RANDOM DISCOUNTED EXPECTED UTILITY*

JOSE APESTEGUIA†, MIGUEL A. BALLESTER‡, AND ÁNGELO GUTIÉRREZ-DAZA§

Abstract. This paper introduces the random discounted expected utility (R-DEU) model, which we have developed as a means to deal with heterogeneous risk and time preferences. The R-DEU model provides an explicit linkage between preference and choice heterogeneity. We prove it has solid comparative statics, discuss its identification, and demonstrate its computational convenience. Finally, we use two distinct experimental datasets to illustrate the advantages of the R-DEU model over common alternatives for estimating heterogeneity in preferences across individuals.

Keywords: Heterogeneity; Risk and Time Preferences; Comparative Statics; Random Utility Models.

JEL classification numbers: C01; D01.

1. Introduction

Economic situations simultaneously involving risk and time pervade most spheres of everyday life, and heterogeneity of behavior is the rule. In this paper, we develop a model for the treatment of heterogeneous risk and time preferences. For standard experimental design environments, we establish the model's predicted choice probabilities and show that it has intuitive comparative statics. We also demonstrate that it is easily implementable in practice, and that it accounts remarkably well for the observed heterogeneity of choice in two key experimental designs. Overall, we provide a well-founded and convenient novel framework for the analysis of heterogeneous risk and time preferences.

Date: July, 2024.

*This paper supersedes our paper "Random Utility Models for the Joint Treatment of Risk and Time Preferences." We thank Daniel Benjamin, Holger Gerhardt, and Jesse Shapiro for helpful comments, and Rafael Suchy for outstanding research assistance. Financial support by FEDER/Ministerio de Ciencia de Innovación (Agencia Estatal de Investigación) through Grant PGC2018-098949-B-I00 and through the Severo Ochoa Program for Centers of Excellence in R&D (Barcelona School of Economics CEX2019-000915-S), and Balliol College is gratefully acknowledged. Codes to replicate all the empirical results in this paper are available at https://github.com/agutieda/R-DEU. The views in this paper do not necessarily reflect those of Banco de México.

† ICREA, Universitat Pompeu Fabra and BSE. E-mail: `jose.apesteguia@upf.edu`.

‡ University of Oxford. E-mail: `miguel.ballester@economics.ox.ac.uk`.

§ Banco de México. E-mail: `angelo.gutierrez@banxico.org.mx`.

Our stochastic model is based on a probability distribution over a given collection of utility functions. This enables us to establish a direct link between preference and choice heterogeneity. We adopt the most standard family of utilities for the treatment of risk and time, namely, discounted expected utilities, and thus name the model random discounted expected utility (R-DEU).[1] We study it under the two main experimental risk and time elicitation mechanisms: double multiple price lists (DMPL) and convex budgets (CB). The sharp contrast between these two mechanisms, one involving binary choices and the other a continuous choice space, enables us to show that the model is very flexible.

The adoption of random utility models (RUMs) with stochastic preference parameters in empirical applications has been slow partly due to their computational complexity. The computation of choice probabilities in these models involves numerical integration over multiple variables, difficulting the analytical study of their properties and their econometric identification. In the case of discounted expected utility, this demands integrating the joint distribution of two variables: the discounting factor and the curvature of the monetary utility function. We show, however, that the R-DEU model is analytically tractable: given any curvature of the monetary function, we prove that there is always an ordered structure linking discounting and choices. Thus, the conditional choice probabilities for any given curvature can be computed straightforwardly and then easily aggregated, rendering the model theoretically and empirically convenient.

Using the above conditional choice probability approach, we then establish, for the first time, the stochastic comparative statics of the R-DEU. We analyze shifts and spreads of the probability distribution over the two main components of discounted expected utility: curvature and discounting. Although the theoretical treatment of comparative statics involving more than one parameter is challenging, the results are consistent with common understanding. First, we find that a shift in the probability distribution towards higher discounting has an effect only in problems involving time, where it shifts choices towards earlier options. Second, a shift in the probability distribution towards larger curvatures has an effect in all types of problems: generating choice shifts (i) towards safer options in multiple price lists involving risk, (ii) towards earlier options in multiple price lists

[1]Not to be confused with the Rank-Dependent Expected Utility Model introduced by Quiggin (1982).

involving time, and (iii) towards smoother consumptions in convex sets. Furthermore, a wider spread in any variable in the probability distribution leads to higher choice stochasticity. These results are fundamental in providing the economic literature with a well-founded framework for the proper interpretation and estimation of the variables of interest, i.e., discounting and curvature.

All the former results are for general discounted expected utility representations and unrestricted probability distributions. We then discuss the implications of these results for standard parameterizations. Following common practice in the literature, we consider constant relative risk aversion (CRRA) monetary utility functions, which are determined by a parameter r describing the curvature.[2] Hence, every utility is characterized by a pair of parameters (r, δ), where δ captures discounting. To give further intuition of the properties of the model, its identification, and empirical implementation, we also consider the standard special case where r and δ follow a bivariate normal distribution. This added parametric structure further accentuates the convenience of the model: the choice computation requires the evaluation of conditional and marginal distributions of a bivariate normal, which are themselves normal distributions. It follows that the parametric assumption reduces the dimensionality of the numerical problem, making the computation of choice probabilities routine. Moreover, shifts and spreads in the probability distribution are the result of variations in the first two moments of the distribution, facilitating the identification in the parametric case.

After establishing the theoretical grounds of the model, we illustrate its empirical advantages with a structural estimation exercise using data from two major exemplars of the type of elicitation mechanisms considered: Andersen,Harrison, Lau, and Rutstrom (2008) (hereafter AHLR) and Andreoni and Sprenger (2012b) (hereafter AS). We compare the aggregate and individual-level estimates of the R-DEU model with those obtained using the empirical strategies employed in the respective papers and subsequent literature.

[2]The literature contemplates a variety of formulations of this utility function, some of which have perverse implications after introducing time considerations. In Section 2.1, we discuss the necessary conditions for the appropriate use of a CRRA family in an environment involving both risk and time.

The literature has often used iid-additive RUMs to analyze DMPL experimental designs, where decision problems are binary, involve only risk or time considerations, and a single-dated lottery defines alternatives. Theoretically, implementing this approach with standard representations of discounted expected utility could lead to paradoxical predictions and perverse comparative statics properties (Wilcox, 2011; Apesteguia and Ballester, 2018), thus hindering a thorough understanding of risk and time preferences. Moreover, we show empirically that the R-DEU model offers a better overall fit. It also performs better than recent iid-additive RUM implementations using Wilcox's (2011) correction. There are stark differences across models at the individual level: the R-DEU model delivers reasonable estimates of risk and time preferences, which are highly correlated with commonly used estimates obtained from decision switching within risk and time tasks in DMPL designs. On the contrary, the iid-additive RUMs are only weakly correlated with these semi-parametric estimates and take implausible values for a specific subset of individuals that we identify.

In CB experimental designs, menus are continuous, involve risk and time considerations, and each alternative grants a pair of dated lotteries. For this experimental design, the literature has often relied on estimating risk and time preferences using non-linear least squares, assuming a unique discounted expected utility. To do so, researchers introduce randomness by perturbing the first-order condition of a constrained utility-maximization problem. The randomness introduced in this approach lacks a behavioral foundation in that it does not explicitly connect heterogeneity of choice with heterogeneity of preferences. Moreover, this approach is not well suited to understanding the large heterogeneity in choices observed in the data. Another approach in the analysis of CB datasets uses, as in the case of DMPLs, iid-additive RUMs. Empirically, this multinomial extension tends to deliver estimated utility functions that are convex as a way to explain the pervasive share of corner solutions observed in CB settings, at the cost of leaving unexplained the large fraction of choices in the intermediate range of budget sets. In contrast to these approaches, we show how R-DEU empirically accounts for the observed prevalence of corner and interior choices while delivering plausible estimates of discounting and the curvature of the utility function.

To conclude, the theoretical results and empirical applications illustrate the usefulness of the R-DEU model as a robust and unifying framework for estimating risk and time preferences with experimental data while accounting for the large heterogeneity in choices between and across individuals. The entire exercise of the paper is related to recent methodological literature on preference estimations in a variety of settings (see, e.g., DellaVigna, 2018; Cattaneo et al., 2020; Dardanoni et al., 2020; Aguiar and Kashaev, 2021; Barseghyan et al., 2021; Lau and Yoo, 2023). Our paper stands apart from this literature in that it focuses jointly on risk and time preferences and establishes the comparative statics of the model.

2. Random Discounted Expected Utility

A lottery is a finite collection of monetary prizes and associated probabilities, i.e., a vector of the form $l = [p_1, \ldots, p_n, \ldots, p_N; x_1, \ldots, x_n, \ldots x_N]$, with $p_n \geq 0$, $\sum_{n=1}^{N} p_n = 1$, and $x_n \geq 0$. A dated lottery (l, t) is formed by a lottery and a moment in time $t \geq 0$, in which the resulting prize is awarded.[3]

Discounted expected utility (DEU) is the most commonly-used deterministic model of behavior for the study of risk and time preferences. We consider a family $\{u_r\}_{r\in\mathbb{R}}$ of continuous and strictly increasing utility functions over money, that are normalized to satisfy $u_r(\omega) = 0$ at a baseline wealth level $\omega > 0$. We impose three basic assumptions on the family of monetary utilities. First, it must include the linear monetary utility, that we denote by $r = 0$. Second, the family is strictly ordered by concavity, i.e., $r < r'$ means that $u_{r'}$ is "strictly more concave" than u_r.[4] Third, convexity and concavity are unbounded when r tends to $-\infty$ and $+\infty$, respectively. Many families satisfy these basic requirements, including the widely used CRRA or the constant absolute risk aversion (CARA) utility functions. The discount factor of the individual is denoted by $e^{-\delta}$ with

[3]When necessary, we denote the prizes, payoffs, and the number of outcomes of dated lotteries $j = 1, \ldots, J$ as x_n^j, p_n^j, and N_j. We also assume, as it is typically done, that the awarded monetary prizes are consumed on reception.

[4]Formally, u is "strictly more concave" than u' if there exists an increasing and strictly concave function ϕ such that $u'(x) = \phi(u(x))$ for every x. As a result, utilities with $r > 0$ (resp., $r < 0$) represent risk aversion (resp., risk loving). As shown in Pratt (1964), this condition is equivalent to the condition that the certainty equivalent of u is strictly lower than the certainty equivalent of u' for any lottery l. See also Proposition 6.C.2 in Mas-Colell, Whinston and Greene (1995) and Propositions 6.7 and 6.8 in Kreps (2013).

$\delta \in \mathbb{R}$.[5] Given parameters $(r, \delta) \in \mathbb{R}^2$, the DEU evaluation of a sequence of dated lotteries $(l^1, t^1), \ldots, (l^J, t^J)$ is:

$$DEU_{r,\delta}((l^1, t^1), \ldots, (l^J, t^J)) = \sum_{j=1}^{J} e^{-\delta t^j} \sum_{n=1}^{N_j} p_n^j u_r(\omega + x_n^j). \tag{2.1}$$

We are now in a position to define the stochastic model that we analyze in the paper, that we call Random Discounted Expected Utility (R-DEU) model. Let f be a measurable density with full support over $\mathbb{R}^2$, capturing the prevalence of each possible DEU preference. At the moment of choice from a decision problem, parameters (r, δ) are realized with probability $f(r, \delta)$, and the alternative that maximizes $DEU_{r,\delta}$ within the decision problem is selected.[6]

We now describe formally the decision problems involved in the two settings analyzed in this paper. One of the most prominent settings in the experimental literature involves the use of the so-called double multiple price lists (DMPLs) as in Andersen et al. (2008), where decision problems are binary, involve only either risk or time considerations, and alternatives are defined by a single dated lottery.[7] In a risk decision problem, each of the two alternatives corresponds to a single two-state contingent lottery with prizes awarded in the present. That is, given $x_1^1 > x_1^0 > x_2^0 > x_2^1$ and $p \in (0, 1)$, the associated risk menu is $A_{\mathcal{R}} = \{0_{\mathcal{R}}, 1_{\mathcal{R}}\}$ where $0_{\mathcal{R}} = ([p, 1-p; x_1^0, x_2^0], 0)$ and $1_{\mathcal{R}} = ([p, 1-p; x_1^1, x_2^1], 0)$.[8] In a time decision problem, each of the two alternatives is composed by a unique dated degenerate lottery. That is, given $t^0 < t^1$ and $x^0 < x^1$, the associated time menu is $A_{\mathcal{T}} = \{0_{\mathcal{T}}, 1_{\mathcal{T}}\}$ where $0_{\mathcal{T}} = ([1; x^0], t^0)$ and $1_{\mathcal{T}} = ([1; x^1], t^1)$. Given the binary nature of risk and time menus, the R-DEU choice probabilities in any decision problem are determined by the choice probability of one of the two alternatives in the menu, say $0_{\mathcal{R}}$ in a risk menu and $0_{\mathcal{T}}$

---

[5] Hence, $\delta > 0$ (resp., $\delta < 0$) represents impatience or delay aversion (resp., delay loving). We write the discount factor in this way for convenience; it allows us to use a simple bivariate normal in the parametric estimation. Note that, alternatively, we could simply write $(1 + d)^{-1} = e^{-\delta} \in \mathbb{R}_{++}$ with d representing the discount factor in the positive reals.

