RANDOM DISCOUNTED EXPECTED UTILITY*

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

Abstract. This paper introduces the random discounted expected utility (RDEU) model, which we have developed as a means to deal with heterogeneous risk and time preferences. The RDEU 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 RDEU 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

Date: October, 2023.

*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. All the estimation programs used in this paper are available for use at https://github.com/agutieda/RDEU. 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`.

a well-founded and convenient novel framework for the analysis of heterogeneous risk and time preferences.

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 (RDEU). 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 this type of random utility models (RUMs) with multiple 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. 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 this curse of dimensionality does not apply to the RDEU model. 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 RDEU. 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 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 CRRA monetary utility functions, which are determined by a parameter r describing the curvature.[1] 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 et al. (2008) (hereafter AHLR) and Andreoni and Sprenger (2012b) (hereafter AS). We compare the aggregate and individual-level estimates of the RDEU model with those obtained using the empirical strategies employed in the respective papers and subsequent literature.

The literature has often used iid-additive RUMs in the analysis of DMPL experimental designs. 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 RDEU model offers a better overall fit. It also performs better than more

[1] 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.

recent iid-additive RUM implementations using Wilcox's (2011) correction. At the individual level, there are stark differences across models: the RDEU 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.[2] 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, the literature has often relied on estimating risk and time preferences using non-linear least squares, assuming a unique discounted expected utility. For this purpose, 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 understand 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 RDEU 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 RDEU 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). Our paper stands apart from this literature in that it focuses on risk and time preferences and establishes the comparative statics of the model.

[2] See, for instance, Holt and Laury (2002).

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.

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.[3] 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 CARA. The discount factor of the individual is denoted by $e^{-\delta}$ with $\delta \in \mathbb{R}$.[4] 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:[5]

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

We are now in a position to define the stochastic model that we analyze in the paper, that we call Random Discounted Expected Utility (RDEU) 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]

[3]More 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).

[4]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 $\frac{1}{1+d} = e^{-\delta} \in \mathbb{R}_{++}$ with d representing the discount factor in the positive reals.

[5]We assume, as it is typically done, that the awarded monetary prizes are consumed on reception.

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

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 RDEU 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 $0_{\mathcal{T}}$ 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 $([p^0, 1-$

[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 RDEU.

[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).

$p^0; (1-a)x^0, 0], t^0), ([p^1, 1-p^1; ax^1, 0], t^1)$. Given the continuous nature of convex menus, the RDEU 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 utilities.

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

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

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

[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 in Appendix D.

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 $0_{\mathcal{R}}$ 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 $\mathrm{med}(F)$ 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 $\mathrm{med}(F) = \mathrm{med}(G)$, $F(z) > G(z)$ whenever $z < \mathrm{med}(F)$ and $F(z) < G(z)$ whenever $z > \mathrm{med}(F)$.

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

[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 $K(A_{\mathcal{R}})$.

(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* $(r_1, \delta_1) = (.95, .094)$ *and* $(r_2, \delta_2) = (0, .105)$. *Although* $\delta_1 < \delta_2$ *(or equivalently* $e^{-\delta_1} = .91 > .9 = e^{-\delta_2}$*) 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 RDEU choice probabilities would now require 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.[13] 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

[13]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 $r \geq K(A_\mathcal{T}|\delta)$.

show that keeping constant the marginal distribution of one parameter, an expansion of the conditional distributions of the other always creates more stochasticity.[14]

Proposition 2. *For every pair of RDEUs, 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 $K(A_{\mathcal{T}}|\delta) \neq med(F_{r|\delta})$, $|\mathcal{P}_f(0_{\mathcal{T}}, A_{\mathcal{T}}) - \frac{1}{2}| < |\mathcal{P}_g(0_{\mathcal{T}}, A_{\mathcal{T}}) - \frac{1}{2}|$.*

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, Claim 1 states that the analysis of choice probabilities in RDEU 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, Claims 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, Claim 2 states that, by increasing either the mean of δ or the mean of r, we generate a larger

[14]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.

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, Claim 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 $\{A_{\mathcal{T},c}, A_{\mathcal{T},d}, A_{\mathcal{T},e}\}$ that 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 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 RDEU 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.

Example 2. *Let* $\omega \to 0$ *and* $(\mu_r, \sigma_r, \mu_\delta, \sigma_\delta, \rho) = (0.7, 0.7, .05, .02, -.5)$. *Consider first risk menus. The choice probability of* $([.5, .5; 50, 40], 0)$ *versus* $([.5, .5; 68, 25], 0)$ *is approximately* $1-\Phi(0) = .5$, *and since the threshold of this problem is* $.7$, *we have* $\mu_r = .7$. *The choice probability of* $([.5, .5; 50, 40], 0)$ *versus* $([.5, .5; 95, 25], 0)$ *is* $1 - \Phi(1) = .16$, *and since the threshold of this problem is* 1.4, $\sigma_r = 1.4 - .7 = .7$. *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* $.5$, *and since the constant of this menu is approximately* $\frac{.5}{3}$, *we have* $\mu_\delta = \frac{.5}{3}(1 - .7) = .05$. *Now, consider the choice probability of* $([1; 70 - \epsilon], 0)$, *with* ϵ *small, versus* $([1; 70], 1)$ *that is equal to* $.99$. *This corresponds to two and a half standard deviations of the normal, and since the constant in this case is* 0, *it follows that* $\frac{\mu_\delta}{\sigma_\delta} = \Phi^{-1}(.99) = 2.5$, *and hence* $\sigma_\delta = .02$. *Finally, the time menu involving* $([1; 68], 0)$ *and* $([1; 70], 1)$ *has constant* $.029$, *which is equal to the ratio of standard deviations. Hence, since the choice probability of the earlier option is approximately* $.98$ *(which corresponds to two standard deviations of the normal), it must be* $\rho = \frac{1}{2}\left[\frac{\frac{-.3}{.7} + 2.5}{2}\right]^2 - 1$ *which is approximately* $-.5$.

4. Double Multiple Price Lists: Empirical Illustration

In this section, we illustrate the empirical application of the RDEU 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 253 individuals with four risk tasks, each comprising up to ten risk menus. The monetary prizes of the safe and risky alternative, $(x_1^0, x_2^0, x_1^1, x_2^1)$, varied between tasks. 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 sharing the same early prize x^0. All menus in a given time task 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. Following these authors, we also assume the integrated average daily wealth value ω is common across individuals and equal to 118 Danish kroner (DKK).

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.[15] 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.[16] 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.

[15]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 (2021).

[16]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.

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 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. The mid-point of this interval is frequently used 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. 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 SPE of θ 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.

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.[17] The standard deviation of μ_r across individuals is 0.608, implying that almost half the variance in r comes from variation between

[17] Table 4 in the Appendix reports the corresponding summary statistics for the individual estimates of $(\sigma_r, \sigma_\delta)$.

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 RDEU 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 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 RDEU model.[18] Finally, note that the AHLR dataset satisfies the conditions discussed in Proposition 3. It follows that the RDEU 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 $\delta = \mu_\delta$)

[18] 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.

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).[19] 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 generates further problems in the power expression of equation (4.1) involving imaginary numbers and leading to smaller choice probabilities of better alternatives.[20]

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.

[19]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.

[20]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 RDEU.

(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.[21] The model is thus characterized by the parameters $(\mu_r, \sigma_r, \mu_\delta, \sigma_\delta)$, and we label it as WILCOX.[22]

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 RDEU, LUCE and WILCOX models, respectively.

The estimated average risk aversion coefficient μ_r under the RDEU 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 comparative statics of the RDEU model discussed in the previous section. The close relationship between the semi-parametric estimates and parametric estimates from the RDEU model illustrates the intuitive and close mapping of the parameters in the RDEU model to the variation in choices and menus in the dataset.

Comparing the results of RDEU 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 RDEU 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

[21]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.

[22]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.

the estimated models, we can see that the RDEU 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 RDEU 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 RDEU 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 RDEU 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 RDEU model. We can see that each dot in both scatterplots is close to the 45-degree line, confirming the close relationship between the RDEU and SPE estimates observed in the summary statistics of Table 2. Table 4 and Figure 2, in the appendix, 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 in 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 in blue 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 RDEU 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 RDEUs, 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 $G_{r|\delta}$, $\mathcal{P}_f([0,\overline{a}], A_{\mathcal{C}}) - \mathcal{P}_f([0,\underline{a}], A_{\mathcal{C}}) \geq \mathcal{P}_g([0,\overline{a}], A_{\mathcal{C}}) - \mathcal{P}_g([0,\underline{a}], A_{\mathcal{C}})$ for every $0 < \underline{a} < \bar{e} < \overline{a} < 1$.*

(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 $(\mu_r, \sigma_r, \mu_\delta, \sigma_\delta, \rho) = (0.7, 0.7, .05, .02, -.5)$. Take first the convex problem defined by probabilities $p^0 = 1$ and $p^1 = .8$, 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 $a = .48$ is $.7$. Since the cumulative choice probability at $a = .48$ is $.5$, we learn that $\mu_r = .7$, and from the above expression, $\sigma_r = .7$. 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 RDEU 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 RDEU 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 $\alpha^1 = [a_1, a_2]$, $\alpha^2 = [a_2, a_3]$, $\ldots$, $\alpha^S = [a_S, a_{S+1}]$, 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 RDEU 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:

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

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 RDEU 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.[23]

[23] 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 underlying parameters. RDEU formalizes this principle in terms of behavioral variation, allowing to produce explicit probabilistic predictions.

