logit-logistic random variable: $u_{2nt} = \Lambda(\zeta_{nt})$ or $\zeta_{nt} = \Lambda^{-1}(u_{2nt})$, where $\Lambda(.)$, $\Lambda^{-1}(.)$ and ζ_{nt} are as defined earlier, and the bounds for $u_{2nt} \in (0, 1)$ follow from the income monotonicity of utilities. The probability that subject n chooses lottery B in task t is given by $\Pr(u_{2nt} > \bar{u}_{2nt}) = \Pr(\zeta_{nt} > \Lambda^{-1}(\bar{u}_{2nt}))$, and has an analytic solution identical to equation (14) once we set $\mu_z = \mu_\omega$, $\sigma_z = \sigma_\omega$ and $\Lambda^{-1}(\bar{u}_{2nt}) = w_{nt}$. The resulting EUT model is thus dual to the pooled logit model under the representative agent assumption, and can be generalized using the RE and mixed logit models to account for between-subject heterogeneity in μ_z and σ_z.

EUT does not exhaust models with potentially unidimensional risk preference parameters.[41] A prominent alternative is Yaari's [1987] Dual Theory of Choice under Risk (DT), which attributes risk attitude to non-linear probability weighting rather than non-linear utility curvature: It can be seen as a special case of RDU in equation (9) that constrains the utility function to be linear, by setting $U(m_{Lk} | \omega) = m_{Lk}$. When DT displays the single crossing property, one can apply our dual estimation method to estimate its RP variants, just as with the EUT model. Given the Prelec PWF in (10) or the power PWF in (B2) below, lottery pairs based on the logic of MPL induce single crossing in the DT index function as required by our method.[42]

B2. Empirical Illustration: Dual Theory Models

We consider the dual estimation of a RP variant of Yaari's [1987] Dual Theory (DT), using data from a field experiment reported in Andersen, Harrison, Lau and Rutström [2014]. As summarized in section 2 of our main text, the battery of choice tasks in the experiment was based on the same logic as the multiple price list of Holt and Laury [2002]: Each lottery pair can be written as $A = \{(m_{A1}, (1 - p_2)), (m_{A2}, p_2)\}$ and $B = \{(m_{B1}, (1 - p_2)), (m_{B2}, p_2)\}$ where $m_{B1} < m_{A1} < m_{A2} < m_{B2}$ and the probability of the higher prize p_2 is identical across the two lotteries.

Under DT, the subject's evaluation of lottery $L \in \{A, B\}$ in such pairs may be specified as

$$DT_L(\eta) = [(1 - \pi(p_2 | \eta)] \times m_{L1} + \pi(p_2 | \eta) \times m_{L2} \tag{B1}$$

where $\pi(. | \eta)$ is some probability weighting function (PWF) of a unidimensional parameter η. Before placing a functional form on $\pi(. | \eta)$, define $\bar{p}_2 = (m_{A1} - m_{B1}) / [(m_{B2} - m_{B1}) - (m_{A2} - m_{A1})]$, and let $\Delta DT(\eta) = DT_B(\eta) - DT_A(\eta)$ denote the subject's valuation difference between the two lotteries. Next, note that $\Delta DT(\eta) > 0$ if $\pi(p_2 | \eta) > \bar{p}_2$ and $\Delta DT(\eta) < 0$ if $\pi(p_2 | \eta) < \bar{p}_2$: That is, $\bar{p}_2$ measures the threshold level of the weighted probability of the higher prize that determines whether the subject prefers lottery B or lottery A. It follows that as long as $\pi(p_2 | \eta)$ is monotone increasing or decreasing

[41] More generally, one may also apply our approach outside the analysis of risk preferences. For example, consider the estimation of the RP model of exponential discounting behavior that satisfies the single crossing property (*e.g.*, Apesteguia and Ballester [2018]). Since an exponential discounting factor lies in the unit interval, it lends itself to the use of a logit-logistic RP urn.

[42] In the MPL design, two lotteries have the same expected value if the probability of the better prize in each lottery is equal to $\bar{p}_2 = (m_{A1} - m_{B1}) / [(m_{B2} - m_{B1}) - (m_{A2} - m_{A1})]$ because the prizes satisfy the inequality $m_{B1} < m_{A1} < m_{A2} < m_{B2}$. With an abuse of notation, let $\pi(P | \varphi)$ denote an arbitrary PWF of a unidimensional parameter φ, which needs not be the Prelec function in (10). The indifference point under DT is the value of φ that solves the equation $\pi(p_2 | \varphi) = \bar{p}_2$, where p_2 is the objective probability of the better prize.

in η at a given p_2, $\Delta DT(\eta)$ displays the single crossing property as required by the dual estimation method, and one can calculate the indifference point η^* that solves $\Delta DT(\eta^*) = 0$ by inverting the identity $\pi(p_2 | \eta^*) = \bar{p}_2$ with respect to η^*.

Prelec's [1998] PWF in equation (10) is a functional form that induces single crossing in $\Delta DT(\eta)$, given the lottery pairs in question. In the present application, for the ease of interpretation, we adopt another functional form that is consistent with the single crossing property. Specifically, we parameterize $\pi(. | \eta)$ as a power function

$$\pi(p_2 | \eta) = (p_2)^\eta \tag{B2}$$

where the power coefficient η is a risk aversion parameter, in the sense that the decision makers with $\eta \in (0, 1)$ display optimism (*i.e.*, $\pi(p_2 | \eta) > p_2$) which induces risk seeking and those with $\eta > 1$ display pessimism (*i.e.*, $\pi(p_2 | \eta) < p_2$) which induces risk aversion. The indifference point η^* then has a simple analytic solution, namely $\eta^* = \ln(\bar{p}_2) / p_2$.

As with the main text, let n index subjects and t index choice tasks. Let $\ln(\eta_{nt})$ denote a primitive RP parameter that randomly varies across subjects and choice tasks, and $\ln(\eta^*_{nt})$ denote an independent variable that records the natural log of the indifference point η^*_{nt}. Suppose that the RP urn for $\ln(\eta_{nt})$ is a logistic density function $f(\ln(\eta_{nt}) | \mu_\eta, \sigma_\eta)$ of the risk aversion parameter μ_η and the noise parameter σ_η. Since $\ln(1) = 0$, a positive (negative) value of μ_η indicates that the subject is risk averse (risk seeking) in the absence of behavioral noise.