[6] Given that f is assumed to be measurable, indifferences between maximal alternatives are inessential, and will be obviated in the paper.

[7] See also Burks et al. (2009), Dohmen et al. (2010), Tanaka et al. (2010), Benjamin et al. (2013), Falk et al. (2018) or Jagelka (2021). In Appendix B, we study a hybrid version where both risk and time considerations are simultaneously active. (see, Ahlbrecht and Weber (1997), Coble and Lusk (2010), Baucells and Heukamp (2012) and Cheung (2015)).

[8] Sometimes, $p \in \{0, 1\}$ is considered. These cases are trivial since one of the two lotteries is dominated and, hence, predicted a zero probability of choice by R-DEU.

in a time menu. Denote by $\Gamma(0_{\mathcal{R}}, A_{\mathcal{R}}) \subseteq \mathbb{R}^2$ (resp., $\Gamma(0_{\mathcal{T}}, A_{\mathcal{T}}) \subseteq \mathbb{R}^2$) the collection of all parameter combinations (r, δ) for which the maximizer of $DEU_{r,\delta}$ within menu $A_{\mathcal{R}}$ (resp., $A_{\mathcal{T}}$) is $0_{\mathcal{R}}$ (resp., $0_{\mathcal{T}}$). The f-measures of these sets describe the choice probabilities:

$$\mathcal{P}_f(0_{\mathcal{R}}, A_{\mathcal{R}}) = \int_{\Gamma(0_{\mathcal{R}}, A_{\mathcal{R}})} f(r, \delta)\mathrm{d}(r, \delta),$$

$$\mathcal{P}_f(0_{\mathcal{T}}, A_{\mathcal{T}}) = \int_{\Gamma(0_{\mathcal{T}}, A_{\mathcal{T}})} f(r, \delta)\mathrm{d}(r, \delta).$$

In an alternative setting pioneered by Andreoni and Sprenger (2012a,b), subjects are faced with the so-called convex menus.[9] These menus are continuous, involve risk and time considerations, and each alternative grants a pair of dated lotteries. Formally, given $x^0 \leq x^1$, $t^0 < t^1$ and $p^0, p^1 \in (0, 1]$, the associated convex menu is $A_{\mathcal{C}} = [0, 1]$ where alternative $a \in A_{\mathcal{C}}$ is defined by the sequence of two dated lotteries $([p^0, 1 - p^0; (1 - a)x^0, 0], t^0), ([p^1, 1 - p^1; ax^1, 0], t^1)$. Given the continuous nature of convex menus, the R-DEU choice probabilities are determined by the cumulative choice probability of selecting alternatives below any given value $a \in [0, 1]$. Denote by $\Gamma([0, a], A_{\mathcal{C}}) \subseteq \mathbb{R}^2$ the collection of all parameter combinations (r, δ) for which the maximizer of $DEU_{r,\delta}$ within menu $A_{\mathcal{C}}$ is an alternative below a. The f-measures of these sets describe the choice probabilities:

$$\mathcal{P}_f([0, a], A_{\mathcal{C}}) = \int_{\Gamma([0,a], A_{\mathcal{C}})} f(r, \delta)\mathrm{d}(r, \delta).$$

2.1. **Parametric Version.** All our theoretical results are for general discounted expected utility representations and unrestricted probability distributions. However, specific parameterizations are often times useful, both as an illustration of the main insights of a theoretical result and as a practical tool in an estimation exercise. We illustrate the intuition of every theoretical result using CRRA monetary utility functions.

Given parameter $r \in \mathbb{R}$, the CRRA utility evaluation of an extra prize $x \geq 0$ is:

$$u_r^{crra}(\omega + x) = \begin{cases} \frac{(\omega + x)^{1-r} - \omega^{1-r}}{1-r} & \text{whenever } r \neq 1; \\ \log(\omega + x) - \log \omega & \text{otherwise.} \end{cases}$$

[9] Convex menus are being used extensively for the study of a variety of economic preferences. See, e.g., Choi et al. (2007), Fisman et al. (2007), Augenblick et al. (2015), Carvalho et al. (2016), Alan and Ertac (2018), and Kim et al. (2018).

Let us briefly comment on the rationale behind the constants chosen for the CRRA family, since the literature contemplates many different formulations and not all of them are appropriate when both risk and time are involved. First, as in the case of the study of risk preferences alone, parameter r can be assumed to belong to $\mathbb{R}$, but note that this necessitates the baseline wealth assumption $\omega > 0$. Otherwise, monetary utilities $\{u_r^{crra}\}_{r\geq 1}$ would not be well-defined for null prizes. Second, we then need to guarantee that all monetary utilities are strictly increasing and, hence, the raw power function $(\omega + x)^{1-r}$ must be re-scaled with the constant $\frac{1}{1-r}$. Third, since this re-scaling creates negative utilities whenever $r > 1$, the addition of the constant $-\frac{\omega^{1-r}}{1-r}$ guarantees positive utilities, makes $u_r^{crra}(\omega) = 0$, facilitating the analysis of lotteries involving null prizes, and implies the standard continuity property $\lim_{r\to 1} u_r^{crra}(\omega + x) = u_1^{crra}(\omega + x)$.[10]

When focusing on the parametric case, we will also impose some restrictions on the probability distribution f. The computational methods discussed in Appendix C allow the efficient estimation of the model for any distribution characterized by a finite vector of parameters $\theta \in \Theta$.[11] We illustrate with the case where (r, δ) follows a bivariate normal distribution, so that $\theta \equiv (\mu_r, \sigma_r, \mu_\delta, \sigma_\delta, \rho)$, where μ_z and σ_z are the corresponding mean and standard deviation of parameter $z \in \{r, \delta\}$, and ρ is the correlation coefficient between r and δ. This assumption provides a natural benchmark to compare the model and the empirical results in the following sections to other models in the literature. As we show below, it also allows for simple expressions of the choice probabilities in the model, which we will exploit to provide conditions allowing the identification of θ in each experimental setting.

3. Double Multiple Price Lists: Theory

3.1. **Risk Menus.** Given that in these decision problems all the action takes place in the present, the discount parameter δ plays no role. Moreover, the type of lotteries at stake always creates an intuitive, ordered, structure of choices for parameter r. For every risk menu $A_{\mathcal{R}}$ we show below that there is a real-value constant $K(A_{\mathcal{R}})$ such that alternative $0_{\mathcal{R}}$ is chosen if and only if $r \geq K(A_{\mathcal{R}})$. Hence, the choice probability of alternative $0_{\mathcal{R}}$

---

[10] A discussion on the role of wealth ω in CRRA utilities can be read in Appendix E.

[11] We illustrate an example with truncated normal and beta distributions Appendix D.

is the f-measure of the rectangular set $\Gamma(0_{\mathcal{R}}, A_{\mathcal{R}}) = \{(r, \delta) : r \geq K(A_{\mathcal{R}})\}$, that can be conveniently computed by using the marginal CDF of r, denoted F^r.

Moreover, comparative statics related to shifts and spreads of parameter r follow immediately, and are in full alignment with our most basic intuitions. When the mass of the marginal distribution of r is shifted towards larger values, the choice probability of the safer alternative is guaranteed to strictly increase. When the mass of the marginal distribution of r is brought away from its median, the choice probability of both alternatives strictly approaches one half, i.e., behavior becomes strictly more stochastic.[12]

To formalize these ideas, we need to define standard domination and expansion notions using CDFs. Formally, let F and G be two CDFs over the random variable z, with domain in an open interval, and denote by $\text{med}(F)$ the median of distribution F. Then, we say that: (i) F dominates G if $F(z) < G(z)$ holds for all values of z and (ii) F expands G if $\text{med}(F) = \text{med}(G)$, $F(z) > G(z)$ whenever $z < \text{med}(F)$ and $F(z) < G(z)$ whenever $z > \text{med}(F)$.[13]

Proposition 1. *For every pair of R-DEUs, f and g, and every menu $A_{\mathcal{R}}$:*

(1) $\mathcal{P}_f(0_{\mathcal{R}}, A_{\mathcal{R}}) = 1 - F^r(K(A_{\mathcal{R}}))$.

(2) *If F^r dominates G^r, $\mathcal{P}_f(0_{\mathcal{R}}, A_{\mathcal{R}}) > \mathcal{P}_g(0_{\mathcal{R}}, A_{\mathcal{R}})$.*

(3) *If F^r expands G^r with $K(A_{\mathcal{R}}) \neq$ med(F^r), $|\mathcal{P}_f(0_{\mathcal{R}}, A_{\mathcal{R}}) - \frac{1}{2}| < |\mathcal{P}_g(0_{\mathcal{R}}, A_{\mathcal{R}}) - \frac{1}{2}|$.*

3.2. **Time Menus.** When time is at stake, understanding behavior is slightly more complicated because the discount parameter δ, on its own, is hardly informative about behavior.

Example 1. *Let two DEU–CRRA individuals with $\omega = 100$ and preference parameters $(r_1, \delta_1) = (0.95, 0.094)$ and $(r_2, \delta_2) = (0, 0.105)$. Although $\delta_1 < \delta_2$ (or equivalently $e^{-\delta_1} = 0.91 > 0.9 = e^{-\delta_2}$) may suggest that individual 1 is more patient, it is immediate to see that she is indeed the only one that prefers $([1; 71.5], 0)$ to $([1; 80], 1)$.*

The joint consideration of both parameters is then required to fully understand the predictions of DEU, and consequently, the R-DEU choice probabilities would now require

[12] There is an obvious exception to this principle when choice stochasticity is already maximal, with both alternatives being chosen with the same probability 1/2. This happens when the median of F^r coincides with the separating threshold $K(A_{\mathcal{R}})$.

[13] The proof of every proposition is available in Appendix A, which is available in this link.

to compute a double integration. Fortunately, we show below that the analysis renders again simple after conditioning on parameter r, because this always generates an intuitive, ordered, structure of choices over the discounting parameter δ. For any given time menu $A_{\mathcal{T}}$ and any value of r, we show below that there exists a menu-dependent constant $K(A_{\mathcal{T}}|r) \in \mathbb{R}_+$ such that the earlier alternative $0_{\mathcal{T}}$ is selected if and only if $\delta \geq K(A_{\mathcal{T}}|r)$, i.e., $\Gamma(0_{\mathcal{T}}, A_{\mathcal{T}}) = \{(r, \delta) : \delta \geq K(A_{\mathcal{T}}|r)\}$. As a result, the choice probability of alternative $0_{\mathcal{T}}$ can be conveniently expressed by means of the choice probabilities conditional on parameter r. In short, we evaluate the conditional CDFs of parameter δ on parameter r, that we denote by $F_{\delta|r}$, at the corresponding threshold $K(A_{\mathcal{T}}|r)$, and then aggregate across values of r using its marginal density, that we denote by f^r. Proposition 2 builds upon this ordered structure, showing that the thresholds $\{K(A_{\mathcal{T}}|r)\}_{r\in\mathbb{R}}$ are strictly decreasing in r, and constitute a bijection from $\mathbb{R}$ to $\mathbb{R}_{++}$, which can thus be inverted.[14] Hence, comparative statics of shifts are immediate, as keeping constant the marginal distribution of r (resp., δ), and shifting upwards the conditional distributions of δ (resp., r) guarantee an increase in the choice probability of the earlier alternative $0_{\mathcal{T}}$. Second, with respect to spreads, we can again show that keeping constant the marginal distribution of one parameter, an expansion of the conditional distributions of the other always creates more stochasticity.[15]

Proposition 2. *For every pair of R-DEUs, f and g, and every menu $A_{\mathcal{T}}$:*

(1) $\mathcal{P}_f(0_{\mathcal{T}}, A_{\mathcal{T}}) = 1 - \int_r F_{\delta|r} K(A_{\mathcal{T}}|r) f^r(r)\mathrm{d}r = 1 - \int_{\delta>0} F_{r|\delta} K(A_{\mathcal{T}}|\delta) f^\delta(\delta)\mathrm{d}\delta$.