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 3 presents the estimated parameters under each model. Regarding risk aversion, the RDEU 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 RDEU model allows it to match the large fraction of corner solutions with a slightly concave utility function.[24]

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 RDEU model, which is a substitute for a larger discount factor to explain choices between $a = 0$ and $a = 1$ at the corners. Finally, we 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.

[24]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.

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 RDEU 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.[25] It can be seen that the RDEU model matches the choice patterns observed in convex budgets both in the full sample and for particular menus.

The RDEU 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 RDEU 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 in the Appendix 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.

[25]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. The graphs for all the menus can be found in the supplementary material to this paper.

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.

Appendix A. Proofs

Proof of Proposition 1: Consider a menu $A_{\mathcal{R}} = \{0_{\mathcal{R}}, 1_{\mathcal{R}}\} = \{([p, 1-p; x_1^0, x_2^0], 0), ([p, 1-p; x_1^1, x_2^1], 0)\}$, such that $x_1^1 > x_1^0 > x_2^0 > x_2^1$ and $p \in (0,1)$. Consider $r < r'$. Construct the affine transformations $v_r, v_{r'}$ of $u_r, u_{r'}$ satisfying $v_r(\omega + x_2^1) = v_{r'}(\omega + x_2^1) = 0$ and $v_r(\omega + x_2^0) = v_{r'}(\omega + x_2^0) = 1$. By strict monotonicity of the original utility functions, it must be $v_r(\omega + x_1^0) > 1$ and $v_{r'}(\omega + x_1^0) > 1$. We now claim that $v_{r'}(\omega + x_1^0) < v_r(\omega + x_1^0)$ must hold, and this will be proved by contradiction. Assume that $v_{r'}(\omega + x_1^0) \geq v_r(\omega + x_1^0) > 1$. In this case, we can consider the lotteries $[p^*, 1-p^*; x_1^0, x_2^1]$ and $[1; x_2^0]$, with $\frac{1}{v_{r'}(\omega + x_1^0)} \leq p^* \leq \frac{1}{v_r(\omega + x_1^0)}$. It is immediate to see that the expected utility constructed upon v_r leads to, at least, weakly prefer lottery $[1; x_2^0]$ while the expected utility constructed upon $v_{r'}$ leads to, at least, weakly prefer lottery $[p^*, 1-p^*; x_1^0, x_2^1]$. This contradicts the fact that $v_{r'}$, being a strict concave transformation of v_r, must have a strictly lower certainty equivalent for the second, riskier lottery and hence, we have proved that $v_{r'}(\omega + x_1^0) < v_r(\omega + x_1^0)$.

We now claim that $v_{r'}(\omega + x_1^1) - v_{r'}(\omega + x_1^0) < v_r(\omega + x_1^1) - v_r(\omega + x_1^0)$ must hold, and prove it by contradiction. If it were not true, given that we already proved $v_{r'}(\omega + x_1^0) < v_r(\omega + x_1^0)$, we would have $\frac{v_{r'}(\omega + x_1^1) - v_{r'}(\omega + x_1^0)}{v_{r'}(\omega + x_1^0)} > \frac{v_r(\omega + x_1^1) - v_r(\omega + x_1^0)}{v_r(\omega + x_1^0)}$. Considering the lotteries $[p', 1-p'; x_1^1, x_2^1]$ and $[1; x_1^0]$, with $\frac{v_{r'}(\omega + x_1^1) - v_{r'}(\omega + x_1^0)}{v_{r'}(\omega + x_1^0)} > \frac{1-p'}{p'} > \frac{v_r(\omega + x_1^1) - v_r(\omega + x_1^0)}{v_r(\omega + x_1^0)}$, the expected utility constructed upon v_r would lead to the choice of $[1; x_1^0]$ while the expected utility constructed upon $v_{r'}$ would lead to the choice of $[p', 1-p'; x_1^1, x_2^1]$. This is again a contradiction with the concavity assumption, and concludes the argument; it must then be $v_{r'}(\omega + x_1^1) - v_{r'}(\omega + x_1^0) < v_r(\omega + x_1^1) - v_r(\omega + x_1^0)$.

We now claim that for every (r,δ) and (r',δ') such that $r'>r$, if $DEU_{r,\delta}(0_{\mathcal{R}}) \geq DEU_{r,\delta}(1_{\mathcal{R}})$, then $DEU_{r',\delta'}(0_{\mathcal{R}}) > DEU_{r',\delta'}(1_{\mathcal{R}})$. To see this, suppose that $DEU_{r,\delta}(0_{\mathcal{R}}) \geq DEU_{r,\delta}(1_{\mathcal{R}})$. This is equivalent to claim that the expected utility of lottery $[p,1-p;x_1^0,x_2^0]$ is greater than the expected utility of lottery $[p,1-p;x_1^1,x_2^1]$ when the monetary utility u_r is used. That is equivalent to claim that the expected utility of lottery $[p,1-p;x_1^0,x_2^0]$ is greater than the expected utility of lottery $[p,1-p;x_1^1,x_2^1]$ when the monetary utility v_r is used, and can be written as $\frac{1-p}{p} \geq \frac{v_r(\omega+x_1^1)-v_r(\omega+x_1^0)}{v_r(\omega+x_2^0)-v_r(\omega+x_2^1)} = v_r(\omega+x_1^1)-v_r(\omega+x_1^0)$. From our previous claims, we know that it must be $\frac{1-p}{p} > v_{r'}(\omega+x_1^1)-v_{r'}(\omega+x_1^0) = \frac{v_{r'}(\omega+x_1^1)-v_{r'}(\omega+x_1^0)}{v_{r'}(\omega+x_2^0)-v_{r'}(\omega+x_2^1)}$, which implies that the first lottery is strictly preferred to the second using $v_{r'}$ or, alternatively, using $u_{r'}$. This implies $DEU_{r',\delta'}(0_{\mathcal{R}}) > DEU_{r',\delta'}(1_{\mathcal{R}})$ and concludes the argument. With the unbounded curvature assumption, the certainty equivalent of both lotteries must converge to the maximum and minimum payout when r tends to $-\infty$ and $+\infty$, respectively. That is, there are values of r for which $0_{\mathcal{R}}$ and $1_{\mathcal{R}}$ are preferred. As a result, there must be a unique $K(A_{\mathcal{R}}) \in \mathbb{R}$ such that alternative $0_{\mathcal{R}}$ is preferred if and only if $r \geq K(A_{\mathcal{R}})$ which leads to Claim 1 in the proposition. For Claim 2, note that whenever F^r dominates G^r, it must be $\mathcal{P}_f(0_{\mathcal{R}},A_{\mathcal{R}}) = 1-F^r(K(A_{\mathcal{R}})) > 1-G^r(K(A_{\mathcal{R}})) = \mathcal{P}_g(0_{\mathcal{R}},A_{\mathcal{R}})$. For Claim 3, notice that the assumption requires us to consider two cases, $\text{med}(F^r) > K(A_{\mathcal{R}})$ or $\text{med}(F^r) < K(A_{\mathcal{R}})$. In the first case, since F^r expands G^r, it must be $F^r(K(A_{\mathcal{R}})) > G^r(K(A_{\mathcal{R}})) > 1/2$, while in the second case, it must be that $F^r(K(A_{\mathcal{R}})) < G^r(K(A_{\mathcal{R}})) < 1/2$, concluding the proof. ■

Proof of Proposition 2: Consider a menu $A_{\mathcal{T}} = \{0_{\mathcal{T}}, 1_{\mathcal{T}}\} = \{([1;x^0],t^0),([1;x^1],t^1)\}$ such that $t^0 < t^1$ and $x^0 < x^1$. From the definition of DEU, it follows immediately that

$$DEU_{r,\delta}(0_{\mathcal{T}}) \geq DEU_{r,\delta}(1_{\mathcal{T}}) \Leftrightarrow \delta \geq K(A_{\mathcal{T}}|r) = \frac{1}{t^1-t^0}\log\frac{u_r(\omega+x^1)}{u_r(\omega+x^0)}.$$

Strict monotonicity of u_r guarantees that this threshold is always a positive real value and, hence, the first expression in Claim 1, and Claim 2a, follow immediately.