The analytic steps that we have described for the EUT model with CRRA utility can be directly adapted to accommodate the DT model with the power PWF, once we set ω_{nt} to $\ln(\eta_{nt})$ and w_{nt} to $\ln(\eta^*_{nt})$. This includes the use of the RE logit and mixed logit in the WTP space to capture between-subject heterogeneity in the RP urn. Table B3 reports the three types of logit models which are dual to the DT model. In brief, we draw qualitatively similar conclusions as we did with the EUT model. We find evidence of risk aversion for the representative agent (pooled logit), as well as the average decision maker in the subject population (RE logit and WTP space model). Between-subject heterogeneity in the risk aversion parameter, estimated for the RE logit (WTP space) model, implies that 81.1% (81.6%) of decision makers are risk averse; the corresponding estimate from the EUT model is 77.2% (73.0%).

B3. Empirical Illustration: Incorporating Observed Heterogeneity

We can extend the dual estimation method to incorporate observed heterogeneity in risk preferences, simply by adding relevant independent variables to the dual logit models in the usual manner. To facilitate discussion, suppose that the RP model of interest is the EUT model with CRRA utility reported in section 3, and let $female_n$ denote a binary indicator variable that is equal to one if subject n is female and zero otherwise. To start off with, suppose now that we add this variable to the dual pooled logit model as follows

$$L_{nt}(b_0, b_1, \beta_1) = \Lambda(b_0 + b_1 female_n + \beta_1 w_{nt}) \tag{B3}$$

which can be seen as a specification that replaces a constant intercept β_0 in (12) with a demographic intercept $(b_0 + b_1 female_n)$. Just as the constant specification (12) is dual to the logistic EUT model with the risk aversion parameter $\mu_\omega = -\beta_0/\beta_1$ and the noise parameter $\sigma_\omega = 1/\beta_1$, the demographic specification (B3) is dual to the logistic EUT model with the risk aversion parameter $(m_0 +$

$m_1$female$_n$) = ($-b_0/\beta_1 - b_1/\beta_1$) and the noise parameter $\sigma_\omega = 1/\beta_1$. This insight extends directly to the RE logit (15) and the mixed logit model in the WTP space in (17); in the latter model, we can also replace a constant intercept τ in the log-precision parameter, $\ln(\lambda_n) \sim N(\tau, \sigma_\tau^2)$, with a demographic intercept, thereby inducing observed heteroskedasticity in the log-noise parameter of the primal model.

Table B4 reports the demographic extension of each dual logit model in Table 1. To make the link between the two tables clearer, we use β_0:base and β_0:female to denote b_0 and b_1, and likewise we use μ_ω:base and μ_ω:female to denote m_0 and m_1. Other parameter labels with base and female suffixes can be similarly associated with the baseline intercept and the demographic slope coefficient. As with section 3, we use the *logit and xtlogit, re* commands in *Stata* to estimate the pooled and RE logit models. Hole's [2016] *mixlogitwtp* command does not allow one to include observed heterogeneity in the log-precision parameter of the WTP space model. We therefore use the *gmnl* command by Gu, Knox and Hole [2013] and estimate the WTP space model as a special case of the GMNL-II model (Fiebig, Keane, Louviere, and Wasi [2010]), an approach inspired by Greene and Hensher [2010; p.416].[43]

The pooled logit model reported in the top panel of Table B4 suggests that while both men and women are risk averse, women tend to be more risk averse than men. The coefficient of relative risk aversion is equal to 0.416 for men (μ_ω:base) and 0.664 (μ_ω:base + μ_ω:female) for women; the female-male difference of 0.248 (μ_ω:female) is significantly greater than zero (p-value < 0.001). We draw qualitatively similar conclusions from the RE logit model in the middle panel and the WTP space model in the bottom panel, which account for unobserved between-subject heterogeneity in risk aversion on top of the observed heterogeneity. The WTP space model also includes an extra parameter E[ln $\sigma_{\omega n}$]:female, which captures whether women's risk preferences tend to show greater or smaller random fluctuations than men's. We do not find evidence of such demographic heteroskedasticity: The point estimate of 0.176 is small in magnitude compared to the overall intercept (E[ln $\sigma_{\omega n}$]:base = −1.139) and is not significantly different from zero at the 5% significance level (p-value = 0.076).

B4. Empirical Illustration: Dual Estimation with Latent Class Models

The latent class logit model is perhaps the most widely used alternative to mixed logit models that assume normally distributed random coefficients: See, for example, Oviedo and Yoo [2017] and Doiron and Yoo [2020]. From an econometric perspective, the latent class logit can be seen as a finite mixture of *C* different pooled logit models, where *C* is a number to be preset prior to estimation and each of the component models has its own values of parameters β_0 and β_1. The log-likelihood function of the latent class logit model is equal to a weighted average of the log-likelihood functions of the pooled logit models, where the weights are parameters to be estimated alongside the *C* different vectors of β_0 and β_1.

We can apply the latent class logit model to the dual estimation of the RP-EUT model with CRRA utility (for that matter, to the dual estimation of any RP model which displays the single crossing property). The primal model can be motivated as a model which assumes that there are *C* different classes or types of decision makers in the subject population, where each type of decision maker has their own logistic RP urn, with their own values of risk aversion parameter μ_ω and noise

[43] We thank Arne Risa Hole for alerting us to this reference.

parameter σ_ω. The weight on each class in the log-likelihood function can be seen as the population share of that class. Much as the mixed logit model in the WTP space, the latent class logit model is a model that accounts for between-subject heterogeneity in both risk aversion and noise since the μ_ω and σ_ω are allowed to vary across classes.