(2) (a) *If $F^r = G^r$, and for all r $F_{\delta|r}$ dominates $G_{\delta|r}$, $\mathcal{P}_f(0_{\mathcal{T}}, A_{\mathcal{T}}) \geq \mathcal{P}_g(0_{\mathcal{T}}, A_{\mathcal{T}})$.*

(b) *If $F^\delta = G^\delta$, and for all δ $F_{r|\delta}$ dominates $G_{r|\delta}$, $\mathcal{P}_f(0_{\mathcal{T}}, A_{\mathcal{T}}) \geq \mathcal{P}_g(0_{\mathcal{T}}, A_{\mathcal{T}})$.*

(3) (a) *If $F^r = G^r$, and for all r $F_{\delta|r}$ expands $G_{\delta|r}$ with $K(A_{\mathcal{T}}|r) \neq med(F_{\delta|r})$, $|\mathcal{P}_f(0_{\mathcal{T}}, A_{\mathcal{T}}) - \frac{1}{2}| < |\mathcal{P}_g(0_{\mathcal{T}}, A_{\mathcal{T}}) - \frac{1}{2}|$.*

(b) *If $F^\delta = G^\delta$, and for all δ $F_{r|\delta}$ expands $G_{r|\delta}$ with $K(A_{\mathcal{T}}|\delta) \neq med(F_{r|\delta})$, $|\mathcal{P}_f(0_{\mathcal{T}}, A_{\mathcal{T}}) - \frac{1}{2}| < |\mathcal{P}_g(0_{\mathcal{T}}, A_{\mathcal{T}}) - \frac{1}{2}|$.*

---

[14] In other words, conditioning on δ also renders ordered choices over parameter r. Whenever $\delta \leq 0$, $1_{\mathcal{T}}$ is always chosen. Whenever $\delta > 0$ there is a menu-dependent constant $K(A_{\mathcal{T}}|\delta) \in \mathbb{R}$ such that $0_{\mathcal{T}}$ is chosen if and only if $r \geq K(A_{\mathcal{T}}|\delta)$.

[15] As in the case of risk, the median of each conditional distribution $F_{\delta|r}$ must be different to the corresponding threshold $K(A_{\mathcal{T}}|r)$ when expansions of δ are considered, with an analogous expression for the case of r.

3.3. **Implications for the Parametric Version.** The general results of Propositions 1 and 2 have the following implications when using the particular case of CRRA and the bivariate normal. In the case of CRRA, the thresholds described in Proposition 1 simply correspond to the unique value of r that solves the equation $\frac{1-p}{p} = \frac{(\omega+x_1^1)^{1-r}-(\omega+x_1^0)^{1-r}}{(\omega+x_2^0)^{1-r}-(\omega+x_2^1)^{1-r}}$. In the bivariate normal, the marginal distribution of parameter r is normally distributed, with parameters μ_r and σ_r. Putting both things together, part 1 states that the analysis of choice probabilities in R-DEU is a straightforward computational exercise. Moreover, dominating shifts and expansions of F^r are the result of an increase in, respectively, μ_r and σ_r. Hence, parts 2 and 3 inform the analyst that straightforward intuitions are in place. An increase in the median of parameter r creates always a larger probability of choice for the safer alternative, while an increase in the variance of parameter r generates more choice stochasticity.

Similarly, we can read Proposition 2 from the parametric point of view. With CRRA, the threshold map can be written as $K(A_{\mathcal{T}}|r) = \frac{1}{t^1-t^0} \log \left[\frac{(\omega+x^1)^{1-r}-\omega^{1-r}}{(\omega+x^0)^{1-r}-\omega^{1-r}}\right]$. With the bivariate normal, all conditionals $F_{\delta|r}$ are also normal, with mean $\mu_\delta + \frac{\sigma_\delta}{\sigma_r}\rho(r-\mu_r)$ and standard deviation $\sqrt{1-\rho^2} \cdot \sigma_\delta$. This, combined with the already-mentioned normality of f^r makes the computation of probabilities a straightforward exercise. Moreover, considering $z, z' \in \{r, \delta\}$ with $z \neq z'$, an increase of μ_z leaves unaffected the marginal $F^{z'}$ while generating a dominating shift in all conditionals $F_{z|z'}$. Hence, part 2 states that, by increasing either the mean of δ or the mean of r, we generate a larger choice probability for the earlier alternative. Third, increasing σ_z leaves unaffected the marginal $F^{z'}$ and, under the appropriate correction of the covariance, it generates the expansion of all conditionals $F_{z|z'}$. Hence, part 3 states that an increase of the variance of either r or δ, with the appropriate correction of the covariance, will produce more choice stochasticity.

Propositions 1 and 2 set the basis for the non-parametric identification of the model. We now study the identification of parameters θ, under the assumption that (r, δ) follows a bivariate normal distribution. Consider a DMPL dataset $\mathcal{O}$ consisting of a set of observations of the choice of a subject, or group of subjects, when presented with a risk or time menu. Assume that the dataset has M of such menus, denoted as A_m for $m = 1, \ldots, M$. The following proposition shows that two risk menus and three time menus, with properties commonly found in existing experimental datasets, are sufficient to identify θ.

Proposition 3. *Suppose that the dataset $\mathcal{O}$ contains:*

(a) *Two risk menus $\{A_{\mathcal{R},a}, A_{\mathcal{R},b}\}$ such that $K(A_{\mathcal{R},a}) \neq K(A_{\mathcal{R},b})$.*

(b) *Three time menus $\{A_{\mathcal{T},c}, A_{\mathcal{T},d}, A_{\mathcal{T},e}\}$ differing in one of three dimensions:*

(i) the delay $t_m^1 - t_m^0$ (ii) the current prize x_m^0 (iii) the future prize x_m^1.

Then, θ is identified.

Intuitively, the proof of Proposition 3 shows that we can use variation in the indifference thresholds $K(A_{\mathcal{R}})$ across risk menus to identify the parameters (μ_r, σ_r) characterizing the marginal distribution of r. Conditional on (μ_r, σ_r), we can use the variation in the delay across time menus offering the same prizes to recover $(\mu_\delta, \sigma_\delta, \rho)$. Alternatively, one can use variation in the implicit return rate across time menus (that is, variation in (x^0, x^1)) to replace variation in delays.

Under standard regularity conditions, identification of θ implies the consistency of maximum likelihood estimators of this parameter vector. This property guarantees that an analyst can recover the population value of θ with a large enough sample of observations. Nevertheless, one may be concerned about the behavior of these estimators with small experimental samples. The following result shows that the true parameters can be inferred with as few as five menus, alleviating these concerns.

Proposition 4. *For any $\theta \in \Theta$, there exist five DMPL menus that allow to infer its value exactly.*

In experimental settings, risk and time menus are usually tailored to include variation that allows researchers to infer a set of values of risk aversion and discounting under the assumption that subjects have deterministic preferences and maximize their discounted expected utility. As discussed in the next section, researchers use switches in choices across menus with different indifference thresholds to estimate an interval containing the point value of a subject's risk aversion coefficient (assuming a CRRA utility function) or her discount rate, given a value of the risk aversion coefficient. Proposition 3 shows that the same type of variation allows researchers to infer the parameters of R-DEU representation under parametric assumptions. Moreover, Proposition 4 shows that researchers can also tailor the DMPL menus of an experimental design to maximize their ability to infer a set of parameter values. We conclude the discussion with an example.[16]

Example 2. *Let* $\omega \to 0$ *and* $(\mu_r, \sigma_r, \mu_\delta, \sigma_\delta, \rho) = (0.7, 0.7, 0.05, 0.02, -0.5)$. *Consider first risk menus. The choice probability of* $([0.5, 0.5; 50, 40], 0)$ *versus* $([0.5, 0.5; 68, 25], 0)$ *is approximately* $1 - \Phi(0) = 0.5$, *and since the threshold of this problem is* 0.7, *we have* $\mu_r = 0.7$. *The choice probability of* $([0.5, 0.5; 50, 40], 0)$ *versus* $([0.5, 0.5; 95, 25], 0)$ *is* $1 - \Phi(1) = 0.16$, *and since the threshold of this problem is* 1.4, $\sigma_r = 1.4 - 0.7 = 0.7$. *Consider now time menus. As argued in the proof of Proposition 4, when* $\omega \to 0$ *the threshold map becomes the piece-wise linear map* $\min\{0, K(A_\mathcal{T})(1 - r)\}$, *which can be approximated by the linear map* $K(A_\mathcal{T})(1 - r)$ *that passes through the point* $(1, 0)$ *and has slope* $-K(A_\mathcal{T})$. *Then, consider the choice probability of* $([1; 59], 0)$ *versus* $([1; 70], 1)$ *which is approximately* 0.5, *and since the constant of this menu is approximately* $\frac{0.5}{3}$, *we have* $\mu_\delta = \frac{0.5}{3}(1 - 0.7) = 0.05$. *Now, consider the choice probability of* $([1; 70 - \epsilon], 0)$, *with* ϵ *small, versus* $([1; 70], 1)$ *that is equal to* 0.99. *This corresponds to two and a half standard deviations of the normal, and since the constant in this case is* 0, *it follows that* $\frac{\mu_\delta}{\sigma_\delta} = \Phi^{-1}(0.99) = 2.5$, *and hence* $\sigma_\delta = 0.02$. *Finally, the time menu involving* $([1; 68], 0)$ *and* $([1; 70], 1)$ *has constant* 0.029, *which is equal to the ratio of standard deviations. Hence, since the choice probability of the earlier option is approximately* 0.98 *(which corresponds to two standard deviations of the normal), it must be* $\rho = \frac{1}{2}\left[\frac{\frac{-0.3}{0.7} + 2.5}{2}\right]^2 - 1$ *which is approximately* -0.5.

4. Double Multiple Price Lists: Empirical Illustration

In this section, we illustrate the empirical application of the R-DEU model to DMPL datasets by estimating the parametric version of the model and comparing its results to those obtained employing alternative structural models previously used in the literature. For this purpose, we use data from AHLR. In this experimental study, the authors presented a sample 253 individuals representative of the Danish population with four risk tasks. Each task was comprised of up to ten risk menus. The monetary prizes of the

[16]Following the standard convention, we let $\Phi(\cdot)$ and $\phi(\cdot)$ denote the CDF and PDF of the standard normal distribution.

safe and risky alternative, $(x_1^0, x_2^0, x_1^1, x_2^1)$, varied between each task. All menus shared the same prizes within a given task but differed in the payoff probabilities (p^0, p^1). The experiment also presented each individual with six time tasks of up to 10 time menus which shared the same early prize x^0. All menus in a given time task also shared the same payoff delay k but varied in the value of the delayed prize x^1. The delay k and payoff dates (t^0, t^1) changed across tasks.[17] Following these authors, we also assume the integrated average daily wealth value ω is common across individuals and equal to 118 Danish kroner (DKK) in 2003, equivalent to approximately 30 USD that year.

We are interested in estimating risk aversion and discounting at both population and individual levels. For this reason, we restrict the analysis to a subsample from the original dataset, satisfying the following restrictions: first, we discard observations corresponding to four risk menus and six time menus containing dominated lotteries.[18] Second, we drop from the sample individuals reporting indifference between the two alternatives in some tasks. Finally, we focus on individuals whose choice switches from the safe lottery to the risky one in at least one of the four risk tasks and also switch from the early lottery to the delayed one in at least one of the six time tasks. In other words, we drop from the sample individuals who made the same choice in all the risk or time tasks.[19] These restrictions leave us with an estimation sample of 202 individuals, each facing up to 36 risk menus and up to 54 time menus.

Before discussing the methodology for estimating the structural models, we present the results from a semi-parametric estimator based on an elicitation procedure frequently used in the literature. These estimates provide a useful benchmark and will give a first picture of the degree of heterogeneity in preferences within and across individuals in the dataset.

4.1. **Semi-Parametric Estimation.** Employing multiple price lists to produce interval estimates of r and δ is common. In a given risk task, one can identify the two adjacent

[17]An example of these tasks shown in Appendix F.

[18]We do this for expositional purposes. Extending the model by adding a tremble probability to include menus with dominated alternatives is straightforward. See, for instance, Apesteguia and Ballester (2018) and Jagelka (2024).

[19]This restriction is not necessary to compute estimates at the population level. However, it is necessary to obtain comparable estimates across individuals and models. The reason is that, in all models we consider here, variation in choices is required to point-identify the parameters associated with the coefficient of risk aversion and the discount rate of an individual. If choices are the same in all menus, we can only set-identify these parameters.

menus where a subject's decision switches from the safe lottery to the risky one. The indifference threshold $K(A_{\mathcal{R}})$ of these menus provides a lower and upper bound of the interval of values of r consistent with this switch. Alternative estimators can be computed from these intervals, but it is common practice to use the midpoint of this interval as a point estimate of r. Using this procedure in the AHLR dataset results in four estimates for each subject, which we can interpret as draws from the individuals' distribution of r.[20] We can thus compute estimates of μ_r and σ_r for each individual from the average and standard deviation of the elicited draws. We can also compute population estimates of these parameters by pooling all individual draws.

Conditional on a value of r, we can follow a similar procedure to obtain draws of δ from the indifference thresholds $K(A_{\mathcal{T}}|r)$ from the adjacent menus in a time task where the choice of the individual switches from the early to the delayed lottery. We repeat this procedure across the six time tasks using each of the four draws of r obtained for this individual, obtaining 24 draws of δ for each individual. We use these draws to compute individual and population estimates of μ_δ and σ_δ as before. We also compute estimates of ρ from the sample correlation of these draws. We label these as *semi-parametric estimates* (SPE) of θ since they are obtained under parametric assumptions of the utility function but do not specify any particular distribution for (r, δ).