We now claim that the threshold map $\{K(A_{\mathcal{T}}|r)\}_{r\in\mathbb{R}}$ is strictly decreasing in r. To see this, notice that scalar transformations leave DEU decisions unaffected. Hence, we can select the scalar transformation v_r of u_r for which $v_r(\omega+x^1)=1$ holds and, then, we are only required to show that $\frac{-\log v_r(\omega+x^0)}{t^1-t^0}$ is strictly decreasing in r or, equivalently, that $v_r(\omega+x^0)$ is strictly increasing in r. Suppose by contradiction that this is not the case, i.e., $v_r(\omega + x^0) \geq v_{r'}(\omega + x^0)$ with $r < r'$. By considering the lotteries $[p^*, 1-p^*; \omega, \omega + x^1]$ and $[1; \omega + x^0]$, where $v_r(\omega + x^0) \geq p^* \geq v_{r'}(\omega + x^0)$, it is immediate to see that the expected utility, using v_r, is larger for the first lottery than for the second, while the expected utility, using $v_{r'}$, is larger for the second lottery than for the first, a contradiction with the strict concavity assumption. The threshold map is thus strictly decreasing in r.

The unbounded curvature assumption also proves that the map is onto for $\mathbb{R}_{++}$ and hence, it is a bijective map between $\mathbb{R}$ and $\mathbb{R}_{++}$. Thus, it can be inverted to obtain the strictly decreasing thresholds $\{K(A_{\mathcal{T}}|\delta)\}_{\delta \in \mathbb{R}_{++}}$, such that, for a given $\delta > 0$, alternative $0_{\mathcal{T}}$ is chosen if and only if r is above this threshold. For $\delta \leq 0$ alternative $1_{\mathcal{T}}$ is always chosen. Hence, $\mathcal{P}_f(0_{\mathcal{T}}, A_{\mathcal{T}}) = \int_r (1 - F_{\delta|r}(K(A_{\mathcal{T}}|r))) f^r(r) \mathrm{d}r = \int_{\delta > 0} (1 - F_{r|\delta}(K(A_{\mathcal{T}}|\delta))) f^\delta(\delta) \mathrm{d}\delta$, where f^δ is the marginal density of δ. The second expression in Claim 1, and Claim 2b, follow.

For Claim 3a, we just need to reproduce the logic of Proposition 1, expanding separately each of the conditional distributions $F_{\delta|r}$. This always creates a strictly larger conditional stochasticity. From there, we need to prove that the argument extends to the weighted aggregation of all these conditional distributions. To see this, notice that the continuity of the map $\{K(A_{\mathcal{T}}|r)\}_{r \in \mathbb{R}}$ guarantees that all conditional medians of δ lie on the same side of the threshold map. As a result, the same alternative, either $0_{\mathcal{T}}$ or $1_{\mathcal{T}}$, is chosen more often in each of the conditionals, and the expansion argument extends to the aggregation. For Claim 3b, a similar argument holds by expanding the conditionals $F_{r|\delta}$ and using the continuity of the inverse map. ∎

Proof of Proposition 3: We start by showing that $\theta_r \equiv (\mu_r, \sigma_r)$ is identified. Assume, on the contrary, that this is not the case: there exist θ'_r and θ^*_r in Θ_r such that $\theta'_r \neq \theta^*_r$ and $\mathcal{P}_{\theta'}(1_{\mathcal{R}}, A_{\mathcal{R}}) = \mathcal{P}_{\theta^*}(1_{\mathcal{R}}, A_{\mathcal{R}})$. Using Proposition 1, we can write this equality as:

$$\mathcal{P}_{\theta'_r}(r \leq K(A_{\mathcal{R}})) = \Phi\left(\frac{K(A_{\mathcal{R}}) - \mu'_r}{\sigma'_r}\right) = \Phi\left(\frac{K(A_{\mathcal{R}}) - \mu^*_r}{\sigma^*_r}\right) = \mathcal{P}_{\theta^*_r}(r \leq K(A_{\mathcal{R}})).$$

Since the $\Phi(\cdot)$ is a strictly monotonic function, the last equality implies that:

$$-\frac{\mu^*_r}{\sigma^*_r} + \frac{1}{\sigma^*_r} K(A_{\mathcal{R}}) = -\frac{\mu'_r}{\sigma'_r} + \frac{1}{\sigma'_r} K(A_{\mathcal{R}}), \tag{A.1}$$

for every menu $A_{\mathcal{R}}$. Now, since $\theta'_r \neq \theta^*_r$, there are three possible cases: $\mu'_r \neq \mu^*_r$ and $\sigma'_r = \sigma^*_r$; $\mu'_r = \mu^*_t$ and $\sigma'_r \neq \sigma^*_r$; and $\mu'_r \neq \mu^*_r$ and $\sigma'_r \neq \sigma^*_r$. In the first case, equation (A.1) implies $\mu'_r = \mu^*_r$, leading to a contradiction. Consider now the second and third cases where $\sigma_r' \neq \sigma_r^*$. Evaluating (A.1) for $A_{\mathcal{R},a}$ and $A_{\mathcal{R},b}$ and combining the resulting expressions yields:

$$\left(\frac{1}{\sigma_r^*} - \frac{1}{\sigma_r'}\right)\left(K\left(A_b^{\mathcal{R}}\right) - K\left(A_a^{\mathcal{R}}\right)\right) = 0. \tag{A.2}$$

Since $K\left(A_{\mathcal{R},a}\right) \neq K\left(A_{\mathcal{R},b}\right)$ by assumption (a), it must be the case that $\sigma_r' = \sigma_r^*$, arriving to a contradiction. We thus conclude that θ_r is identified. The next step is to show that $\theta_\delta \equiv (\mu_\delta, \sigma_\delta, \rho)$ is identified. Fix (μ_r, σ_r) and assume, on the contrary, that θ_δ is not identified: there exist θ_δ' and θ_δ^* in Θ such that $\theta_\delta' \neq \theta_\delta^*$ and $\mathcal{P}_{\theta_\delta'}\left(0_{\mathcal{T}}, A_{\mathcal{T}}\right) = \mathcal{P}_{\theta_\delta^*}\left(0_{\mathcal{T}}, A_{\mathcal{T}}\right)$. Using Proposition 2, the equality of probabilities implies:

$$\int_{-\infty}^{\infty}\left\{\Phi\left(\frac{K\left(A_{\mathcal{T}}|r\right) - \mu_{\delta|r}'}{\sigma_{\delta|r}'}\right) - \Phi\left(\frac{K\left(A_{\mathcal{T}}|r\right) - \mu_{\delta|r}^*}{\sigma_{\delta|r}^*}\right)\right\}\phi\left(\frac{r-\mu_r}{\sigma_r}\right)dr = 0,$$

with $\mu_{\delta|r} = \mu_\delta + \rho\sigma_\delta v\left(r\right)$, $\sigma_{\delta|r} = \sigma_\delta\sqrt{1-\rho^2}$, and $\nu\left(r\right) \equiv \frac{r-\mu_r}{\sigma_r}$. The term in brackets in the previous expression is bounded in $[-1,1]$. By the continuity and monotonicity of $\phi\left(\cdot\right)$ and $\Phi\left(\cdot\right)$, there exists $\overline{r}_m \in \mathbb{R}$ for each one of the menus $\{A_{\mathcal{T},c}, A_{\mathcal{T},d}, A_{\mathcal{T},e}\}$ such that:

$$-\frac{\mu_\delta'}{\sigma_{\delta|r}'} - \alpha' v\left(\overline{r}_m\right) + \frac{1}{\sigma_{\delta|r}'}K\left(A_{\mathcal{T},m}|\overline{r}_m\right) = -\frac{\mu_\delta^*}{\sigma_{\delta|r}^*} - \alpha^* v\left(\overline{r}_m\right) + \frac{1}{\sigma_{\delta|r}^*}K\left(A_{\mathcal{T},m}|\overline{r}_m\right), \tag{A.3}$$

where $\alpha \equiv \rho/\sqrt{1-\rho^2}$ and we use the fact that $\sigma_{\delta|r}$ is independent of $\overline{r}_m$. Now, any of the three conditions in Assumption (b) implies $K\left(A_{\mathcal{T},c}|r\right) \neq K\left(A_{\mathcal{T},d}|r\right) \neq K\left(A_{\mathcal{T},e}|r\right) \neq K\left(A_{\mathcal{T},c}|r\right)$ for any $r \in \mathbb{R}$. Using this result and the Implicit Function Theorem, we conclude that that $\overline{r}_c \neq \overline{r}_d \neq \overline{r}_e \neq \overline{r}_c$. Setting $m = c$ in (A.3) and subtracting the corresponding expression for $m = d$, we get:

$$\left[\alpha' - \alpha^*\right]\left(v\left(\overline{r}_c\right) - v\left(\overline{r}_d\right)\right) + \left[\frac{1}{\sigma_{\delta|r}'} - \frac{1}{\sigma_{\delta|r}^*}\right]\left(K\left(A_{\mathcal{T},c}|\overline{r}_c\right) - K\left(A_{\mathcal{T},d}|\overline{r}_d\right)\right) = 0. \tag{A.4}$$