The weighted average structure of the latent class logit model's log-likelihood function opens up another way of addressing observed heterogeneity in risk preferences. Instead of specifying μ_ω and σ_ω to vary with observed characteristics as we have illustrated in section B1, we may let the mixture weights or class shares to vary with observed characteristics. When the observed characteristics are demographic characteristics like female_n, the primal RP model postulates that there are fundamentally *C* different types of RP urns and some types are more prevalent in certain demographic groups than others.

In Table B5, we use the *lclogit2* command in *Stata* (Yoo [2020]) to fit a 3-class latent class logit model to the Andersen, Harrison, Lau and Rutström [2014] data set. This model is dual to a 3-class RP-EUT model with CRRA utility, and we specify the share of each class to vary between male and female sub-populations by including female_n as an independent variable in the share equation. The risk aversion parameter μ_ω (noise parameter σ_ω) is estimated to be 0.357 (0.233) for class 1; 0.541 (0.652) for class 2 and −0.118 (2.041) for class 3. The baseline class share estimates suggest that class 1 makes up 61.8% of the male population; class 2 makes up 21.5%; and class 3 makes up 16.7%. Finally, we find that women are more risk averse than men in the sense that the most risk averse class makes up a greater share of the female sub-population: the share of class 2 in the female sub-population is 14 percentage points greater, and that of class 1 is 10.1 percentage points smaller. While the share of the only risk seeking type, class 3, is also smaller by 3.9 percentage points in the female sub-population, this difference is not statistically significant (p-value = 0.267).

Additional References

Gu, Yuanyuan; Hole, Arne Risa; and Knox, Stephanie, "Fitting the generalized multionomial logit model in Stata," *Stata Journal*, 13(2), 2013, 382-397.

Doiron, Denise, and Yoo, Hong Il, "Stated Preferences over Job Characteristics: a Panel Study," *Canadian Journal of Economics*, 53, 2020, 43-82.

Oviedo, Jose L, and Yoo, Hong Il, "A Latent Class Nested Logit Model for Rank-Ordered Data with Application to Cork Oak Reforestation," *Environmental and Resource Economics*, 68, 2017, 1021-1051.

Yoo, Hong Il, "lclogit2: An enhanced command to fit latent class conditional logit models," *Stata Journal*, 20(2), 2020, 405-425.

Table B1: EUT Model with Non-parametric Utility for {$0, $5, $10}

Parameter	Estimate	Standard Error	*p*-value	95% Confidence Interval	
			A. Pooled Logit (Log-likelihood = −746.988)		
β_0	−1.538	0.197	<0.001	−1.925	−1.152
β_1	1.443	0.162	<0.001	1.126	1.761
$\mu_\zeta = -\beta_0/\beta_1$	1.066	0.107	<0.001	0.857	1.275
$\sigma_\zeta = 1/\beta_1$	0.693	0.078	<0.001	0.541	0.845
			B. Random Effects Logit (Log-likelihood = −638.381)		
β_0	−2.066	0.257	<0.001	−2.570	−1.562
β_1	1.945	0.217	<0.001	1.521	2.370
σ_0	1.464	0.192	<0.001	1.086	1.841
$E[\mu_{\zeta n}] = -\beta_0/\beta_1$	1.062	0.109	<0.001	0.848	1.276
$SD[\mu_{\zeta n}] = \sigma_0/\beta_1$	0.752	0.112	<0.001	0.533	0.972
$\sigma_\zeta = 1/\beta_1$	0.514	0.057	<0.001	0.402	0.626
			C. Mixed Logit in WTP Space (Log-likelihood = −633.076)		
β_0	−1.012	0.083	<0.001	−1.175	−0.850
τ	0.717	0.106	<0.001	0.509	0.925
σ_0	0.799	0.099	<0.001	0.604	0.994
σ_τ	0.534	0.115	<0.001	0.759	0.310
$E[\mu_{\zeta n}] = -\beta_0$	1.012	0.083	<0.001	1.175	0.850
$E[\ln \sigma_{\zeta n}] = -\tau$	−0.717	0.106	<0.001	−0.509	−0.925
$SD[\mu_{\zeta n}] = \sigma_0$	0.799	0.099	<0.001	0.604	0.994
$SD[\ln \sigma_{\zeta n}] = \sigma_\tau$	0.534	0.115	<0.001	0.759	0.310

Notes: All models have been estimated using the Harrison and Rutström [2008] data set. Standard errors have been adjusted for clustering at the subject level, except in panel C. The *mixlogitwtp* (version 1.1.0) command in *Stata* does not support clustered standard errors.

Table B2: EUT Model with Non-parametric Utility for {$5, $10, $15}

Parameter	Estimate	Standard Error	*p*-value	95% Confidence Interval	
			A. Pooled Logit (Log-likelihood = −704.567)		
β_0	0.160	0.143	0.263	−0.120	0.439
β_1	1.060	0.140	<0.001	0.785	1.335
$\mu_\zeta = -\beta_0/\beta_1$	−0.151	0.143	0.293	−0.431	0.130
$\sigma_\zeta = 1/\beta_1$	0.943	0.125	<0.001	0.699	1.188
			B. Random Effects Logit (Log-likelihood = −639.532)		
β_0	0.204	0.178	0.251	−0.145	0.553
β_1	1.310	0.174	<0.001	0.970	1.650
σ_0	1.134	0.148	<0.001	0.844	1.425
$E[\mu_{\zeta n}] = -\beta_0/\beta_1$	−0.156	0.146	0.284	−0.441	0.129
$SD[\mu_{\zeta n}] = \sigma_0/\beta_1$	0.866	0.145	<0.001	0.582	1.150
$\sigma_\zeta = 1/\beta_1$	0.764	0.101	<0.001	0.565	0.962
			C. Mixed Logit in WTP Space (Log-likelihood = −639.486)		
β_0	0.163	0.143	0.254	−0.117	0.444
τ	0.268	0.138	0.053	−0.003	0.539
σ_0	0.867	0.149	<0.001	0.575	1.159
σ_τ	0.109	0.344	0.751	−0.564	0.783
$E[\mu_{\zeta n}] = -\beta_0$	−0.163	0.143	0.254	−0.444	0.117
$E[\ln \sigma_{\zeta n}] = -\tau$	−0.268	0.138	0.053	−0.539	0.003
$SD[\mu_{\zeta n}] = \sigma_0$	0.867	0.149	<0.001	0.575	1.159
$SD[\ln \sigma_{\zeta n}] = \sigma_\tau$	0.109	0.344	0.751	−0.564	0.783

Notes: All models have been estimated using the Harrison and Rutström [2008] data set. Standard errors have been adjusted for clustering at the subject level, except in panel C. The *mixlogitwtp* (version 1.1.0) command in *Stata* does not support clustered standard errors.