The last column of Table 1 shows the estimated parameters obtained using the previous procedure and pooling all individual draws. Three results are worth noting. First, the average risk aversion coefficient is 0.715, which aligns with estimates in the experimental literature and the structural estimates reported in AHLR. Similarly, the average (annual) discount rate is 13.8%, which is higher than the one estimated by AHLR but is still on the lower end of estimates obtained from experimental datasets.

[20] This procedure is similar to the one followed by Andreoni and Sprenger (2012a) to estimate the curvature of the utility function using the DMPL design. Other alternatives include choosing the point that makes the subject appear more patient or risk-averse. Without additional assumptions, there is no obvious reason to prefer one alternative over the others in these applications. Through the lens of the R-DEU model, the choice of the midpoint is consistent with the belief that r follows a uniform distribution, conditional on being on the chosen interval. Thus, one can think of the DMPL intervals as producing a piecewise approximation of the CDF of r. The accuracy of this approximation increases as the number of intervals increases and their amplitude becomes small. For the results in this section, choosing other points in the interval has limited impact on the estimated value of σ_r but can change the value of μ_r estimated for each individual. However, a change in location affects all individuals similarly and has little effect on the relationship observed between the semi-parametric estimator and the other estimators analyzed.

Second, there is a large degree of heterogeneity in preferences. The standard deviation of all the draws of r and δ in the sample is 0.833 and 0.165, respectively. Notably, a large fraction of this variation corresponds to heterogeneity within subjects. To see this, we compute the summary statistics of the estimates of (μ_r, μ_δ) at the individual level and report them in Table 2.[21] The standard deviation of μ_r across individuals is 0.608, implying that almost half the variance in r comes from variation between subjects. Similarly, the standard deviation of μ_δ across individuals is 0.111, indicating that around 44% of the variance in δ comes from variation between subjects.

Finally, there is a large and negative correlation between r and δ. The estimated correlation coefficient, -0.761, is very similar to the correlation between the individual estimates (μ_r, μ_δ), which is -0.807. As discussed in Proposition 2, higher values of r and lower values of δ generate a larger choice probability of the earlier alternative in time menus. Hence, the negative correlation between r and δ is consistent with the observed behavior in the time menus of the dataset.

4.2. **Structural Estimation.** We now turn to the structural estimation of θ using the parametric R-DEU model and, for the sake of comparison, two other structural alternatives. Our dataset contains a collection of menus $\{A_m\}_{m=1}^{M}$ and a set of N observations for $i = 1, \ldots, I$ individuals, which we denote as $\mathcal{O}$. The observation (i, m) records the choice of individual i on menu m as an indicator function $Y_{i,m}$ that takes a value of zero when the individual chooses the early/safe lottery in the menu (denoted as 0_m), and takes a value of one otherwise. To compute the population estimates, we follow the literature and assume preferences admit a representative agent so that $\mathcal{P}_\theta(0_m, A_m)$ is independent of i. Under this assumption, we can write the log-likelihood function of the data, conditional on parameter vector θ, as:

$$\log \mathcal{L}\left(\theta|\mathcal{O}\right) = \frac{1}{N}\sum_{i,m}\Big[(1 - Y_{i,m})\log \mathcal{P}_\theta(0_m, A_m) + Y_{i,m}\log(1 - \mathcal{P}_\theta(0_m, A_m))\Big].$$

We compute the maximum-likelihood estimator of θ by numerically maximizing the previous log-likelihood. This estimator is consistent and asymptotically normal under standard regularity conditions as long as θ is identified. We compute robust standard

[21]Table **??** reports the corresponding summary statistics for the individual estimates of $(\sigma_r, \sigma_\delta)$.

errors of the estimates clustered at the individual level and estimate preference parameters by subject similarly using the subsample of $\mathcal{O}$ corresponding to each individual.

All that is left is to specify $\mathcal{P}_\theta(0_m, A_m)$. Propositions 1 and 2 give the choice probabilities of the R-DEU model.[22] Finally, note that the AHLR dataset satisfies the conditions discussed in Proposition 3. It follows that the R-DEU model is identified, and our estimates are consistent and asymptotically efficient.

We also consider two alternative models previously used in the literature. The first model assumes that there is a unique underlying preference (that is, $r = \mu_r$ and $\delta = \mu_\delta$) subject to iid-additive noise, where choices are given by the following rule:

$$\mathcal{P}_\theta(0_m, A_m) = \frac{DEU_{r,\delta}(0_m)^{\frac{1}{\sigma}}}{DEU_{r,\delta}(0_m)^{\frac{1}{\sigma}} + DEU_{r,\delta}(1_m)^{\frac{1}{\sigma}}}, \tag{4.1}$$

with $DEU_{r,\delta}(0_m)$ and $DEU_{r,\delta}(1_m)$ denoting, respectively, the discounted expected utility of the early/safe lottery and the late/risky lottery in A_m, as defined in equation (2.1). This probability rule follows Luce (1959) and was introduced to the estimation of risk preferences by Holt and Laury (2002). It corresponds to the specification used in AHLR to compute population estimates of risk aversion and discounting with their data. Following these authors, we specify $u_r(x) = \frac{(x+\omega)^{1-r}}{1-r}$ and allow the noise parameter σ to differ between risk and time tasks so that the model is characterized by four parameters: $(\mu_r, \sigma_r, \mu_\delta, \sigma_\delta)$. We label this model as LUCE. It is important to emphasize that this model has several theoretical problems that complicate the interpretation of estimates. First, as shown in Apesteguia and Ballester (2018), the presence of iid-additive shocks makes $\mathcal{P}_\theta(0_m, A_m)$ a non-monotonic function of (r, δ).[23] Consequently, the model is potentially not identified since different values of these parameters may rationalize the same observed probability. Second, the functional form used in the monetary valuations generates the sort of problems discussed in Section 2.1. In particular, valuations are negative when $r > 1$, which in turn

[22] Appendix C describes the numerical method used to evaluate these probabilities efficiently by exploiting the assumption that r and δ follow a bivariate normal distribution.

[23] Notice that, as discussed in Apesteguia and Ballester (2018), the non-monotonicities are driven by the non-linearity of the utility representations; the standard use of mixed-logit models does not share these problems since they typically assume a latent utility that is linear on the parameters of interest.

generates further problems in the power expression of equation (4.1) involving imaginary numbers and leading to smaller choice probabilities of better alternatives.[24]

The second model also assumes deterministic preferences but considers instead the following specification of the probability of choosing the early/safe lottery and the late/risky lottery in menu A_m:

$$\mathcal{P}_\theta(0_m, A_m) = \Phi\left(\frac{DEU_{r,\delta}(0_m) - DEU_{r,\delta}(1_m)}{\nu_m \sigma}\right), \tag{4.2}$$

where ν_m is a menu-specific normalizing constant and σ is a noise parameter taking different values in risk and time tasks. This model is based on the "contextual error" specification proposed by Wilcox (2011) and applied empirically by Andersen et al. (2014) and Harrison et al. (2020). Following these authors, we assume $\nu_m = 1$ for menus in time tasks and set ν_m equal to the maximum utility across prizes in A_m minus the minimum utility across prizes in the same menu.[25] The model is thus characterized by the parameters $(\mu_r, \sigma_r, \mu_\delta, \sigma_\delta)$, and we label it as WILCOX.[26]

4.3. **Population Estimates.** We begin by reporting the estimated parameters at the population level in Table 1. The second, third, and fourth columns show the estimates of the R-DEU, LUCE and WILCOX models, respectively.

The estimated average risk aversion coefficient μ_r under the R-DEU model is 0.781, slightly higher than that obtained from the SPE. On the other hand, the estimated value of σ_r is 0.895, which is very similar to the one obtained using the SPE estimates. The estimated mean discount rate is 12.5%, which is also close but slightly lower than the 13.8% from the SPE. The estimated σ_δ is slightly lower but close in magnitude to the standard deviation from the SPE. Finally, the estimated correlation of -0.958 is large in magnitude and negative, consistent with the results from the SPE estimates and the

[24]Some of these theoretical problems also apply to the next alternative model and to the iid-additive RUM used in the empirical analysis of CB settings. In what follows, we will focus on the empirical comparisons with R-DEU.

[25]Unlike these authors, we use the CDF of a normal distribution instead of a logistic distribution to map the latent index to probabilities. This difference is not important for the results.

[26]The literature has also considered the use of random coefficient models, or mixed-logit models, for structurally estimating risk and time preferences (see, for instance, Andersen et al. (2008) and Andersen et al. (2014)). These models are very flexible and allow two levels of variation, one at the individual level and one at the population level. At the individual level, they are akin to the iid-additive RUM model above. At the population level, they allow variability in both r and δ across individuals. However, they share some of the theoretical problems of the two previous models, so we do not consider them here.

comparative statics of the R-DEU model discussed in the previous section. The close relationship between the semi-parametric estimates and parametric estimates from the R-DEU model illustrates the intuitive and close mapping of the parameters in the R-DEU model to the variation in choices and menus in the dataset.

Comparing the results of R-DEU model with those of LUCE and WILCOX, we can see that the population estimates of μ_r and μ_δ are very similar across models and, in the case of μ_r, we cannot reject the hypothesis that these are statistically equal. However, the estimated values of σ_r and σ_δ are quite different across models. In the case of the R-DEU model, both parameters have a direct mapping to the variance of r and δ in the SPE. On the other hand, the LUCE and WILCOX models treat these as noise parameters related to the volatility of the utility shocks. For this reason, their mapping to the data is less straightforward. Comparing the log-likelihoods of the estimated models, we can see that the R-DEU model has a slightly better fit to the data than the other two models due to a greater ability to explain choices in time menus. This is unsurprising since the R-DEU model allows for correlation between r and δ, providing an extra parameter to fit the data. Nevertheless, the differences in fit are small, and the estimated values of average risk aversion μ_r and discounting μ_δ are similar across the three structural models.

4.4. **Individual Estimates.** We now turn attention to the estimates at the individual level. Table 2 shows summary statistics of the estimated values of μ_r and μ_δ for each individual under the corresponding structural model. The last three rows of Table 2 report the Pearson correlation coefficient, the Kendall rank correlation coefficient, and the Spearman rank correlation coefficient between the individual estimates under each structural model and the corresponding SPE.

The moments of the individual estimates are very similar in the R-DEU and SPE models. In particular, the mean and standard deviation of μ_r and μ_δ across individuals are remarkably close to their corresponding population estimates. In addition, we see that all three measures of correlation are positive and very large, providing further evidence of the tight relationship between the semi-parametric and R-DEU estimates, both qualitatively and quantitatively.

In contrast, the mean and standard deviation of the individual estimates obtained using the LUCE and WILCOX models differ substantially from their SPE and population counterparts. The mean risk aversion coefficient in both models is negative, and the standard deviation is an order of magnitude larger than the population estimate. Similarly, the mean and standard deviation of the individual estimates of μ_δ presents implausible large values under the two alternative models. The Pearson correlation with the SPE in both cases is close to zero, although the rank correlation measures are higher in comparison. This suggests that the puzzling results are driven by a share of individuals for which the models estimate implausible values of μ_r and μ_δ. Looking at the quantiles of the distribution of μ_r and μ_δ under LUCE and WILCOX, we can see that the presence of a large mass of atypical values in the tails of the distribution provokes the unexpected values for the mean and standard deviation.

To understand the differences in performance across structural models, Figure 1 displays scatterplots of the estimated values of μ_r and μ_δ against the corresponding value obtained using the semi-parametric estimates. The latter provides a good benchmark of the values of the risk-aversion coefficient and the discount rate we would expect from each individual, given their choices across tasks.

The first row shows the corresponding plots for the R-DEU model. We can see that each dot in both scatterplots is close to the 45-degree line, confirming the close relationship between the R-DEU and SPE estimates observed in the summary statistics of Table 2. Table 3 and Figure 2 show that this close relationship also holds for the individual estimates of σ_r and σ_δ.

The second row in Figure 1 shows the corresponding scatterplots for the LUCE model. Three things are worth highlighting. First, there is a stark upper bound in the estimates of μ_r obtained using this model. This bound follows from the problems mentioned above when $r > 1$. This is not an issue on the population estimates, given that, for this particular dataset, the average risk aversion coefficient across individuals is below the threshold, as suggested by the SPE. However, it is a problematic restriction for individual estimates. According to the SPE, around 36% choices in risk menus are consistent with $r > 1$, and one-third of the sample subjects have $\mu_r > 1$. Second, the estimated μ_δ tracks, on average, the corresponding SPE. For low values of μ_δ, the LUCE estimate is higher than the corresponding SPE. This is a consequence of the upper bound on μ_r: suppose a subject with $r > 1$ chooses the delayed lottery over the early one in many time menus. This behavior is consistent with having large values of r, low values of δ, or both. However, the upper bound makes the model underestimate the risk aversion coefficient of this individual. Consequently, it has to over-estimate this subject's δ to rationalize her choices in time menus. Finally, note that there are several individuals for which the LUCE model estimates extremely low values of r. Similarly, the model estimates implausible large or negative discount rates for many subjects in the sample.