Repeating this procedure for menus $A_{\mathcal{T},c}$ and $A_{\mathcal{T},e}$, we get:

$$\left[\alpha' - \alpha^*\right]\left(v\left(\overline{r}_c\right) - v\left(\overline{r}_e\right)\right) + \left[\frac{1}{\sigma_{\delta|r}'} - \frac{1}{\sigma_{\delta|r}^*}\right]\left(K\left(A_{\mathcal{T},c}|\overline{r}_c\right) - K\left(A_{\mathcal{T},e}|\overline{r}_e\right)\right) = 0. \tag{A.5}$$

Using (A.5) to replace $\left[1/\sigma_{\delta|r}' - 1/\sigma_{\delta|r}^*\right]$ in (A.4), we get:

$$\left[\alpha' - \alpha^*\right]\left[\overline{r}_c - \overline{r}_d\right]\left[1 - \left(\frac{\overline{r}_c - \overline{r}_e}{\overline{r}_c - \overline{r}_d}\right)\left(\frac{K\left(A_{\mathcal{T},c}|\overline{r}_c\right) - K\left(A_{\mathcal{T},d}|\overline{r}_d\right)}{K\left(A_{\mathcal{T},c}|\overline{r}_c\right) - K\left(A_{\mathcal{T},e}|\overline{r}_e\right)}\right)\right] = 0. \tag{A.6}$$

Since $K\left(A_{\mathcal{T},c}|\overline{r}_c\right) \neq K\left(A_{\mathcal{T},d}|\overline{r}_d\right) \neq K\left(A_{\mathcal{T},e}|\overline{r}_e\right) \neq K\left(A_{\mathcal{T},c}|\overline{r}_c\right)$, equation (A.6) implies that $\alpha' = \alpha^*$, which in turn implies $\rho' = \rho^*$. This result and equation (A.5) implies that $\sigma'_{\delta|r} = \sigma^*_{\delta|r}$, so that $\sigma'_\delta = \sigma^*_\delta$. Using equation (A.3), we conclude that $\mu'_\delta = \mu^*_\delta$, arriving to a contradiction. This concludes the proof. ■

Proof of Proposition 4: To simplify the presentation of the result and provide neat intuitions, we use two relatively mild assumptions: (i) the probability of the event $\{r > 1, \delta < 0\}$ is small and (ii) there are two time menus $A_{\mathcal{T}_1}$ and $A_{\mathcal{T}_2}$ with the same payouts where the probabilities of selecting options $0_{\mathcal{T}_1}$ and $1_{\mathcal{T}_2}$ are greater than $\frac{1}{2}$. To motivate (i), notice that $\delta < 0$ already corresponds to the rare event in which the individual has a strict preference for the future, and we are compounding this with the extra effect of a more-than-logarithmic curvature. To motivate (ii), notice that we can make the differences in the timings as small or as large as desired.

Now, from Proposition 1 we know that $\mathcal{P}_f(0_\mathcal{R}, A_\mathcal{R}) = 1 - F^r(K(A_\mathcal{R}))$, which in the parametric version reads as $\mathcal{P}_f(0_\mathcal{R}, A_\mathcal{R}) = 1 - \Phi(\frac{K(A_\mathcal{R})-\mu_r}{\sigma_r})$. Fix any value of $p \in (0,1)$, three payouts $x_1^0 > x_2^0 > x_2^1 > 0$, and consider risk menus that vary only on the payout x_1^1. The thresholds of these menus are strictly increasing in x_1^1 and form a bijection with the real numbers. Hence, there exist two menus $A_{\mathcal{R}_1}$ and $A_{\mathcal{R}_2}$ such that the choice probabilities for options $0_{\mathcal{R}_1}$ and $0_{\mathcal{R}_2}$ are equal to $1 - \Phi(0)$ and $1 - \Phi(1)$, respectively. It must then be $\mu_r = K(A_{\mathcal{R}_1})$ and $\sigma_r = K(A_{\mathcal{R}_2}) - \mu_r = K(A_{\mathcal{R}_2}) - K(A_{\mathcal{R}_1})$.

We now discuss time menus and, by using sufficiently large payouts in all our arguments, we can assume w.l.o.g. that behavior corresponds to the case $\omega \to 0$. In this limit case the conditional threshold map becomes piece-wise linear: (a) when $r < 1$, $K(A_\mathcal{T}|r) = \frac{\log \frac{x^1}{x^0}}{t^1-t^0}(1-r) \equiv K(A_\mathcal{T})(1-r) > 0$, where $K(A_\mathcal{T})$ is a menu-dependent constant, and (b) when $r \geq 1$, $K(A_\mathcal{T}|r)$ becomes null. Hence, the probability of choosing $0_\mathcal{T}$ corresponds to the probability that δ lies above $\min\{0, K(A_\mathcal{T})(1-r)\}$ and, given our assumption (i), this can be approximated by the probability that δ lies above $K(A_\mathcal{T})(1-r)$. This is the probability that the random variable $K(A_\mathcal{T})(1-r) - \delta$ lies below zero. Given normality this random variable is also normal with mean $-\mu_\delta - K(A_\mathcal{T})\mu_r + K(A_\mathcal{T})$ and standard deviation $\sqrt{K^2(A_\mathcal{T})\sigma_r^2 + \sigma_\delta^2 + 2\rho K(A_\mathcal{T})\sigma_r\sigma_\delta}$. Thus the choice probability of $0_\mathcal{T}$ is approximately $\Phi\big(\frac{\mu_\delta + K(A_\mathcal{T})\mu_r - K(A_\mathcal{T})}{\sqrt{K^2(A_\mathcal{T})\sigma_r^2 + \sigma_\delta^2 + 2\rho K(A_\mathcal{T})\sigma_r\sigma_\delta}}\big)$.

By assumption (ii) there exist two time menus $A_{\mathcal{T}_1}$ and $A_{\mathcal{T}_2}$ with the same payouts x^0, x^1 such that the choice probabilities of $0_\mathcal{T}$ are above and below $\Phi(0) = \frac{1}{2}$, respectively. Due to the stationarity of DEU, we can assume w.l.o.g. that the earlier payout takes place in the present in both cases. By continuity there must exist a unique t, and hence a menu $A_{\mathcal{T}_3} = \{([1; x^0], 0), ([1; x^1], t)\}$ such that the choice probability of option

$0_{\mathcal{T}_3}$ is exactly $\Phi(0)$. Hence, $\mu_\delta = K(A_{\mathcal{T}_3})(1-\mu_r) = K(A_{\mathcal{T}_3})(1-K(A_{\mathcal{R}_1}))$. Second, consider any sequence of time problems $\{A_{\mathcal{T}^n}\}$ such that $\lim_n K(A_{\mathcal{T}^n}) = 0$. Denote by q the limit of the choice probabilities of option $0_{\mathcal{T}^n}$. We know that $q = \Phi(\frac{\mu_\delta}{\sigma_\delta})$ and, hence, it must be $\sigma_\delta = \frac{\mu_\delta}{\Phi^{-1}(q)} = \frac{K(A_{\mathcal{T}_3})(1-K(A_{\mathcal{R}_1}))}{\Phi^{-1}(q)}$. Finally, by fixing again any three parameters in a time menu and varying the fourth, we know that there exists a unique time menu $A_{\mathcal{T}_4}$ in such a family for which $K(A_{\mathcal{T}_4}) = \frac{\sigma_\delta}{\sigma_r}$.[26] Denote by q' the choice probability of option $0_{\mathcal{T}_4}$. It must then be $q' = \Phi(\frac{\frac{\mu_r-1}{\sigma_r}+\frac{\mu_\delta}{\sigma_\delta}}{\sqrt{2(1+\rho)}})$. Notice that the right-hand side map is either strictly increasing or strictly decreasing in ρ, which allows to obtain ρ: $\rho = \frac{1}{2}\left[\frac{\frac{\mu_r-1}{\sigma_r}+\frac{\mu_\delta}{\sigma_\delta}}{\Phi^{-1}(q')}\right]^2 - 1 = \frac{1}{2}\left[\frac{\frac{K(A_{\mathcal{R}_1})-1}{K(A_{\mathcal{R}_2})-K(A_{\mathcal{R}_1})}+\Phi^{-1}(q)}{\Phi^{-1}(q')}\right]^2 - 1$. This concludes the proof.■