Table B3: RDU Model with Linear Utility and Power PWF

Parameter	Estimate	Standard Error	p-value	95% Confidence Interval	
			A. Pooled Logit (Log-likelihood = −7508.914)		
β_0	−1.096	0.063	<0.001	−1.220	−0.972
β_1	1.559	0.070	<0.001	1.422	1.696
$\mu_\eta = -\beta_0/\beta_1$	0.703	0.040	<0.001	0.625	0.781
$\sigma_\eta = 1/\beta_1$	0.641	0.029	<0.001	0.585	0.698
			B. Random Effects Logit (Log-likelihood = −6125.813)		
β_0	−1.587	0.112	<0.001	−1.807	−1.368
β_1	2.187	0.108	<0.001	1.975	2.400
σ_0	1.800	0.138	<0.001	1.530	2.070
$E[\mu_{\eta n}] = -\beta_0/\beta_1$	0.726	0.044	<0.001	0.639	0.813
$SD[\mu_{\eta n}] = \sigma_0/\beta_1$	0.823	0.057	<0.001	0.710	0.935
$\sigma_\eta = 1/\beta_1$	0.457	0.023	<0.001	0.413	0.502
			C. Mixed Logit in WTP Space (Log-likelihood = −5439.673)		
β_0	−0.782	0.030	<0.001	−0.841	−0.723
τ	1.094	0.051	<0.001	0.995	1.193
σ_0	0.869	0.030	<0.001	0.810	0.927
σ_τ	1.010	0.046	<0.001	0.920	1.101
$\varrho_{0\tau}$	0.180	0.042	<0.001	0.098	0.262
$E[\mu_{\eta n}] = -\beta_0$	0.782	0.030	<0.001	0.723	0.841
$E[\ln \sigma_{\eta n}] = -\tau$	−1.094	0.051	<0.001	−0.539	0.003
$SD[\mu_{\eta n}] = \sigma_0$	0.869	0.030	<0.001	0.575	1.159
$SD[\ln \sigma_{\eta n}] = \sigma_\tau$	1.010	0.046	<0.001	−0.564	0.783
$\varrho[\mu_{\eta n}, \ln \sigma_{\eta n}] = \varrho_{0\tau}$	0.180	0.042	<0.001	−0.242	0.132

Notes: All models have been estimated using the Andersen, Harrison, Lau and Rutström [2014] data set. Standard errors have been adjusted for clustering at the subject level, except in panel C. The *mixlogitwtp* (version 1.1.0) command in *Stata* does not support clustered standard errors. The coefficient $\varrho_{0\tau}$ is the correlation coefficient between the random intercept and the random log-precision parameter.

Table B4: EUT Model with CRRA Utility and Observed Heterogeneity

Parameter	Estimate	Standard Error	p-value	95% Confidence Interval	
A. Pooled Logit with Observed Heterogeneity in Risk Aversion (Log-likelihood = −7442.767)					
β_0:base	−0.730	0.089	<0.001	−0.904	−0.556
β_0:female	−0.436	0.113	<0.001	−0.658	−0.213
β_1	1.756	0.089	<0.001	1.582	1.930
μ_ω:base	0.416	0.046	<0.001	0.326	0.505
μ_ω:female	0.248	0.065	<0.001	0.120	0.376
$\sigma_\omega = 1/\beta_1$	0.569	0.029	<0.001	0.513	0.626
B. Random Effects Logit with Observed Heterogeneity in Risk Aversion (Log-likelihood = −6137.296)					
β_0:base	−1.015	0.140	<0.001	−1.290	−0.740
β_0:female	−0.645	0.182	<0.001	−1.001	−0.289
β_1	2.411	0.133	<0.001	2.150	2.672
σ_0	1.746	0.140	<0.001	1.471	2.021
$E[\mu_{\omega n}]$:base	0.421	0.053	<0.001	0.317	0.525
$E[\mu_{\omega n}]$:female	0.268	0.075	<0.001	0.120	0.415
$SD[\mu_{\omega n}] = \sigma_0/\beta_1$	0.724	0.055	<0.001	0.616	0.832
$\sigma_\omega = 1/\beta_1$	0.415	0.023	<0.001	0.370	0.460
C. Mixed Logit in WTP Space with Observed Heterogeneity in Risk Aversion & Noise (Log-likelihood = −5395.158)					
β_0:base	−0.366	0.038	<0.001	−0.440	−0.291
β_0:female	−0.357	0.047	<0.001	−0.450	−0.264
τ:base	1.139	0.029	<0.001	1.011	1.267
τ:female	−0.176	0.099	0.076	−0.370	0.018
σ_0	0.885	0.026	<0.001	0.828	0.943
σ_τ	1.108	0.046	<0.001	1.018	1.197
$E[\mu_{\omega n}]$:base	0.366	0.038	<0.001	0.291	0.440
$E[\mu_{\omega n}]$:female	0.357	0.047	<0.001	0.264	0.450
$E[\ln \sigma_{\omega n}]$:base	−1.139	0.029	<0.001	−1.267	−1.011
$E[\ln \sigma_{\omega n}]$:female	0.176	0.099	0.076	0.018	0.370
$SD[\mu_{\omega n}] = \sigma_0$	0.724	0.055	<0.001	0.616	0.832
$SD[\ln \sigma_{\omega n}] = \sigma_\tau$	0.415	0.023	<0.001	0.370	0.460

Notes: All models have been estimated using the Andersen, Harrison, Lau and Rutström [2014] data set. Standard errors have been adjusted for clustering at the subject level, except in panel C. The *gmnl* (version 1.1.0) command in *Stata* does not support clustered standard errors.