At first glance, this behavior is disconnected from their choices as it seems uncorrelated with their SPE. To understand the source of this erratic behavior, we distinguish two groups of subjects in each scatterplot. In the left column of Figure 1, the first group (plotted as blue circles) corresponds to 103 subjects who switched from the safe to the risky lottery in all four risk tasks. The remaining 99 subjects (shown in orange) compose the second group. These subjects did not switch in at least one of the four risk tasks. The right column of Figure 1 shows as circles the 112 subjects who switched from the early lottery to the delayed lottery in all six time tasks. Finally, we show in purple the remaining 90 subjects that did not switch in at least one of the six time tasks. We can see a clear pattern: the subjects for which the LUCE model estimates implausible values of μ_r and μ_δ are usually subjects who did not switch in at least one of the tasks. These subjects display relatively extreme preferences together with some degree of choice stochasticity, which the LUCE model is unable to capture. Importantly, these subjects are not simple outliers as they compose almost half of the sample in this experimental setting.

The corresponding results for the WILCOX model are shown in the third row of Figure 1. Compared to the LUCE model, the model does a better job capturing the heterogeneity in μ_r and μ_δ reflected in the SPE estimates. However, it also delivers implausibly large values of μ_r for a large part of the sample. The scatterplot shows that these large estimates are usually obtained for subjects who did not switch in at least one of the risk tasks. Consequently, the model also estimates values of μ_δ close to zero for many subjects with both low and large SPE. As discussed before, with large values of r, the individual is more likely to choose delayed lotteries. Suppose this individual chooses the early lottery in many of the time menus in the dataset. In that case, the model needs to compensate the large value of r with extremely low values of δ to rationalize her choices. Finally, the model also estimates implausible large values of μ_δ for subjects who did not switch in at least one of the time tasks.

The results suggest that two alternative approaches to structurally estimate risk and time preferences, the LUCE and WILCOX models, are not well suited to capture the large heterogeneity in preferences between and within subjects on DMPL datasets. In contrast, the R-DEU model offers a flexible framework with solid theoretical foundations, clear identification restrictions, and an intuitive connection with choices in DMPL data both at the population and individual levels.

5. Convex Menus: Theory

Although convex menus may seem more convoluted, the analysis can be analogously simplified by conditioning again on parameter r. This creates an intuitive, ordered structure of choices for parameter δ. Having fixed r, each $a \in [0,1)$ has an associated threshold $K(a, A_{\mathcal{C}}|r) \in (0,1]$, such that the choice is below a if and only if the value of parameter δ lies above the threshold. That is, $\Gamma(a, A_{\mathcal{C}}) = \{(r,\delta) : \delta \geq K(a, A_{\mathcal{C}}|r)\}$. In the case of convex monetary utilities, $r \leq 0$, the threshold is unique, independent of a, as only corner solutions have non-null probability. In the case of strictly concave monetary utilities, $r > 0$, the threshold $K(a, A_{\mathcal{C}}|r)$ corresponds to the unique value of δ for which the first-order condition holds for alternative a, i.e., to the value of δ for which the derivative of $DEU_{r,\delta}(a)$ with respect to a is equal to zero. The computation of the choice probabilities follows, again, from the weighted consideration of all conditional distributions $F_{\delta|r}$.

The comparative statics of shifts in parameter δ are the continuous analogous of the case of $A_{\mathcal{T}}$. To understand the case of shifts in r, we now show that whenever $r > 0$, the map $\{K(a, A_{\mathcal{C}}|r)\}_{r\in\mathbb{R}}$ is strictly increasing in r if and only if $a > \bar{e} = \frac{x^0}{x^0+x^1}$, and strictly decreasing whenever $a < \bar{e}$. The value $\bar{e}$ is no coincidence, as it describes the allocation that equalizes the two prizes, and hence the two wealths, across periods t^0 and t^1. Hence, we can show that fixing the marginal distribution of δ and shifting upwards the conditional distributions of r, we generate a larger probability of choice for any neighborhood of $\bar{e}$, i.e., choices become more *balanced.* This comparative statics exercise neatly reflects the role of $r > 0$ in $A_{\mathcal{C}}$ as inter-temporal substitution.

The comparative statics of spreads of the parameters are similar to the case of $A_{\mathcal{T}}$, simply accounting for the continuity of the choice variable. In the binary case of $A_{\mathcal{T}}$ the trade-off between earlier versus future prizes does necessarily involve the choice of alternative $0_{\mathcal{T}}$ versus alternative $1_{\mathcal{T}}$. In the current continuous case, this trade-off has alternative $\bar{e}$ as the critical value. Alternatives below (resp., above) $\bar{e}$ allocate a larger potential prize to the earlier period (resp., later period). We now show that, keeping constant the marginal distribution of one parameter, an expansion of the conditional distributions of the other parameter always brings the cumulative choice probability $\mathcal{P}_f([0, \bar{e}], A_{\mathcal{C}})$ closer to 1/2. That is, the probabilities of choices below and above $\bar{e}$ become closer, implying that behavior is now more stochastic.

Proposition 5. *For every pair of R-DEUs, f and g, and every menu $A_{\mathcal{C}}$:*

(1) $\mathcal{P}_f([0, a], A_{\mathcal{C}}) = 1 - \int_r F_{\delta|r}(K(a, A_{\mathcal{C}}|r))f^r(r)\mathrm{d}r$.

(2) (a) *If $F^r = G^r$, and for all r $G_{\delta|r}$ dominates $F_{\delta|r}$, $\mathcal{P}_f([0, a], A_{\mathcal{C}}) \geq \mathcal{P}_g([0, a], A_{\mathcal{C}})$ for every $a \in [0, 1)$.*

(b) *If $F^\delta = G^\delta$, and for all δ $F_{r|\delta}$ dominates $G_{r|\delta}$, $\mathcal{P}_f([0, \bar{a}], A_{\mathcal{C}}) - \mathcal{P}_f([0, \underline{a}], A_{\mathcal{C}}) \geq \mathcal{P}_g([0, \bar{a}], A_{\mathcal{C}}) - \mathcal{P}_g([0, \underline{a}], A_{\mathcal{C}})$ for every $0 < \underline{a} < \bar{e} < \bar{a} < 1$.*

(3) (a) *If $F^r = G^r$, and for all r $F_{\delta|r}$ expands $G_{\delta|r}$ with $K(\bar{e}, A_{\mathcal{C}}|r) \neq$ med $F(\delta|r)$, $|\mathcal{P}_f([0, \bar{e}], A_{\mathcal{C}}) - \frac{1}{2}| < |\mathcal{P}_g([0, \bar{e}], A_{\mathcal{C}}) - \frac{1}{2}|$.*

(b) *If $F^\delta = G^\delta$, and for all δ $F_{r|\delta}$ expands $G_{r|\delta}$ with $K(\bar{e}, A_{\mathcal{C}}|\delta) \neq$ med $F(r|\delta)$, $|\mathcal{P}_f([0, \bar{e}], A_{\mathcal{C}}) - \frac{1}{2}| < |\mathcal{P}_g([0, \bar{e}], A_{\mathcal{C}}) - \frac{1}{2}|$.*

5.1. **Implications for the Parametric Version.** The implications of Proposition 5 for the parametric case of the CRRA and the bivariate normal are in line with the general discussion. With convex utilities, the unique relevant threshold for δ separates the choice of $a = 0$ and $a = 1$ and it corresponds to $K(A_{\mathcal{C}}|r) = \frac{1}{t^1 - t^0} \log \frac{p^1}{p^0} + K(A_{\mathcal{T}}|r) = \frac{1}{t^1 - t^0} \log \frac{p^1}{p^0} + \frac{1}{t^1 - t^0} \log \left[\frac{(\omega + x^1)^{1-r} - \omega^{1-r}}{(\omega + x^0)^{1-r} - \omega^{1-r}} \right]$. In the concave part, the threshold for δ determining a choice below a can be obtained from the first-order condition, and corresponds to

$K(A_{\mathcal{C}}|r) = \frac{1}{t^1 - t^0} \log \frac{p^1}{p^0} + \frac{1}{t^1 - t^0} \log \frac{x^1}{x^0} + \frac{1}{t^1 - t^0} r \log \left[\frac{(1-a)x^0 + \omega}{ax^1 + \omega}\right]$. There are three terms in the expression; the first term is the same than the first term of the convex case and depends on the probabilities, the second term is the limit when $r \to 0$ of the second term of the convex case, and the third term is unique to the concave case. The latter one shows that solutions are a linear function of r. Moreover, solutions are in general interior, and whenever ω tends to 0, they are always interior. Since the choice probabilities are again built on the basis of the conditional probabilities that are normally distributed, computation is routine. Ceteris paribus, an increase in the median of δ generates larger choice probabilities for alternatives allocating more resources to the earlier period. Given the convex nature of the menu, increasing the median of r has mostly a smoothing effect, equalizing the prizes across the two time periods. As before, increasing either the variance of δ or of r leads to more choice stochasticity.

As in the case of DMPLs, Proposition 5 sets the basis for the identification of the model. We now study the identification of θ under parametric assumptions. Consider a convex budget dataset $\mathcal{O}$ consisting of a set of observations of the tokens allocated by an individual, or group of individuals, when presented with a set of convex menus $A_{\mathcal{C},m}$ indexed by $m = 1, \ldots, M$. The following result shows that variation in pay-off delay and variation in either the payoff probability or the return rate implicit across convex menus is sufficient to identify θ.

Proposition 6. *Suppose that the dataset $\mathcal{O}$ contains five convex menus with relatively large payoffs, such that $\omega/x^1 \to 0$ and $\omega/x^0 \to 0$, satisfying the following conditions:*

(a) *Two of the menus $\{A_{\mathcal{C},a}, A_{\mathcal{C},b}\}$ are such that (i) $t^1_a - t^0_a = t^1_b - t^0_b$ and (ii) $p^1_a/p^0_a \neq p^1_b/p^0_b$ or $x^1_a/x^0_a \neq x^1_b/x^0_b$.*

(b) *The three remaining menus $\{A_{\mathcal{C},c}, A_{\mathcal{C},d}, A_{\mathcal{C},e}\}$ differ only in one of three dimensions: (i) the delay $t^1_m - t^0_m$ (ii) the current prize x^0_m (iii) the future prize x^1_m.*

Then, θ is identified.

Intuitively, the proof of Proposition 6 shows that one can use two moments of the data to identify the distribution of r: the share of interior choices and the elasticity of the response of token allocations to either payoff probabilities or return rates. Using these moments to identify (μ_r, σ_r), one can then focus on identifying $(\mu_\delta, \sigma_r, \rho)$ from the corner solutions in the data. Specifically, when $r < 0$, the problem of the subject is analogous to the discrete choice in time menus studied in DMPL settings. We can thus use the same conditions used to identify θ_δ in Proposition 3 to identify these parameters from the predicted behavior at corner allocations.

As in the case of DMPL lotteries, it is also possible to infer θ using a small number of convex menus.

Proposition 7. *For any $\theta \in \Theta$, there exist five convex menus that allow to infer its value exactly.*

Example 3. *Consider again the case where $\omega \to 0$ and parameters $(\mu_r, \sigma_r, \mu_\delta, \sigma_\delta, \rho) = (0.7, 0.7, 0.05, 0.02, -0.5)$. Take first the convex problem defined by probabilities $p^0 = 1$ and $p^1 = 0.8$, payouts $x^0 = 15$ and $x^1 = 20$, and timings $t^0 = 0$ and $t^1 = 0 + \epsilon$, with ϵ small. The choice probability of $a = 1$ corresponds to one negative standard deviation, and hence $\frac{\mu_r}{\sigma_r} = 1$. The risk aversion level above which the choice is below $a = 0.48$ is 0.7. Since the cumulative choice probability at $a = 0.48$ is 0.5, we learn that $\mu_r = 0.7$, and from the above expression, $\sigma_r = 0.7$. We can now consider the convex version of the time problems described in Example 1 by fixing $p^1 = p^0$, and reproduce the analysis there with a hypothetical discrete choice problem in which we aggregate all observed probabilities of options below $1/2$ and options above $1/2$.*

6. Convex Menus: Empirical Illustration

We now illustrate the empirical application of the R-DEU model to convex budgets using data from the experimental design in AS. In this study, the authors present 80 subjects with 84 convex menus. In each menu, the subject receives 100 tokens and decides how many to allocate between two dates: t^0 and t^1. Every token allocated in t^0 is transformed into dollars at a rate q^0 so that $x^0 = 100q^0$. Similarly, every token allocated in t^1 is exchanged into dollars at a rate q^1 so that $x^1 = 100q^1$. The prizes x^0 and x^1 are rewarded with probabilities p^0 and p^1, respectively. Otherwise, the subject received nothing. All menus fixed t^0 to 7 days, and q^0 to 0.20 USD per token, while varying the remaining menu characteristics. Consequently, the empirical design satisfies the conditions for identification of the R-DEU model discussed in Proposition 6.