Proof of Proposition 5: Consider a menu $A_{\mathcal{C}}$ defined by $(p^0, x^0, t^0; p^1, x^1, t^1)$. We first claim that, for every $r \in \mathbb{R}$, the argument that maximizes $DEU_{r,\delta}$ is decreasing in δ. To see this, consider any pair of parameters (r, δ), and let $a^* \in [0,1]$ be the argument that maximizes $DEU_{r,\delta}$. If $a^* = 1$, we are done. Consider then the case of $a^* < 1$ and any alternative $a^* < a \leq 1$. Given the optimality of a^*, we know that $DEU_{\delta,r}(a) \leq DEU_{\delta,r}(a^*)$, i.e., $e^{-\delta t^0}p^0u_r(\omega + (1-a)x^0) + e^{-\delta t^1}p^1u_r(\omega + ax^1) \leq e^{-\delta t^0}p^0u_r(\omega+(1-a^*)x^0)+e^{-\delta t^1}p^1u_r(\omega+a^*x^1)$ holds. The latter inequality is equivalent to $p^0u_r(\omega+(1-a)x^0)+e^{-\delta(t^1-t^0)}p^1u_r(\omega+ax^1) \leq p^0u_r(\omega+(1-a^*)x^0)+e^{-\delta(t^1-t^0)}p^1u_r(\omega+a^*x^1)$. Now, it is evident that an increase of δ leaves unaffected the first term in both the left and the right hand sides but decreases more significantly the second term of the left hand side, because the function u_r is strictly increasing. Hence, alternative a^* must be preferred to alternative a for the larger δ, and the argument maximizing DEU must be a^* or smaller. We have proved our claim. Hence, given $r \in \mathbb{R}$, we can define $K(a, A_{\mathcal{C}}|r)$, $a \in [0,1)$, as the infimum of the values of δ for which any alternative in $[0,a]$ is the DEU maximizer. Hence, the maximizer of $DEU_{r,\delta}$ is below a if and only if δ lies above $K(a, A_{\mathcal{C}}|r)$, i.e., $\Gamma(a, A) = \{(r,\delta) : \delta \leq K(a, A_{\mathcal{C}}|r)\}$, and Claims 1 and 2a follow.

We now study the structure of the thresholds. We start with the case of convex monetary utilities, i.e., $r \leq 0$. Convexity and the fact that $u_r(\omega) = 0$ guarantee that $e^{-\delta t^0}p^0u_r(\omega + (1-a)x^0) + e^{-\delta t^1}p^1u_r(\omega + ax^1) \leq e^{-\delta t^0}p^0[au_r(\omega) + (1-a)u_r(\omega + x^0)] +$

[26]This is in general different to $A_{\mathcal{T}_3}$. Otherwise, notice that the mapping $\Phi(\frac{\mu_\delta+K(A_\mathcal{T})\mu_r-K(A_\mathcal{T})}{\sqrt{K^2(A_\mathcal{T})\sigma_r^2+\sigma_\delta^2+2\rho K(A_\mathcal{T})\sigma_r\sigma_\delta}})$ is strictly monotone in ρ, and hence the parameter can be recovered using some other time menu.

$e^{-\delta t^1}p^1[(1-a)u_r(\omega)+au_r(\omega+x^1)] = e^{-\delta t^0}p^0(1-a)u_r(\omega+x^0)+e^{-\delta t^1}p^1 a u_r(\omega+x^1) \leq \max\{e^{-\delta t^0}p^0 u_r(\omega+x^0), e^{-\delta t^1}p^1 u_r(\omega+x^1)\}$. Hence, only alternatives 0 or 1 can be the maximizers of $DEU_{r,\delta}$. Thus, for every menu $A_{\mathcal{C}}$ there is a unique threshold $K(a, A_{\mathcal{C}}|r) \in \mathbb{R}$, independent of a, that corresponds to the δ that, given r, equalizes the DEU value of 0 and 1. This value is $\frac{1}{t^1-t^0}\log\left[\frac{p^1 u_r(\omega+x^1)}{p^0 u_r(\omega+x^0)}\right]$, that can also be written as $\frac{1}{t^1-t^0}\log\frac{p^1}{p^0}+K(A_{\mathcal{T}}|r)$, with $K(A_{\mathcal{T}}|r)$ referring to the hypothetical time menu in which prizes x^0 and x^1 are offered at periods t^0 and t^1, without considering the probability of these prizes. Proposition 2 argued that $K(A_{\mathcal{T}}|r)$ is strictly decreasing, and hence $K(a, A_{\mathcal{C}}|r)$ is also strictly decreasing whenever $r \leq 0$.

We now analyze strictly concave utilities, $r > 0$. We start by claiming that the threshold $K(a, A_{\mathcal{C}}|r)$ is decreasing for every $a > \bar{e}$, and increasing for every $a < \bar{e}$. We start with the former, assuming by contradiction that $0 < r < r'$ but $K(a, A_{\mathcal{C}}|r) < K(a, A_{\mathcal{C}}|r')$ for some $a > \bar{e}$. Using continuity and the definition of the thresholds, there must exist δ^* with $K(a, A_{\mathcal{C}}|r) < \delta^* < K(a, A_{\mathcal{C}}|r')$ such that the maximizer for DEU_{r,δ^*} is a^*, with $\bar{e} < a^* < a$. Consider any $a' > a^*$. It must be $DEU_{r,\delta^*}(a^*) \geq DEU_{r,\delta^*}(a')$, i.e., $e^{-\delta^* t^0}p^0 u_r(\omega+(1-a^*)x^0)+e^{-\delta^* t^1}p^1 u_r(\omega+a^*x^1) \geq e^{-\delta^* t^0}p^0 u_r(\omega+(1-a')x^0)+e^{-\delta^* t^1}p^1 u_r(\omega+a'x^1)$. Dividing both terms by the positive constant $p^0e^{-\delta^* t^0}+p^1e^{-\delta^* t^1}$ and denoting $p = \frac{p^0e^{-\delta^* t^0}}{p^0e^{-\delta^* t^0}+p^1e^{-\delta^* t^1}}$, the former expression can be written as $pu_r(\omega+(1-a^*)x^0)+(1-p)u_r(\omega+a^*x^1) \geq pu_r(\omega+(1-a')x^0)+(1-p)u_r(\omega+a'x^1)$. Hence, the comparison of these two alternatives is equivalent to that of a risk menu, with a^* corresponding to alternative $0_{\mathcal{R}}$ and a' to alternative $1_{\mathcal{R}}$. Hence, since a^* is preferred at (r, δ^*), we know from Proposition 1 that a^* will also be preferred at (r', δ^*) because $r' > r$.[27] Thus, the maximizer of DEU_{r',δ^*} cannot be above a^*. This contradicts the definition of $K(a, A_{\mathcal{C}}|r')$. That is, the threshold must be decreasing whenever $a > \bar{e}$. Given that the family $\{u_r\}$ is strictly ordered by concavity, the threshold must be strictly decreasing. The proof that the threshold is strictly increasing whenever $a < \bar{e}$ is analogous and thus omitted.

Consider now $0 < \underline{a} < \bar{e} < \overline{a} < 1$ and denote by $\delta^{\bar{e}}$ the value of δ that makes indifferent all the alternatives when $r = 0$. From the previous reasoning, whenever $r > 0$, $K(\underline{a}, A_{\mathcal{C}}|r)$ (resp., $K(\overline{a}, A_{\mathcal{C}}|r)$) is above (resp., below) $\delta^{\bar{e}}$. Thus, for any given $\delta > \delta^{\bar{e}}$ (resp., $\delta < \delta^{\bar{e}}$), a dominating change in the conditional distribution of r creates

[27] Notice that we are maintaining δ^* constant because this value is part of the definition of the lotteries.

an increase in the conditional mass of the set of values of r that lie above the inverse of threshold $K(\underline{a}, A_{\mathcal{C}}|r)$ (resp., $K(\overline{a}, A_{\mathcal{C}}|r)$). Claim 2b then follows.