Table B5: Dual Latent Class Logit of EUT Model with CRRA Utility

(Log-likelihood = −5870.044)

Parameter	Estimate	Standard Error	*p*-value	95% Confidence Interval	
			A. Class 1		
β_0	−1.536	0.098	<0.001	− 1.727	− 1.344
β_1	4.297	0.178	<0.001	3.948	4.646
$\mu_\omega = -\beta_0/\beta_1$	0.357	0.025	<0.001	0.309	0.406
$\sigma_\omega = 1/\beta_1$	0.233	0.010	<0.001	0.214	0.252
Share:base	0.618	0.036	<0.001	0.547	0.690
Share:female	−0.101	0.051	0.047	−0.201	−0.001
			B. Class 2		
β_0	−2.325	0.174	<0.001	−2.666	−1.985
β_1	1.535	0.178	<0.001	1.185	1.884
$\mu_\omega = -\beta_0/\beta_1$	0.541	0.051	<0.001	0.442	0.640
$\sigma_\omega = 1/\beta_1$	0.652	0.076	<0.001	0.503	0.800
Share:base	0.215	0.033	<0.001	0.149	0.280
Share:female	0.140	0.046	0.003	0.049	0.231
			C. Class 3		
β_0	0.507	0.140	<0.001	0.234	0.781
β_1	0.490	0.094	<0.001	0.306	0.674
$\mu_\omega = -\beta_0/\beta_1$	−0.118	0.033	<0.001	−0.182	−0.054
$\sigma_\omega = 1/\beta_1$	2.041	0.390	<0.001	1.276	2.806
Share:base	0.167	0.026	<0.001	0.116	0.219
Share:female	−0.039	0.035	0.267	−0.107	0.030

Notes: The model has been estimated using the Andersen, Harrison, Lau and Rutström [2014] data set. Standard errors have been adjusted for clustering at the subject level.

Online Appendix C: Supporting Results for Section 4

Table C1: EUT Model with CRRA Utility ($\varkappa = 0.015$)

Parameter	Estimate	Standard Error	*p*-value	95% Confidence Interval	
			A. Representative Agent Model (Log-likelihood = −7572.459)		
μ_ω	0.512	0.036	<0.001	0.442	0.582
σ_ω	0.673	0.034	<0.001	0.607	0.739
			B. Random Coefficient Model (Log-likelihood = −5380.227)		
$E[\mu_{\omega n}]$	0.591	0.045	<0.001	0.502	0.679
$E[\ln \sigma_{\omega n}]$	−1.343	0.057	<0.001	−1.454	−1.231
$SD[\mu_{\omega n}]$	0.810	0.033	<0.001	0.745	0.875
$SD[\ln \sigma_{\omega n}]$	1.208	0.053	<0.001	1.105	1.312
$\varrho[\mu_{\omega n}, \ln \sigma_{\omega n}]$	0.129	0.028	<0.001	0.073	0.184

Notes: All models have been estimated using the Andersen, Harrison, Lau and Rutström [2014] data set. In the representative agent model, standard errors have been adjusted for clustering. In the random coefficient model, we do not adjust standard errors for clustering to make them comparable to standard errors in panel C of Table 1.

Table C2: RDU Model with CRRA Utility ($\varkappa$ = 0.015)

Parameter	Estimate	Standard Error	*p*-value	95% Confidence Interval	
			A. Representative Agent Model (Log-likelihood = −7414.671)		
μ_ω	0.635	0.051	<0.001	0.535	0.736
m_φ	0.224	0.078	0.004	0.071	0.377
σ_ω	0.643	0.050	<0.001	0.544	0.742
s_φ	1.017	0.046	<0.001	0.926	1.108
			B. Random Coefficient Model (Log-likelihood = −-5232.844)		
$E[\mu_{\omega n}]$	0.558	0.046	<0.001	0.468	0.648
$E[m_{\varphi n}]$	0.066	0.091	0.469	−0.112	0.244
$E[\sigma^{UN}_{\omega n}]$	0.309	0.049	<0.001	0.214	0.405
$E[s^{UN}_{\varphi n}]$	1.174	0.080	<0.001	1.017	1.331
$SD[\mu_{\omega n}]$	0.697	0.056	<0.001	0.588	0.806
$SD[m_{\varphi n}]$	1.174	0.080	<0.001	1.017	1.331
$SD[\sigma^{UN}_{\omega n}]$	0.367	0.045	<0.001	0.280	0.455
$SD[s^{UN}_{\varphi n}]$	0.013	0.074	0.862	−0.131	0.157
$\varrho[\mu_{\omega n}, \sigma^{UN}_{\omega n}]$	−0.308	0.065	<0.001	−0.434	−0.181

Notes: All models have been estimated using the Andersen, Harrison, Lau and Rutström [2014] data set. All standard errors have been adjusted for clustering at the subject level. The representative model assumes that the log-shape parameter is normally distributed: $\ln(\varphi_{nt}) \sim N(m_\varphi, s_\varphi^2)$. In the random coefficient model, $\sigma^{UN}_{\omega n}$ is a normally distributed random coefficient such that $|\sigma^{UN}_{\omega n}|$ is equal to $\sigma_{\omega n}$, the within-subject scale of ω_{nt}. $s^{UN}_{\varphi n}$ is similarly defined with respect to $s_{\varphi n}$, the within-subject standard deviation of $\ln(\varphi_{nt})$.