The dataset records the share of tokens $a \in [0,1]$ allocated in t^1 by each subject i when presented with each menu in $\{A_m\}_{m=1}^{M}$. Since tokens are not divisible, the experimental implementation discretizes the choice set in S equidistant options $\alpha^1 = [a_1, a_2]$, $\alpha^2 = [a_2, a_3]$, ..., $\alpha^S = [a_S, a_{S+1}]$, with $a_1 = 0$ and $a_{S+1} = 1$. In the data, 93% of the choices correspond to token allocations that are a multiple of 5. Consequently, we set $S = 21$, so that $a_2 = 0.025, a_3 = 0.075, \ldots a_S = 0.975$. As a result, the dataset $\mathcal{O}$ contains a collection of $M = 84$ convex menus faced by $I = 80$ individuals, for a total of $N = 6720$ observations. The observation (i, m) records the choice of individual i on menu m as an indicator function $Y_{i,m}(s)$ taking a value of one when the token allocation is contained in α^s, and zero otherwise. In what follows, we set $\omega = 5$ USD, consistent with the participation payment in AS.

6.1. **Structural Estimation.** To estimate the R-DEU model, we use our parametric restrictions and follow a representative agent approach where the probability that $a_m \in \alpha^s$, denoted as $\mathcal{P}_\theta([a_s, a_{s+1}], A_m)$, is independent of i and given by Proposition 5. We can thus write the log-likelihood function of the data, conditional on parameter vector θ, as:

$$\log \mathcal{L}\left(\theta|\mathcal{O}\right) = \frac{1}{N} \sum_{i,m} \sum_{s} \Big[Y_{i,m}(s) \log \mathcal{P}_\theta([a_s, a_{s+1}], A_m) \Big].$$

As before, maximization of the previous log-likelihood yields a consistent and asymptotically normal estimator of θ under standard regularity conditions. We compute robust standard errors of the estimates clustered at the individual level and compare the estimates of the R-DEU model with two alternative methods.

The first alternative method follows AS in estimating r and δ from the first order condition associated with the convex budget problem using non-linear least squares (NLS) to minimize the distance between predicted and observed allocation of tokens. This method leads to the estimation of two preference parameters, μ_r and μ_δ, without an explicit account for their heterogeneity.[27]

[27]Notice that the NLS method imposes larger penalties to larger deviations from the first-order conditions, which may be read as larger penalties to larger deviations from the mean values of the

The second alternative method employs an iid-additive RUM to estimate risk and time preferences in convex budgets (see Harrison et al. (2013) and Cheung (2015)). In this model, the probability of choosing alternative a in menu A_m is given by:

$$\mathcal{P}_\theta([a_s, a_{s+1}], A_m) = \frac{e^{DEU(\overline{\alpha}^s)}}{e^{DEU(0)} + e^{DEU(\overline{\alpha}^2)} + \ldots + e^{DEU(1)}},$$

with $\overline{\alpha}^1 = 0$, $\overline{\alpha}^S = 1$, and $\overline{\alpha}^s = (a_s + a_{s+1})/2$ for $s = 2, \ldots, S-1$. As in the NLS approach, this model assumes preferences are deterministic so that μ_r and μ_δ are the only two parameters to be estimated. However, choice in this model is stochastic and thus potentially consistent with the large heterogeneity in allocations observed in this type of data.

Table 4 presents the estimated parameters under each model. Regarding risk aversion, the R-DEU model estimates $\mu_r = 0.207$ and $\sigma_r = 0.752$. Note that these are lower than estimates from DMPL designs, probably due to the fact that here the curvature r represents both, risk aversion and intertemporal substitution. However, it is positive and statistically different from zero. This contrasts with the estimates from the iid-additive RUM. The reason for these differences is simple: around 48% of the observations in the dataset correspond to extreme allocations $a = 0$ or $a = 1$. To explain the large presence of corner solutions, the iid-additive RUM requires estimating convex utility functions. The additional flexibility of the R-DEU model allows it to match the large fraction of corner solutions with a slightly concave utility function.[28]

As for the distribution of δ, we estimate an average annual discount rate of approximately 34%. This estimate is close to the 26% estimated using NLS and is almost half the annual rate estimated using the iid-additive RUM. This difference may be explained by the larger concavity of utility estimated in the R-DEU model, which is a substitute for a larger discount factor to explain choices between $a = 0$ and $a = 1$ at the corners. Finally, we

underlying parameters. R-DEU formalizes this principle in terms of behavioral variation, allowing to produce explicit probabilistic predictions.

[28]One advantage of having the whole distribution of r is that we can estimate additional moments of interest and compute their corresponding standard errors using the delta method. For example, one could be interested in $E[r|r > 0]$, which provides information about the average curvature of the utility function inferred from interior allocations. Given our estimates and parametric assumptions, we estimate the value of this moment to be 0.68, with a standard error of 0.067.

estimate a large variability in the discount factor, with a standard deviation of 1.8 and a negative correlation coefficient of -0.16, lower than the one obtained using DMPL data.

The difference in population estimates and log-likelihoods present an incomplete picture of the differences across models. Figure 3 shows the distribution of a across all observations in the data and compares it with the corresponding distribution of choice predicted by the model.

The first thing to note is that the fit of the R-DEU model in the full sample is quite good. The model does a good job matching both the share of corner solutions and the presence of interior choices distributed around $a = 0.5$. The iid-additive RUM, on the other hand, misses both a large share of the corner allocations and the share of interior choices around $a = 0.5$. Finally, since the NLS model does not specify how choice stochasticity emerges, we cannot provide an explicit account for the choice heterogeneity in a given menu. Instead, we can show the choices given by the estimated parameters, both at the menu level and across menus. The second column of Figure 3 shows the observed and predicted frequency of each share choice for a single menu in the dataset, for the three models under consideration.[29] It can be seen that the R-DEU model matches the choice patterns observed in convex budgets both in the full sample and for particular menus.

The R-DEU model does a good job explaining the overall patterns of choice frequency across all menus in the dataset. This is not to say that the R-DEU model is thus able to rationalize any data. The model inherits many of the weaknesses of the assumptions of expected utility and exponential discounting. One example is the common ratio property discussed in AS. Figure 4 shows the predicted choice distribution across tasks sharing the same payoff probabilities. We can see that the distribution is identical across menus with the same ratio p^0/p^1, which is inconsistent with the observed choice patterns in the data. Nevertheless, the tools we introduce can be used to extend the model to account for these and other behavioral considerations. In Appendix D, we illustrate by estimating a small extension of the model assuming $\beta - \delta$ preferences as in AS. We leave a theoretical analysis of this and other extensions for future research.

[29] Figure 3 reports on menu $(x_0, x_1, p_0, p_1, t_0, t_1) = (20, 20, 0.4, 0.5, 7, 45)$. Analogous conclusions are obtained for any one of the 84 menus.

7. Final Remarks

In this paper we have studied preference heterogeneity in the context of the most standard treatment of risk and time preferences, and we have proposed and studied the random discounted expected utility model. By using the ordered structure that links parameters and choice, we have shown that the model is computationally convenient, and well founded in terms of comparative statics. In addition, we have applied the model to two very different datasets, and shown that the model accounts behavior remarkably well in both cases. We believe that this is a promising approach to the treatment of heterogeneity when multiple parameters are involved, such as in the study of social preferences, ambiguity, limited attention, and other relevant behavioral considerations.

References

[1] Aguiar, Victor H. and Nail Kashaev (2021). "Stochastic Revealed Preferences with Measurement Error." *Review of Economic Studies* 88(4):2042–2093. 10.1093/restud/rdaa067

[2] Ahlbrecht, Martin and Martin Weber (1997). "An Empirical Study on Intertemporal Decision Making under Risk." *Management Science*, 43(6):813–826. 10.1287/mnsc.43.6.813

[3] Alan, Sule and Seda Ertac (2018). "Fostering Patience in the Classroom: Results from Randomized Educational Intervention." *Journal of Political Economy*, 126(5):1865–1911. 10.1086/699007

[4] Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and E. Elisabet Rutstrom (2008). "Eliciting Risk and Time Preferences." *Econometrica* 76(3):583–618. 10.1111/j.1468-0262.2008.00848.x

[5] Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and E. Elisabet Rutstrom (2014). "Discounting Behavior: A Reconsideration." *European Economic Review* 71:15–33. 10.1016/j.euroecorev.2014.06.009

[6] Andreoni, James and Charles Sprenger (2012a). "Estimating Time Preferences from Convex Budgets." *American Economic Review*, 102(7):3333–3356. 10.1257/aer.102.7.3333

[7] Andreoni, James and Charles Sprenger (2012b). "Risk Preferences Are Not Time Preferences." *American Economic Review*, 102(7):3357–3376. 10.1257/aer.102.7.3357

[8] Apesteguia, Jose and Miguel A. Ballester (2018). "Monotone Stochastic Choice Models: The Case of Risk and Time Preferences." *Journal of Political Economy*, 126(1):74–106. 10.1086/695504

[9] Augenblick, Ned, Muriel Niederle, and Charles Sprenger (2015). "Working over Time: Dynamic Inconsistency in Real Effort Tasks." *Quarterly Journal of Economics*, 130(3):1067–1115. 10.1093/qje/qjv020

[10] Barseghyan, Levon, Francesca Molinari, and Matthew Thirkettle (2021). "Discrete Choice under Risk with Limited Consideration," *American Economic Review*, 111(6):1972–2006. 10.1257/aer.20190253

[11] Baucells, Manel and Franz H. Heukamp (2012). "Probability and Time Tradeoff." *Management Science*, 58(4):831–842. 10.1287/mnsc.1110.1450

[12] Benjamin, Daniel J., Sebastian A. Brown, and Jesse M. Shapiro (2013). "Who Is 'Behavioral'? Cognitive Ability and Anomalous Preferences." *Journal of the European Economic Association*. 11(6):1231–1255. 10.1111/jeea.12055

[13] Burks, Stephen V., Jeffrey P. Carpenter, Lorenz Goette, and Aldo Rustichini (2009). "Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment." *Proceedings of the National Academy of Sciences*, 106(19):7745–7750. 10.1073/pnas.0812360106

[14] Carvalho, Leandro S., Stephan Meier, and Stephanie W. Wang (2016). "Poverty and Economic Decision-Making: Evidence from Changes in Financial Resources at Payday." *American Economic Review*, 106(2):260–284. 10.1257/aer.20140481

[15] Cattaneo, Matias D., Xiaoxia Ma, Yusufcan Masatlioglu, and Elchin Suleymanov (2020). "A Random Attention Model," *Journal of Political Economy*, 128(7):2796–2836. 10.1086/706861

[16] Cheung, Stephen L. (2015). "Risk Preferences Are Not Time Preferences: On the Elicitation of Time Preference under Conditions of Risk: Comment." *American Economic Review*, 105(7):2242–2260. 10.1257/aer.20120946

[17] Choi, Syngjoo, Raymond Fisman, Douglas Gale, and Shachar Kariv (2007). "Consistency and Heterogeneity of Individual Behavior under Uncertainty," *American Economic Review*, 97(5):1921–1938. 10.1257/aer.97.5.1921

[18] Coble, Keith H. and Jayson L. Lusk (2010). "At the Nexus of Risk and Time Preferences: An Experimental Investigation." *Journal of Risk and Uncertainty*, 41(1):67–79. 10.1007/s11166-010-9096-7

[19] Dardanoni, Valentino, Paola Manzini, Marco Mariotti, and Christopher Tyson (2020). "Inferring Cognitive Heterogeneity from Aggregate Choices." *Econometrica*, 88(3):1269–1296. 10.3982/ECTA16382

[20] DellaVigna, Stefano (2018). "Structural Behavioral Economics." *Handbook of Behavioral Economics-Foundations and Applications 1*, B.D. Bernheim, S. DellaVigna, and D. Laibson (eds.), Elsevier, 613–723. 10.1016/bs.hesbe.2018.07.005

[21] Dohmen, Thomas, Armin Falk, David Huffman, and Uwe Sunde (2010). "Are Risk Aversion and Impatience Related to Cognitive Ability?" *American Economic Review*, 100:1238–1260. 10.1257/aer.100.3.1238

[22] Falk, Armin, Anke Becker, Thomas Dohmen, Benjamin Enke, David Huffman, and Uwe Sunde (2018). "Global Evidence on Economic Preferences." *Quarterly Journal of Economics*, 133(4):1645–1692. *American Economic Review*, 100:1238–1260. 10.1093/qje/qjy013

[23] Fisman, Raymond, Shachar Kariv, and Daniel Markovits (2007). "Individual Preferences for Giving." *American Economic Review*, 97(5):1858–1876. 10.1257/aer.97.5.1858