We now prove Claim 3a. We know that the choice belongs to $[0, \bar{e}]$ if and only if $\delta > K(\bar{e}, A_{\mathcal{C}}|r)$. We can reproduce the analysis of Proposition 2 for the case of expansions in the conditional distribution of δ, and thus, Claim 3a follows. To show Claim 3b, consider any sequence of values $\{a^n\}$, with $a^n > \bar{e}$, such that $\lim_n a^n = \bar{e}$. We know that the choice belongs to $[0, a^n]$ if and only if $\delta > K(a^n, A_{\mathcal{C}}|r)$. Given that for every a^n the threshold is strictly decreasing, we can invert these maps and then reproduce the analysis of Proposition 2 for the case of expansions in the conditional distribution of r, and thus, Claim 3b follows. ∎

Proof of Proposition 6: We start by showing that $\theta_r \equiv (\mu_r, \sigma_r)$ is identified. Assume, on the contrary, that this is not the case: there exists θ' and θ^* in Θ with $\theta'_r \neq \theta^*_r$ such that the distribution of the data is the same under both parameters. Let $\log \widetilde{R} \equiv \log \frac{x^1}{x^0} + \log \frac{p^1}{p^0}$ and $k \equiv t^1 - t^0$. Note that, under the assumption that $x^0/\omega \to 0$ and $x^1/\omega \to 0$, the probability of corner allocations, conditional on $r > 0$, converges to zero. Consequently, the probability of observing corner allocations $a \in \{0, 1\}$ is given by $\mathcal{P}_{\theta'}(a \in \{0,1\}, A_{\mathcal{C}}) = \mathcal{P}_{\theta'}(r \leq 0) = \Phi(-\mu'_r/\sigma'_r)$ which, by the assumption of no-identification, is equal to $\Phi(-\mu^*_r/\sigma^*_r)$. It follows that $\mu'_r/\sigma'_r = \mu^*_r/\sigma^*_r$. Consequently, both $\mu'_r \neq \mu^*_r$ and $\sigma'_r \neq \sigma^*_r$ must hold. Otherwise, $\theta'_r = \theta^*_r$. Now, using the first order condition of the problem, we have:

$$\mathbb{E}_{\theta'}\left[\log\left(\frac{ax^1_m + \omega}{(1-a)\,x^0_m + \omega}\right)|r > 0\right] = -\mathbb{E}_{\theta'}\left[\frac{\delta}{r}|r > 0\right] k_m + \mathbb{E}_{\theta'}\left[\frac{1}{r}|r > 0\right] \log \widetilde{R}_m,$$

for any menu $A_{\mathcal{C},m}$. Combining the corresponding expressions for $A_{\mathcal{C},a}$ and $A_{\mathcal{C},b}$, and using the assumption that $k_a = k_b$, we get:

$$\mathbb{E}_{\theta'}\left[\Delta c|r > 0\right] = \mathbb{E}_{\theta'}\left[\frac{1}{r}|r > 0\right]\left(\log \widetilde{R}_a - \log \widetilde{R}_b\right), \tag{A.7}$$

with

$$\Delta c \equiv \log\left(\frac{ax^1_a + \omega}{(1-a)\,x^0_a + \omega}\right) - \log\left(\frac{ax^1_b + \omega}{(1-a)\,x^0_b + \omega}\right).$$

Since the model is not identified, it must also be the case that $\mathbb{E}_{\theta'}[\Delta c|r > 0] = \mathbb{E}_{\theta^*}[\Delta c|r > 0]$. From equation (A.7), it follows that:

$$\left(\mathbb{E}_{\theta'}\left[\frac{1}{r}|r > 0\right] - \mathbb{E}_{\theta^*}\left[\frac{1}{r}|r > 0\right]\right)\left(\log \widetilde{R}_a - \log \widetilde{R}_b\right) = 0.$$

Since $\widetilde{R}_a \neq \widetilde{R}_b$ by assumption, the previous expression implies that $\mathbb{E}_{\theta'}\left[\frac{1}{r}|r>0\right] = \mathbb{E}_{\theta^*}\left[\frac{1}{r}|r>0\right]$. Using the fact that r follows a normal distribution, we can write $\mathbb{E}_{\theta'}\left[\frac{1}{r}|r>0\right]$ as:

$$\begin{aligned}\mathbb{E}_{\theta'}\left[\frac{1}{r}|r>0\right] &= \frac{1}{1-\Phi\left(-\mu_r'/\sigma_r'\right)}\int_0^\infty \frac{1}{x}\left(\frac{1}{\sigma_r'}\right)\phi\left(\frac{x-\mu_r'}{\sigma_r'}\right)\\ &= \frac{1}{\Phi\left(\mu_r'/\sigma_r'\right)\sigma_r'}\int_0^\infty \frac{1}{z}\phi\left(z-\mu_r'/\sigma_r'\right)dz,\end{aligned}$$

where the second line uses the change of variable $z \equiv x/\sigma$ and the symmetry of the Normal CDF . Finally, for θ^* we must also have:

$$\mathbb{E}_{\theta^*}\left[\frac{1}{r}|r>0\right] = \frac{1}{\Phi\left(\mu_r^*/\sigma_r^*\right)\sigma_r^*}\int_0^\infty \frac{1}{z}\phi\left(z-\mu_r^*/\sigma_r^*\right)dz.$$

Since $\mathbb{E}_{\theta'}\left[\frac{1}{r}|r>0\right] = \mathbb{E}_{\theta^*}\left[\frac{1}{r}|r>0\right]$ and $\mu_r'/\sigma_r' = \mu_r^*/\sigma_r^*$, the previous two expressions imply that $\sigma_r' = \sigma_r^*$, arriving to a contradiction. Hence, θ_r is identified.

The next step is to show that $\theta_\delta \equiv (\mu_\delta, \sigma_\delta, \rho)$ are identified. Fix $\theta_r \equiv (\mu_r, \sigma_r)$ and assume, on the contrary, that θ_δ is not identified: there exists θ_δ' and θ_δ^* in Θ such that $\theta_\delta' \neq \theta_\delta^*$ and $\mathcal{P}_{\theta_\delta'}(a=0, A_\mathcal{C}) = \mathcal{P}_{\theta_\delta^*}(a=0, A_\mathcal{C})$. From Proposition 5, this equality implies: $\mathcal{P}_{\theta_\delta'}(a=0, A_\mathcal{C}|r\leq 0) = \mathcal{P}_{\theta_\delta'}(K(0,A_\mathcal{C}|r)\leq\delta|r\leq 0)\,\mathcal{P}_{\theta_r}(r\leq 0) = \mathcal{P}_{\theta_\delta^*}(K(0,A_\mathcal{C}|r)\leq\delta|r\leq 0)\,\mathcal{P}_{\theta_r}(r\leq 0)$. The last equality implies:

$$\int_{-\infty}^0\left\{\Phi\left(\frac{K\left(0,A_\mathcal{C}|r\right)-\mu_{\delta|r}'}{\sigma_{\delta|r}'}\right)-\Phi\left(\frac{K\left(0,A_\mathcal{C}|r\right)-\mu_{\delta|r}^*}{\sigma_{\delta|r}^*}\right)\right\}dr = 0,$$

At this stage, identification of θ_δ from convex menus is analogous to its identification using time menus. We can thus use assumption (b) and the same steps used in the second part of the proof of Proposition 3 to arrive to a contradiction. ■

Proof of Proposition 7: We start discussing menus in which $t^1 \to t^0$, that can be seen as risk problems only. In these menus, the choice is determined by the optimization of the objective function $p^0\frac{(\omega+(1-a)x^0)^{1-r}-\omega^{1-r}}{1-r} + p^1\frac{(\omega+ax^1)^{1-r}-\omega^{1-r}}{1-r}$. Whenever $r \leq 0$, there are two cases. First, if $p^1x^1 \geq p^0x^0$ the choice is $a=1$. Second, if $p^1x^1 < p^0x^0$, there is $r^* < 0$ such that whenever $r \leq r^*$ the choice is $a=1$ and whenever $r > r^*$ the choice is $a=0$. Now, whenever $r>0$, the solution is interior and choices form a continuous mapping from the corner with larger expectation towards the point $\bar{e} = \frac{x^0}{x^0+x^1}$, which is the limit of choices when $r\to\infty$. Thus, we can consider one such problem, say, one such that $p^1x^1 > p^0x^0$. The observed mass of alternative $a=1$, that we denote by q_1, must be equal to $\Phi(\frac{0-\mu_r}{\sigma_r})$. Now, let a be any value in $(\bar{e}, 1)$ and denote by r_a the value above which the optimal choice falls below a. Denote by q' the observed choice probability below a, that must be equal to $1 - \Phi(\frac{r_a-\mu_r}{\sigma_r})$. This allows to obtain parameters μ_r and σ_r: $\sigma_r = \frac{r_a}{1-\Phi^{-1}(q_1)-\Phi^{-1}(q')}$ and $\mu_r = -\sigma_r \Phi^{-1}(q_1) = -\frac{r_a \Phi^{-1}(q_1)}{1-\Phi^{-1}(q_1)-\Phi^{-1}(q')}$

The rest of the parameters can be identified as follows. As commented in the proof of Proposition 4, for any given ω, the use of large payouts is equivalent to use the case $\omega \to 0$ and, in what follows, we assume large payouts. Whenever $p_1 \to p_0$, the probability of selecting options below $\frac{1}{2}$ is the probability that δ is above $K(A_{\mathcal{T}})(1-r)$.[28] One can then reproduce the proof of Proposition 4 replacing the mass of $0_{\mathcal{T}}$ for the cumulative mass below $\frac{1}{2}$. ■

Appendix B. Hybrid Menus

In a hybrid menu, each of the two alternatives corresponds to a two state-contingent lottery, with the safer lottery awarded earlier in time. Formally, $A_{\mathcal{H}} = \{0_{\mathcal{H}}, 1_{\mathcal{H}}\}$ with $0_{\mathcal{H}} = ([p, 1-p; x_1^0, x_2^0], t^0)$ and $1_{\mathcal{H}} = ([p, 1-p; x_1^1, x_2^1], t^1)$ such that $x_1^1 > x_1^0 > x_2^0 > x_2^1$, $p \in (0, 1)$ and $t^0 < t^1$. The analysis of hybrid menus is analogous to that of time menus by conditioning again on parameter r. For any given hybrid menu $A_{\mathcal{T}}$ and any value of r, there exists a menu-dependent constant $K(A_{\mathcal{H}}|r) \in \mathbb{R}$ such that alternative $0_{\mathcal{H}}$ is selected if and only if $\delta \geq K(A_{\mathcal{H}}|r)$.[29] As a result, the choice probability of alternative $0_{\mathcal{H}}$ is:

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

The effect of shifts and spreads of δ are trivially understood from this structure, by applying the logic of Proposition 2. Understanding the effect of r requires some caution, since the ratio of expected utilities is an object that may be difficult to tame. Fortunately, it can be seen that for standard families of monetary utilities (e.g., CRRA and CARA), the threshold $K(A_{\mathcal{H}}|r)$ is decreasing in r as long as $\delta > 0$.