Table C3: EUT Model with Non-parametric Utility for {$0, $5, $10} and {$5, $10, $15}

Parameter	Estimate	Standard Error	p-value	95% Confidence Interval	
		A. Representative Agent Model for {$0, $5, $10} *(Log-likelihood = −747.763; ϰ = 0.012)*			
μ_ζ	1.082	0.110	<0.001	0.866	1.297
σ_ζ	0.727	0.078	<0.001	0.574	0.880
		B. Representative Agent Model for {$5, $10, $15} *(Log-likelihood = −704.222; ϰ = 0.013)*			
μ_ζ	−0.157	0.153	0.303	−0.457	0.142
σ_ζ	0.987	0.137	<0.001	0.719	1.255
		C. Random Coefficient Model for {$0, $5, $10} *(Log-likelihood = −637.393; ϰ = 0.012)*			
$E[\mu_{\zeta n}]$	1.056	0.107	<0.001	0.846	1.266
$E[\ln \sigma_{\zeta n}]$	−0.630	0.119	<0.001	−0.863	−0.398
$SD[\mu_{\zeta n}]$	0.753	0.094	<0.001	0.569	0.937
$SD[\ln \sigma_{\zeta n}]$	0.559	0.125	<0.001	0.315	0.804
		D. Random Coefficient Model for {$5, $10, $15} *(Log-likelihood = −640.025; ϰ = 0.013)*			
$E[\mu_{\zeta n}]$	−0.173	0.152	0.257	−0.471	0.126
$E[\ln \sigma_{\zeta n}]$	−0.099	0.139	0.477	−0.371	0.173
$SD[\mu_{\zeta n}]$	0.929	0.158	<0.001	0.619	1.238
$SD[\ln \sigma_{\zeta n}]$	0.023	0.220	0.918	−0.409	0.455

Notes: All models have been estimated using the Harrison and Rutström [2008] data set. All standard errors have been adjusted for clustering at the subject level. In the representative agent models, standard errors have been adjusted for clustering. In the random coefficient models, we do not adjust standard errors for clustering to make them comparable to standard errors in the bottom panels, indexed by C, in Tables B1 and B2.

Table C4: EUT Model with Non-parametric Utility for {\$0, \$5, \$10, \$15} ($\varkappa$ = 0.010)

Parameter	Estimate	Standard Error	*p*-value	95% Confidence Interval	
			A. Representative Agent Model (Log-likelihood = −2357.540)		
$\mu\{\zeta2\}$	1.092	0.181	<0.001	0.738	1.446
$\mu\{\zeta3\}$	−0.143	0.082	0.081	−0.305	0.018
$\sigma\{\zeta2\}$	0.980	0.138	<0.001	0.709	1.250
$\sigma\{\zeta3\}$	0.703	0.090	<0.001	0.526	0.881
			B. Random Coefficient Model (Log-likelihood = −2052.437)		
$E[\mu\{\zeta2n\}]$	1.254	0.114	<0.001	1.030	1.478
$E[\mu\{\zeta3n\}]$	−0.144	0.086	0.096	−0.313	0.025
$E[\ln \sigma\{\zeta2n\}]$	−0.165	0.165	0.318	−0.488	0.158
$E[\ln \sigma\{\zeta3n\}]$	−1.164	0.296	<0.001	−1.744	−0.585
$SD[\mu\{\zeta2n\}]$	1.365	0.190	<0.001	0.993	1.736
$SD[\mu\{\zeta3n\}]$	0.531	0.075	<0.001	0.383	0.679
$SD[\ln \sigma\{\zeta2n\}]$	0.611	0.120	<0.001	0.376	0.847
$SD[\ln \sigma\{\zeta3n\}]$	0.851	0.220	<0.001	0.419	1.282

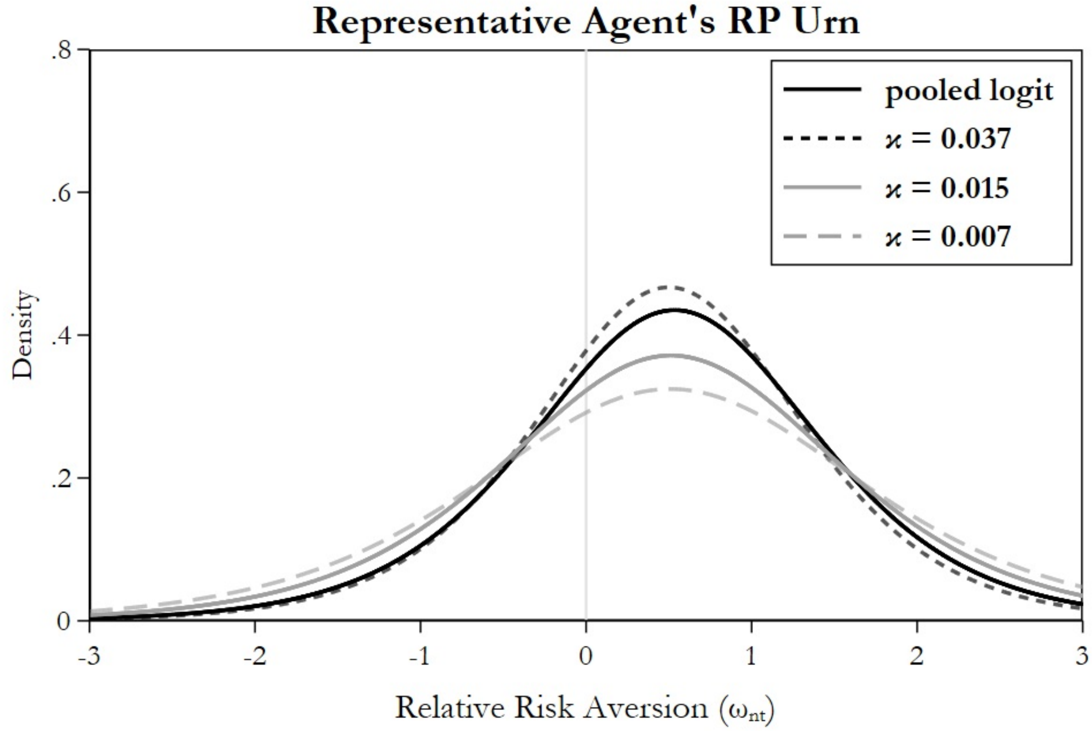
Notes: All models have been estimated using the Harrison and Rutström [2008] data set. All standard errors have been adjusted for clustering at the subject level. $\mu\{\zeta2\}$ refers to the inverted utility parameter μ subscripted by $\zeta2$. Other parameter labels can be interpreted similarly.