[24] Harrison, Glenn W., Morten I. Lau, and E. Elisabet Rutstrom (2013). "Identifying Time Preferences with Experiments: Comment." *Center for the Economic Analysis of Risk Working Paper.*

[25] Harrison, Glenn W., Morten I. Lau, and Hong Il Yoo (2020). "Risk Attitudes, Sample Selection and Attrition in a Longitudinal Field Experiment." *The Review of Economics and Statistics*, 102(3):552–568. 10.1162/rest_a_00845

[26] Holt, Charles A. and Susan K. Laury (2002). "Risk Aversion and Incentive Effects." *American Economic Review*, 92(5):1644–1655. 10.1257/000282802762024700

[27] Jagelka, Tomáš (2024). "Are Economists' Preferences Psychologists' Personality Traits? A Structural Approach." *Journal of Political Economy*, 132(3):910–970. 10.1086/726908

[28] Kim, Hyuncheol Bryant, Syngjoo Choi, Booyuel Kim, and Cristian Pop-Eleches (2018). "The Role of Education Interventions in Improving Economic Rationality." *Science*, 362:83–86. 10.1126/science.aar6987

[29] Kreps, David M. (2013). "Microeconomic Foundations (Vol. 1)." *Princeton University Press.*

[30] Lau, Morten I. and Hyeongwoo Yoo (2023). "A General Approach to Estimating Random Preference Models of Risk Attitudes." *Mimeo.*

[31] Mas-Colell, Andreu, Michael D. Whinston, and Jerry R. Green (1995). "Microeconomic Theory." *Oxford University Press.* DOI: 10.1093/0195073401.001.0001

[32] Pratt, John W. (1964). "Risk Aversion in the Small and in the Large." *Econometrica*, 32(1/2):122–136. 10.2307/1913738

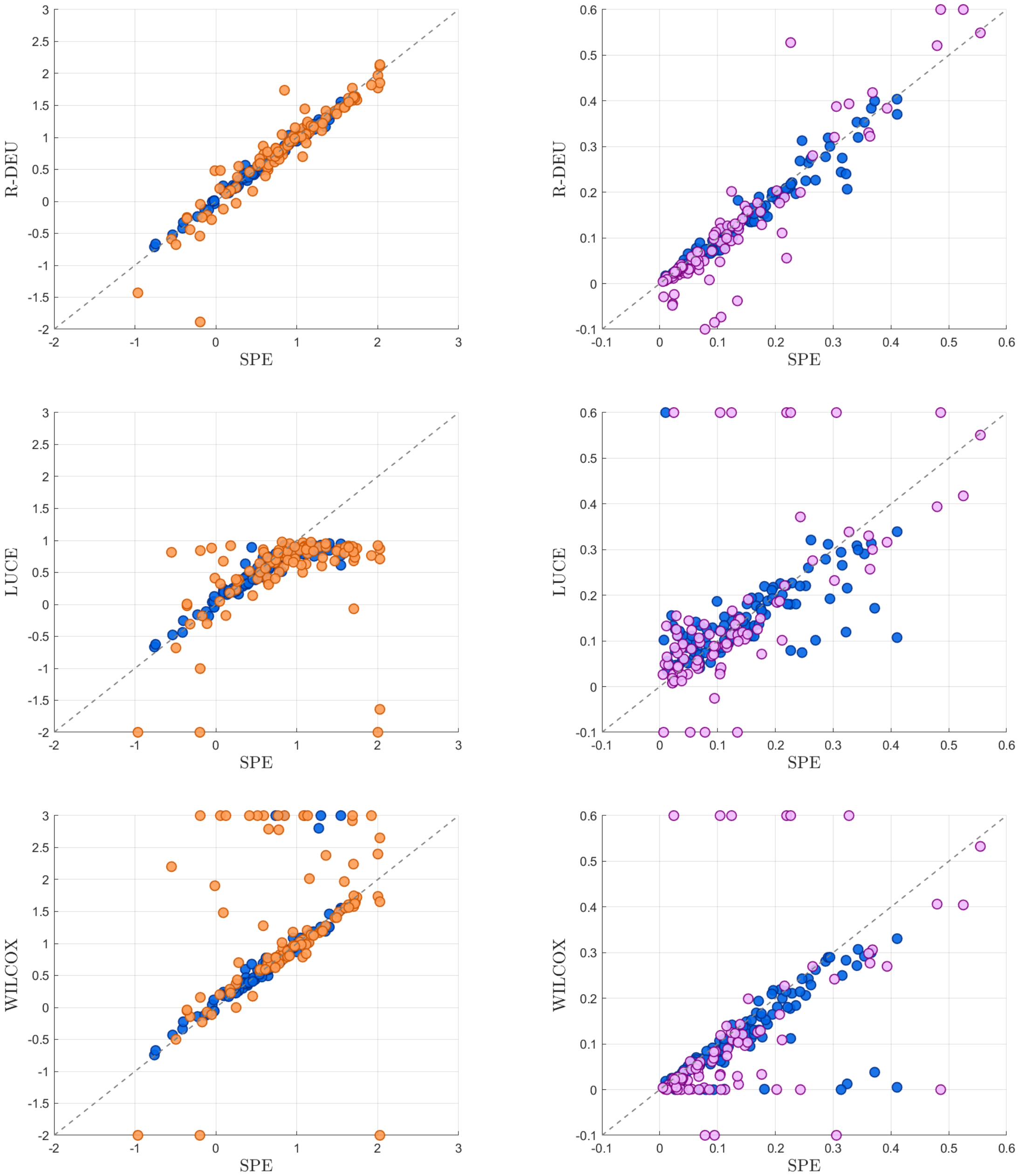
[33] Quiggin, John (1982). "A Theory of Anticipated Utility." *Journal of Economic Behavior & Organization*, 3(4):323–343. 10.1016/0167-2681(82)90008-7

[34] Tanaka, Tomomi, Colin F. Camerer, and Quang Nguyen (2010). "Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam." *American Economic Review*, 100(1):557–571. 10.1257/aer.100.1.557

[35] Wilcox, Nathaniel T. (2011). "Stochastically More Risk Averse: A Contextual Theory of Stochastic Discrete Choice under Risk." *Journal of Econometrics*, 162(1):89–104. 10.1016/j.jeconom.2009.10.012

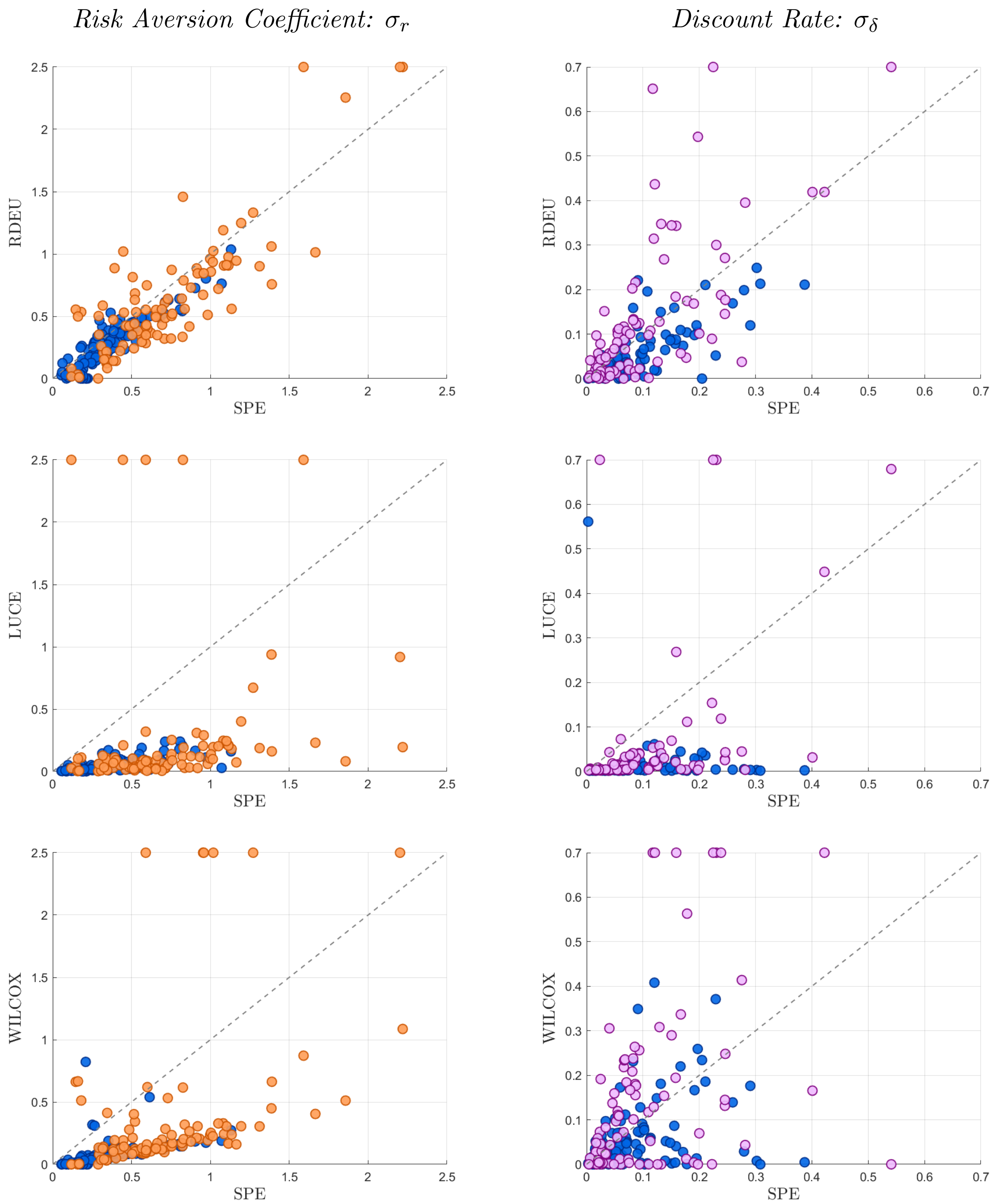
FIGURE 1. Risk and Time Preferences at the Individual Level AHLR

Risk Aversion Coefficient: μ_r *Discount Rate:* μ_δ



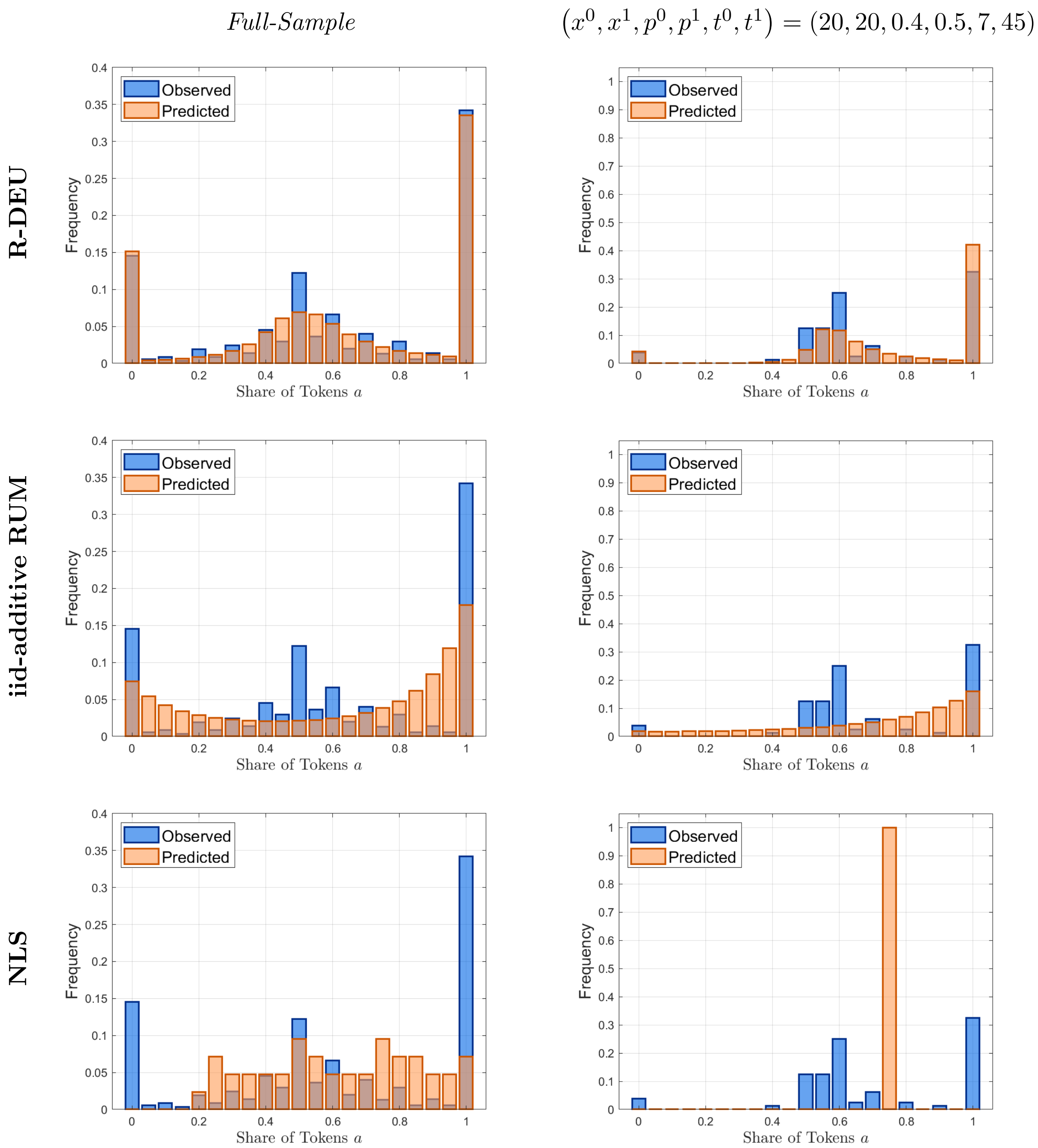
NOTES. The figure shows the estimated average risk aversion coefficient μ_r (first column) and annual discount rate μ_δ (second column) for each individual in the double-multiple price list data from Andersen et al. (2008). Each dot shows a subject's estimate using the corresponding structural model and compares it against the semi-parametric estimate based on the adjacent menus in each risk/time task where the subject's choice switched from the safe/early lottery to the risky/delayed lottery. Subjects who did not switch choices in at least one of the four risk tasks are shown in orange. Subjects who did not switch choices in at least one of the six time tasks are shown in purple. Estimates are truncated to fit the ranges in the plots.