Appendix C. Numerical Evaluation of Choice Probabilities

DMPL. Given a value for parameter vector $\theta \in \Theta$, computation of the log-likelihood defined in Section 4 requires computing $\mathcal{P}_\theta(0_m, A_m)$ for each menu A_m in the dataset $\mathcal{O}$. In general, this requires evaluating a double integral numerically. Given the structure of the RDEU model, this can be done efficiently using Quasi-Monte Carlo (QMC)

[28]Note that unlike in the case of Proposition 4, the probability obtained here is exact.

[29]It is easy to see that $K(A_{\mathcal{H}}|r) = \frac{1}{t^1-t^2} \log\left[\frac{pu_r(x_1^0)+(1-p)u_r(x_2^0)}{pu_r(x_1^1)+(1-p)u_r(x_2^1)}\right]$.

methods: begin by discretizing the support of f_θ it in N_{QMC} nodes $\{r_k, \delta_k\}_{k=1}^{N_{QMC}}$ using low-discrepancy sequences. Let $\mathcal{I}(0_m, A_m|r_k, \delta_k)$ denote an indicator function that takes value of one when $DEU_{r_k,\delta_k}(0_m) > DEU_{r_k,\delta_k}(1_m)$, and zero otherwise. For large enouh N_{QMC}, we have

$$\mathcal{P}_\theta(0_m, A_m) \approx \frac{V}{N_{QMC}} \sum_{k=1}^{N_{QMC}} \mathcal{I}(0_m, A_m|r_k, \delta_k) f_\theta(r_k, \delta_k),$$

where $V \equiv \int_r \int_\delta dr d\delta = (\overline{r} - \underline{r})(\overline{\delta} - \underline{\delta})$ is a normalization constant. We can control the accuracy of the approximation by increasing the number of nodes N_{QMC}. Importantly, the indicator function $\mathcal{I}(0_m, A_m|r_k, \delta_k)$ is independent of θ. It follows that, to compute the maximum-likelihood estimator of θ, this indicator function needs to be computed only once before starting the search of the maximizer, reducing dramatically the estimation time.

We can also use the results in the paper, together with the assumption that f follows a bivariate normal distribution, to improve the estimation algorithm further. For risk menus, Proposition 1 implies $\mathcal{P}_\theta(0_m, A_{\mathcal{R},m}) = 1 - \Phi((K(A_{\mathcal{R},m}) - \mu_r)/\sigma_r)$. Computation of $K(A_{\mathcal{R},m})$ requires solving a non-linear equation numerically for each risk menu. However, these thresholds are independent of θ so we only need to compute them once before estimation. For time menus, Proposition 2 simplifies the double integral characterizing $\mathcal{P}_\theta(0_m, A_{\mathcal{T},m})$ into the following following single-valued integral:

$$\mathcal{P}_\theta(0_m, A_{\mathcal{T},m}) = 1 - \int_r \Phi\left(\frac{K(A_{\mathcal{T},m}|r) - \mu_{\delta|r}}{\sigma_{\delta|r}}\right) \phi\left(\frac{r - \mu_r}{\sigma_r}\right) dr,$$

with $\mu_{\delta|r} \equiv \mu_\delta + \rho \frac{\sigma_\delta}{\sigma_r}(r - \mu_r)$ and $\sigma_{\delta|r} \equiv \sigma_\delta \sqrt{1 - \rho^2}$. This simpler integral can also be evaluated numerically using QMC methods, as discussed before.

Convex Budgets. Computation of the log-likelihood function with convex budgets, as defined in Section 6, is computationally more demanding as it requires evaluating $\mathcal{P}_\theta(a \in \alpha^s, A_m)$ for each of the M menus in the dataset and each of the S options in which the choice set is discretized. As in the iid-additive RUM, we can proceed by rounding each observed allocation a to the midpoint of the option α^s for which $a \in \alpha^s$. This results in S possible observed allocations in the data: $\overline{\alpha}_1 = 0$, $\overline{\alpha}_2 = (a_2 + a_3)/2$, …, $\overline{\alpha}_{S-1} = (a_{S-1} + a_S)/2$ and $\overline{\alpha}_S = 1$. Let $\mathcal{I}(\overline{\alpha}, A_m|r_k, \delta_k)$ denote an indicator function that takes value of 1 when $DEU_{r_k,\delta_k}(\overline{\alpha}) \geq DEU_{r_k,\delta_k}(\overline{\alpha}^*)$ in menu A_m for all $\overline{\alpha}^* \in \{\overline{\alpha}_1, \ldots, \overline{\alpha}_S\}$, and zero otherwise. We can then use the numerical approximation $\mathcal{P}_\theta(a \in \alpha^s, A_m) \approx (V/N_{QMC}) \sum_{k=1}^{N_{QMC}} \mathcal{I}(\overline{\alpha}, A_m|r_k, \delta_k) f_\theta(r_k, \delta_k)$. As in the DMPL case, the indicator function $\mathcal{I}(\overline{\alpha}, A_m|r_k, \delta_k)$ is independent of θ and can be pre-computed before maximization of the log-likelihood function. This method allows for flexible specifications of f and can be easily extended to more general models, as illustrated in Appendix D.

Alternatively, we can exploit the results from Proposition 5 and the assumption that f is normal to compute $\mathcal{P}_\theta([a_s, a_{s+1}], A_m)$ as the single-variable integral:

$$\int_r \left[\Phi\left(\frac{K(A_{s,\mathcal{C}}|r) - \mu_{\delta|r}}{\sigma_{\delta|r}}\right) - \Phi\left(\frac{K(A_{s+1,\mathcal{C}}|r) - \mu_{\delta|r}}{\sigma_{\delta|r}}\right)\right] \phi\left(\frac{r-\mu_r}{\sigma_r}\right) dr,$$

where the thresholds $K(A_{s,\mathcal{C}}|r)$ are defined in Section 5 and are independent of θ.

Appendix D. Extensions: Behavioral Model and Alternative Probability Distributions

The methods introduced in this paper can be extended to allow for additional behavioral features and alternative distributions for the random parameters. We illustrate this in this section using convex budget data from the experimental design in Andreoni and Sprenger (2012a). This dataset is similar to the one used in Section 5 and features 97 subjects facing 45 convex menus with certain payoffs. The main difference with the data used for the baseline analysis in the main text is that some of the menus feature payoffs in the present, which allows estimation of present bias in discounting.

Table 6 summarizes the results of this exercise. To make comparison across models with different distributional assumptions feasible, we report the median and interquantile range (IQR) of the estimated distributions. The second column of Table 6 summarizes the baseline estimates obtained using the RDEU model following the procedure used in Section 6. The estimated curvature of the utility function is statistically zero, similar to the results obtained by Andreoni and Sprenger (2012b), which could be due to the lack of variation in payoff probabilities in this dataset, indicating that the curvature captures only intertemporal substitution. The estimated median of the annual discount rate is small and statistically close to zero. Nevertheless, the model estimates a large degree of heterogeneity in both parameters.

The third column of Table 6 shows the results using Quasi-Monte Carlo (QMC) methods, as discussed in Section C of this Appendix. The results are virtually identical to those obtained using the baseline algorithm developed in the paper. Nevertheless, it takes four times longer to estimate the model using QMC methods, confirming the benefits of exploiting the economic structure of the problem. Despite this, the QMC method is useful for estimating the model with alternative distributions and behavioral features, as discussed below.

The fourth column of Table 6 shows the results for a "constrained" version of the RDE model where δ follows a normal distribution truncated at zero, ruling out the possibility of preference for the future. In this case, we use a Gaussian copula to allow for correlation between r and δ. The estimated median of the annual discount rate is larger and statistically different from zero. The median curvature of the utility function remains close to zero, and the IQR of both parameters is now lower. As expected, restricting the domain of δ reduces the fit to the data, as reflected in the resulting log-likelihood.