Table C5: RDU Model with Non-parametric Utility for {$0, $5, $10, $15} ($\varkappa$ = 0.010)

Parameter	Estimate	Standard Error	*p*-value	95% Confidence Interval	
			A. Representative Agent Model *(Log-likelihood = −2346.771)*		
$\mu\{\zeta 2\}$	1.244	0.158	<0.001	0.934	1.553
$\mu\{\zeta 3\}$	−0.065	0.075	0.389	−0.212	0.083
m_φ	−0.087	0.046	0.059	−0.178	0.003
$\sigma\{\zeta 2\}$	0.906	0.110	<0.001	0.691	1.121
$\sigma\{\zeta 3\}$	0.619	0.061	<0.001	0.500	0.737
s_φ	0.564	0.089	<0.001	0.390	0.739
			B. Random Coefficient Model *(Log-likelihood = −-2004.213)*		
$E[\mu\{\zeta 2n\}]$	1.097	0.101	<0.001	0.900	1.294
$E[\mu\{\zeta 3n\}]$	−0.055	0.104	0.593	−0.259	0.148
$E[m_{\varphi n}]$	0.013	0.035	0.710	−0.055	0.081
$E[\sigma^{UN}\{\zeta 2n\}]$	0.529	0.109	<0.001	0.316	0.743
$E[\sigma^{UN}\{\zeta 3n\}]$	−0.611	0.133	<0.001	−0.871	−0.350
$E[s^{UN}_{\varphi n}]$	0.474	0.081	<0.001	0.316	0.632
$SD[\mu\{\zeta 2n\}]$	1.317	0.173	<0.001	0.977	1.657
$SD[\mu\{\zeta 3n\}]$	0.533	0.120	<0.001	0.297	0.770
$SD[m_{\varphi n}]$	0.482	0.072	<0.001	0.340	0.624
$SD[\sigma^{UN}\{\zeta 2n\}$	0.122	0.051	0.017	0.022	0.223
$SD[\sigma^{UN}\{\zeta 3n\}]$	0.566	0.145	<0.001	0.281	0.850
$SD[s^{UN}_{\varphi n}]$	0.428	0.072	<0.001	0.287	0.568

Notes: All models have been estimated using the Harrison and Rutström [2008] data set. All standard errors have been adjusted for clustering at the subject level. The representative agent model assumes that the log-shape parameter is normally distributed: $\ln(\varphi_{nt}) \sim N(m_\varphi, s_\varphi^2)$. $\mu\{\zeta 2\}$ refers to the inverted utility parameter μ subscripted by $\zeta 2$. Other parameter labels can be interpreted similarly. In the random coefficient model, $\sigma^{UN}\{\zeta 2n\}$ is a normally distributed random coefficient such that $|\sigma^{UN}\{\zeta 2n\}|$ is equal to $\sigma\{\zeta 2n\}$, the within-subject scale of $\zeta\{2nt\}$. $\sigma^{UN}\{\zeta 3n\}$ and $s^{UN}_{\varphi n}$ are similarly defined with respect to $\sigma\{\zeta 3n\}$ and $s_{\varphi n}$, respectively, where the latter refers to the within-subject standard deviation of $\ln(\varphi_{nt})$.

Figure C1: Kernel-Smoothed EUT Model with CRRA Utility



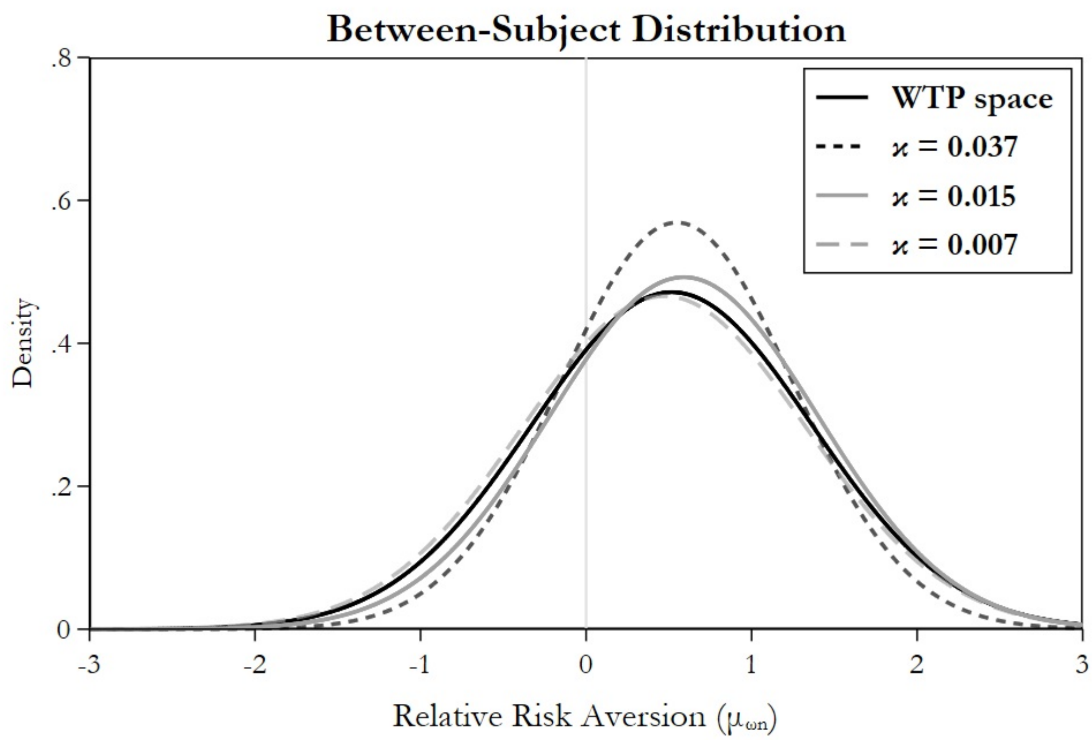
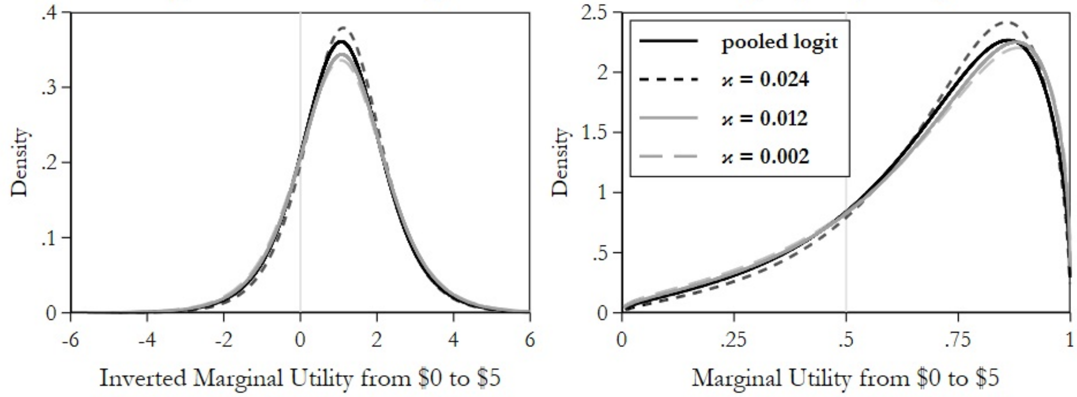