Figure 2. Volatility of Risk and Time Preferences at the Individual Level: AHLR



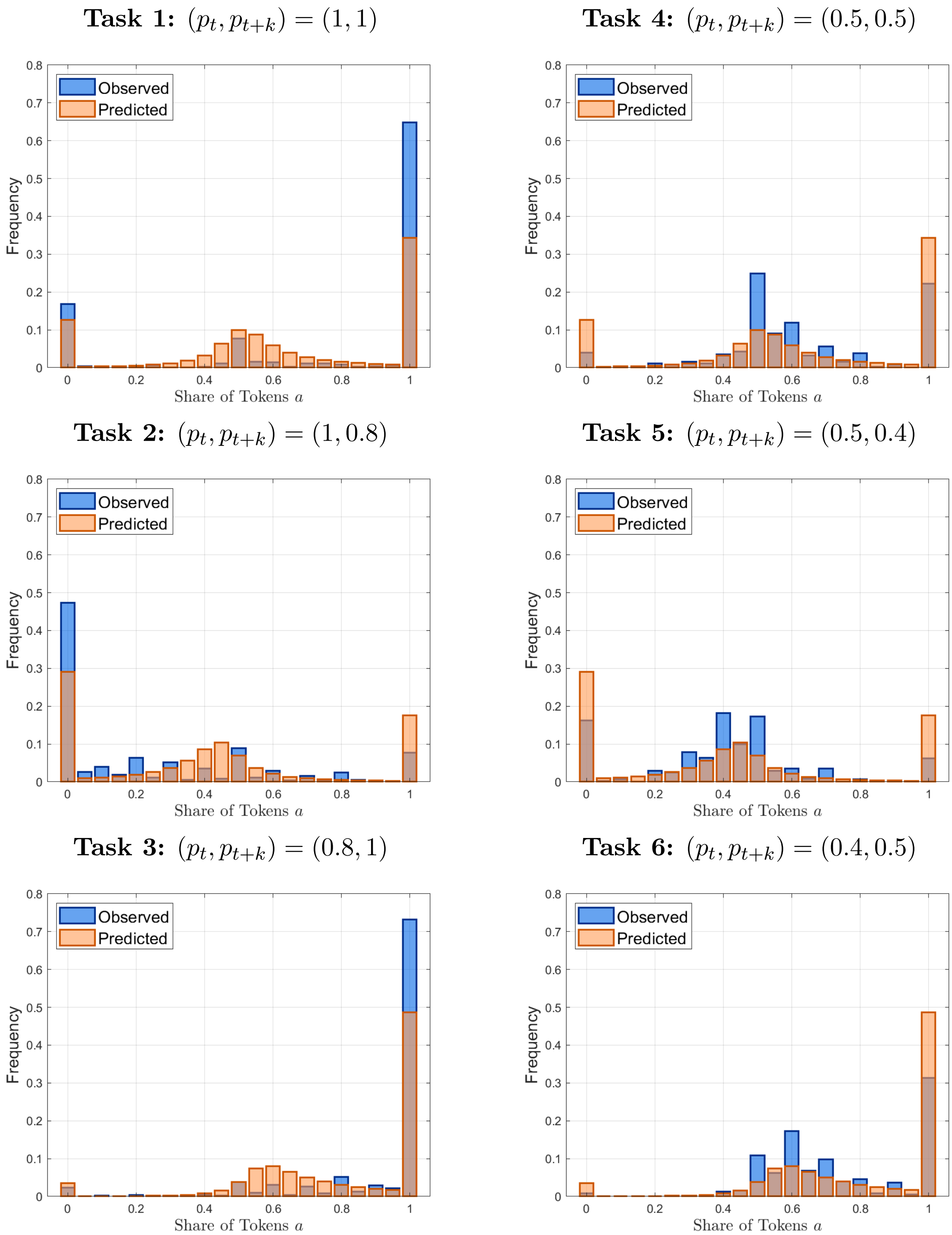
NOTES. The figure shows the estimated standard deviation of the risk aversion coefficient σ_r (first column) and annual discount rate σ_δ (second column) for each individual in the double-multiple price list data from Andersen et al. (2008). Each dot shows a subject's estimate using the corresponding structural model and compares it against the semi-parametric estimate based on the adjacent menus in each risk/time task where the subject's choice switched from the safe/early lottery to the risky/delayed lottery. Subjects who did not switch choices in at least one of the four risk tasks are shown in orange. Subjects who did not switch choices in at least one of the six time tasks are shown in purple. Estimates are truncated to fit the ranges in the plots.

FIGURE 3. Predicted and Observed Choice: AS



NOTES. The figure shows the observed and predicted frequency of choosing each token share a in the convex menu dataset of Andreoni and Sprenger (2012b). The blue bars show the frequency observed in the data. The orange bars show the frequency predicted by the corresponding model using the population estimates reported in Table 3. The left column shows the results for the entire sample, while the right column shows the results for a convex menu with payoffs of 20 USD delivered in 7 and 35 days with probability 0.4 and 0.5, respectively.

FIGURE 4. Predicted and Observed Choice By Risk Condition: AS



NOTES. The figure shows the observed and predicted frequency of choosing each token share a in the convex menu dataset of Andreoni and Sprenger (2012b). The blue bars show the frequency observed in the data. The orange bars show the frequency predicted by the R-DEU model using the population estimates reported in Table 3.

TABLE 1. Aggregate Risk and Time Preferences: AHLR

	R-DEU	**LUCE**	**WILCOX**	**SPE**
μ_r	0.781	0.726	0.777	0.715
	[0.053]	[0.058]	[0.045]	–
σ_r	0.895	0.086	0.220	0.833
	[0.049]	[0.016]	[0.011]	–
μ_δ	0.125	0.101	0.095	0.138
	[0.008]	[0.009]	[0.007]	–
σ_δ	0.125	0.020	0.222	0.165
	[0.010]	[0.006]	[0.017]	–
ρ	−0.958	–	–	−0.761
	[0.016]	–	–	–
Log-Like $\mathcal{L}$				
Risk Menus	−0.450	−0.448	−0.445	–
Time Menus	−0.495	−0.554	−0.557	–
All Menus	−0.481	−0.521	−0.522	–

NOTES.- The table reports estimated risk aversion and discount rates at the aggregate level using different models and data from the double multiple price list design in Andersen et al. (2008). The first three columns show the maximum likelihood estimates from the three structural models described in the main text. The last column shows the population mean and standard deviation of the semi-parametric estimates of r and δ obtained from the adjacent menus in each risk/time task where the choice of each individual in the sample switched from the safe/early lottery to the risky/delayed lottery. Standard errors for each MLE are shown in brackets and are clustered at the individual level.

Table 2. Individual Risk and Time Preferences: AHLR

	Risk Aversion Coefficient μ_r				Discount Rate μ_δ			
Moment	**R-DEU**	**LUCE**	**WILCOX**	**SPE**	**R-DEU**	**LUCE**	**WILCOX**	**SPE**
Mean	0.698	−4.875	−0.305	0.715	0.133	1.525	2.634	0.138
Std. Dev.	0.629	45.04	19.59	0.608	0.167	34.412	20.31	0.111
Min	−1.886	−458.4	−276.7	−0.964	−0.443	−315.6	−34.03	0.006
10th pctl.	−0.122	−0.052	−0.054	−0.064	0.018	0.040	0.001	0.027
25th pctl.	0.366	0.333	0.389	0.285	0.042	0.074	0.024	0.055
Median	0.708	0.656	0.782	0.714	0.098	0.115	0.067	0.110
75th pctl.	1.125	0.842	1.271	1.128	0.182	0.184	0.152	0.194
90th pctl.	1.507	0.894	2.782	1.557	0.315	0.303	0.273	0.308
Max	2.139	0.977	8.123	2.026	1.658	344.4	153.9	0.555
Correlation with SPE								
Pearson' r	0.961	0.079	−0.111	1	0.814	0.029	0.023	1
Kendall's τ	0.899	0.597	0.690	1	0.836	0.544	0.591	1
Spearman's ρ	0.980	0.728	0.756	1	0.943	0.702	0.686	1

NOTES.- The table reports summary statistics of the estimated average risk aversion (μ_r) and discount rates (μ_δ) across individuals using data from the double multiple price list design from Andersen et al. (2008). Each column corresponds to a model described in the main text. The last three rows report, respectively, the Pearson correlation coefficient, the Kendall rank correlation coefficient, and the Spearman rank correlation coefficient between subjects' estimates using a structural model and the semi-parametric estimates obtained from the adjacent menus in each risk/time task where the choice of the individual switched from the safe/early lottery to the risky/delayed lottery.

Table 3. Volatility of Individual Risk and Time Preferences: AHLR

	Risk Aversion Coefficient σ_r				Discount Rate σ_δ			
Moment	**R-DEU**	**LUCE**	**WILCOX**	**SPE**	**R-DEU**	**LUCE**	**WILCOX**	**SPE**
Mean	0.473	19.56	5021	0.539	0.099	0.510	19802	0.091
Std. Dev.	0.568	208.1	70187	0.374	0.247	6.373	139319	0.086
Min	0.001	0.001	0.001	0.056	0.001	0.002	0.001	0.003
10th pctl.	0.050	0.006	0.002	0.172	0.005	0.002	0.001	0.017
25th pctl.	0.190	0.021	0.055	0.273	0.015	0.003	0.005	0.029
Median	0.362	0.051	0.104	0.446	0.042	0.004	0.041	0.061
75th pctl.	0.554	0.112	0.189	0.710	0.107	0.017	0.120	0.120
90th pctl.	0.903	0.211	0.407	1.028	0.211	0.042	0.241	0.207
Max	5.963	2895	1e6	2.219	3.127	90.59	1e6	0.541
Correlation with SPE								
Pearson' r	0.805	−0.055	0.014	1	0.591	0.126	0.136	1
Kendall's τ	0.681	0.471	0.635	1	0.520	0.366	0.310	1
Spearman's ρ	0.850	0.644	0.778	1	0.702	0.504	0.421	1

NOTES.- The table reports summary statistics of the estimated standard deviation of risk aversion (σ_r) and discount rates (σ_δ) across individuals using data from the double multiple price list design from Andersen et al. (2008). Each column corresponds to a model described in the main text. The last three rows report, respectively, the Pearson correlation coefficient, the Kendall rank correlation coefficient, and the Spearman rank correlation coefficient between subjects' estimates using a structural model and the semi-parametric estimates obtained from the adjacent menus in each risk/time task where the choice of the individual switched from the safe/early lottery to the risky/delayed lottery.

TABLE 4. Aggregate Risk and Time Preferences: AS

	R-DEU	**iid-additive RUM**	**NLS**
μ_r	0.207	−0.133	0.317
	[0.062]	[0.020]	[0.017]
σ_r	0.752	–	–
	[0.079]	–	–
μ_δ	0.339	0.571	0.262
	[0.108]	[0.081]	[0.079]
σ_δ	1.805	–	–
	[0.124]	–	–
ρ	−0.164	–	–
	[0.053]	–	–
$\mathcal{L}$	−2.108	−2.519	–

NOTES.- The table reports the maximum-likelihood estimates of risk aversion and discounting at the aggregate level using data of convex menus from the experimental design in Andreoni and Sprenger (2012). Each column reports the estimates for the corresponding structural model discussed in the main text. Standard errors for all estimates, shown in brackets, are clustered at the individual level.