Figure C2: Kernel-Smoothed EUT Model with Non-parametric Utility

Representative Agent's RP Urn for {$0, $5, $10}



Representative Agent's RP Urn for {$5, $10, $15}

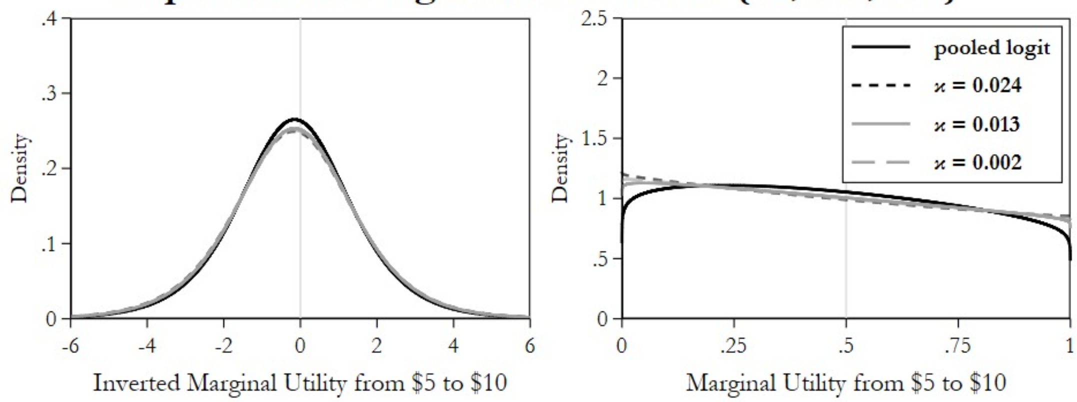
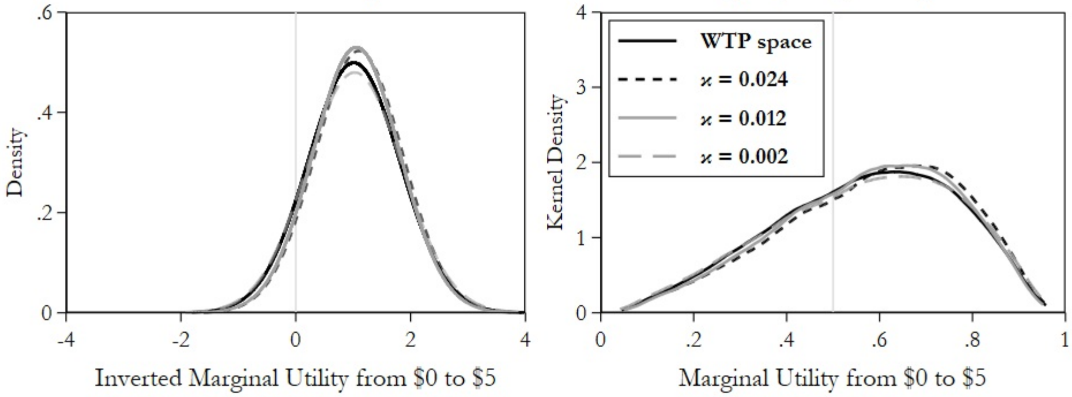


Figure C3: Kernel-Smoothed EUT Model with Non-parametric Utility

Between-Subject Distribution for {$0, $5, $10}



Between-Subject Distribution for {$5, $10, $15}

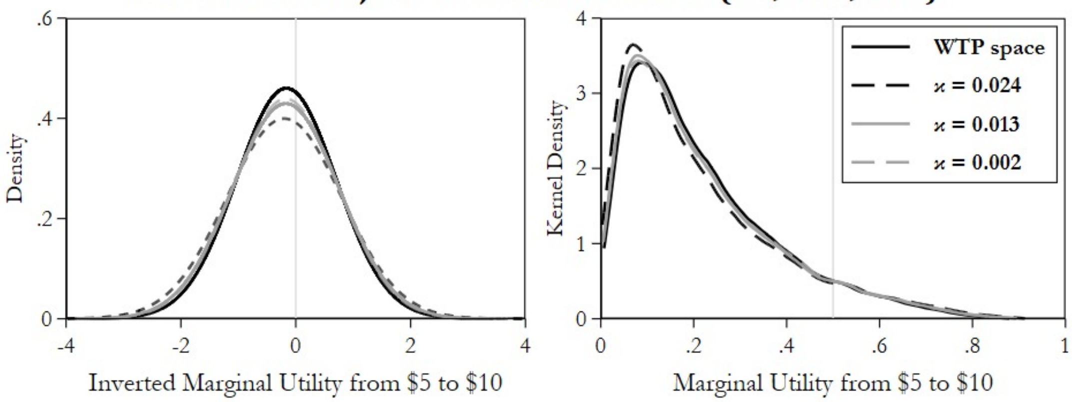


Figure C4: Kernel-Smoothed EUT Model with Non-parametric Utility

Between-Subject Distributions for {$0, $5, $10, $15}