The last two columns of Table 6 show the results for the model extended to allow present bias in discounting so that the discount factor in the model is $\beta e^{-\delta}$ when $t^0 = 0$, and $e^{-\delta}$ otherwise. The first of the two columns shows the results for an "unconstrained" model that assumes parameters r, δ and β follow a multivariate normal distribution with an arbitrary correlation matrix. The results in this case are very similar to those obtained in the RDEU model, indicating a low degree of present bias. The second column shows the results for a "constrained" model where r follows a normal distribution, δ follows a normal distribution truncated at zero, and β follows a beta distribution with support on the unit interval. The estimated distributions of r and δ are similar to those obtained for the constrained RDEU model, but now the median present bias is statistically different from 1. Nevertheless, the improvement in fit is relatively low compared to the constrained RDEU model.

Appendix E. Baseline wealth

We now briefly comment on the role of ω with CRRA monetary utilities. It is immediate to see that in risk menus $A_{\mathcal{R}}$ such that $K(A_{\mathcal{R}}) \neq 0$, $K(A_{\mathcal{R}})$ is strictly increasing (resp., decreasing) in ω whenever $K(A_{\mathcal{R}}) > 0$ (resp., $K(A_{\mathcal{R}}) < 0$).[30] Consequently, ceteris paribus, the alternative with larger expected value will be chosen more often. In time menus $A_{\mathcal{T}}$, every threshold $K(A_{\mathcal{T}}|r)$ converges monotonically to the constant $K(A_{\mathcal{T}}|0)$ as ω increases. The conditional behavior of every r becomes more aligned with the conditional choices of $r = 0$. That is, ceteris paribus, the more risk-averse

[30]In the degenerate case where the expected values of both lotteries coincide, we obviously have $K(A_{\mathcal{R}}) = 0$ for all levels of ω.

(resp., lover) individuals will choose more often the present (resp., future) option. Similarly, in convex menus $A_{\mathcal{C}}$, every threshold map $K(a, A_{\mathcal{C}}|r)$ converges monotonically to the constant map $K(a, A_{\mathcal{C}}|0)$ as ω increases. The conditional behavior of every r becomes more aligned with the conditional choices of $r = 0$ (with interior solutions vanishing).

Since in actual practice it is often assumed zero levels of background wealth, it is interesting to discuss theoretically the limit model when the baseline wealth tends to zero. From the previous discussion, we know that this limit case would create the best conditions for parameter r kicking in all menus. Interestingly, for the case of time menus $A_{\mathcal{T}}$, as discussed in the proof of Proposition 3, the conditional threshold map becomes piece-wise linear: (a) when $r < 1$, $K(A_{\mathcal{T}}|r) = \frac{\log \frac{x^1}{x^0}}{t^1 - t^0}(1-r) \equiv K(A_{\mathcal{T}})(1-r) > 0$, where $K(A_{\mathcal{T}})$ is a menu-dependent constant, and (b) when $r \geq 1$, $K(A_{\mathcal{T}}|r)$ becomes null. Hence, for parameters (r, δ), with $r < 1$, the earlier option $0_{\mathcal{T}}$ is preferred to the later option for such parameters if and only if $\frac{\delta}{1-r} \geq K(A_{\mathcal{T}})$. That is, the expression $\frac{\delta}{1-r}$ represents a simple correction of δ based on the risk parameter r that captures completely time considerations. In other words, the behavior of $DEU_{r,\delta}$ is equivalent to the behavior of $DEU_{0,\frac{\delta}{1-r}}$, and if the analyst is willing to entertain the idea that risk aversion above 1 is not crucial or that risk aversion and delay aversion are somewhat independent phenomena for standard values, independent distributions of r and $\frac{\delta}{1-r}$ can be considered. Importantly, behavior for $r \geq 1$ becomes extreme when wealth is negligible, as alternative $0_{\mathcal{T}}$ is always preferred.

To see the role of baseline wealth in the empirical applications studied in the paper, Table 5 compares the estimates under the baseline choice of ω with those obtained by setting ω to a positive value close to zero. This exercise confirms the previous theoretical discussion: the estimated average and standard deviation of risk aversion falls with the value of ω, and the correlation between r and δ increases. Nevertheless, these changes are quantitativelly small. Notice also that the estimated marginal distribution of δ remains practically unchanged after across values of ω.

References

[1] Aguiar, V.H. and N. Kashaev (2021). "Stochastic Revealed Preferences with Measurement Error." *Review of Economic Studies* 88(4):2042–2093.

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

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

[4] Andersen, S., G.W. Harrison, M.I. Lau, and E.E. Rutstrom (2008). "Eliciting Risk and Time Preferences." *Econometrica* 76(3):583–618.

[5] Andersen, S., G.W. Harrison, M.I. Lau, and E.E. Rutstrom (2014). "Discounting Behavior: A Reconsideration." *European Economic Review* 71:15–33.

[6] Andreoni, J. and C. Sprenger (2012a). "Estimating Time Preferences from Convex Budgets." *American Economic Review*, 102(7):3333–3356.

[7] Andreoni, J. and C. Sprenger (2012b). "Risk Preferences Are Not Time Preferences." *American Economic Review*, 102(7):3357–3376.

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

[9] Augenblick, N., M. Niederle and C. Sprenger. (2015) "Working over Time: Dynamic Inconsistency in Real Effort Tasks." *Quarterly Journal of Economics*, 130(3):1067–1115.

[10] Barseghyan, L., F. Molinari and M. Thirkettle (2021). "Discrete Choice under Risk with Limited Consideration," *American Economic Review*, 111(6):1972–2006.

[11] Baucells, M. and F. H. Heukamp (2012). "Probability and Time Tradeoff." *Management Science*, 58(4):831–842.

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

[13] Burks, S.V, J.P. Carpenter, L. Goette and A. Rustichini (2009). "Cognitive Skills Affect Economic Preferences, Strategic Behavior, and Job Attachment." *Proceedings of the National Academy of Sciences*, 106(19):7745–7750.

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

[15] Cattaneo, M., X. Ma, Y. Masatlioglu and E. Suleymanov (2020). "A Random Attention Model," *Journal of Political Economy*, 128(7):2796–2836.

[16] Cheung, S. (2015). "Risk Preferences Are Not Time Preferences: On the Elicitation of Time Preference under Conditions of Risk: Comment." *American Economic Review*, 105 (7): 2242-60.

[17] Choi, S., R. Fisman, D. Gale, and S. Kariv (2007). "Consistency and Heterogeneity of Individual Behavior under Uncertainty," *American Economic Review*, 97(5):1921–1938.

[18] Coble, K.H. and J. Lusk (2010). "At the Nexus of Risk and Time Preferences: An Experimental Investigation." *Journal of Risk and Uncertainty*, 41(1):67–79.

[19] Dardanoni, V., P. Manzini, M. Mariotti, and C. Tyson (2020). "Inferring Cognitive Heterogeneity from Aggregate Choices." *Econometrica*, 88(3):1269–1296.

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

[21] Dohmen T., A. Falk, D. Huffman and U. Sunde (2010). "Are Risk Aversion and Impatience Related to Cognitive Ability?" *American Economic Review*, 100:1238–1260.

[22] Falk, A., A. Becker, T. Dohmen, B. Enke, D. Huffman and U. Sunde (2018). "Global Evidence on Economic Preferences." *Quarterly Journal of Economics*, 133(4):1645–1692.

[23] Fisman, R., S. Kariv, and D. Markovits (2007). "Individual Preferences for Giving." *American Economic Review*, 97(5):1858–1876.

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

[25] Harrison, G., M. Lau and H. Yoo (2020). "Risk Attitudes, Sample Selection and Attrition in a Longitudinal Field Experiment." *The Review of Economics and Statistics*, 102(3):552–568.

[26] Holt, C. and S. Laury (2002). "Risk Aversion and Incentive Effects ." *American Economic Review*, 92 (5): 1644-1655.

[27] Jagelka, T. (2021). "Are Economists' Preferences Psychologists' Personality Traits?" Mimeo.

[28] Kim, H.B., S. Choi, B. Kim and C. Pop-Eleches (2018). "The Role of Education Interventions in Improving Economic Rationality." *Science*, 362:83–86.

[29] Tanaka, T., C.F. Camerer and Q. Nguyen (2010). "Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam." *American Economic Review*, 100(1):557–571.

[30] Wilcox, N. (2011). "Stochastically More Risk Averse: A Contextual Theory of Stochastic Discrete Choice under Risk." *Journal of Econometrics*, 162(1):89–104.

Table 1. Risk and Time Preferences at the Population Level: AHLR

	RDEU	**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 population 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. Risk and Time Preferences at the Individual Level: AHLR

	Risk Aversion Coefficient μ_r				**Discount Rate μ_δ**			
Moment	**RDEU**	**LUCE**	**WILCOX**	**SPE**	**RDEU**	**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) by 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. Risk and Time Preferences at the Population Level: AS

	RDEU	**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 population 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.

Table 4. Volatility of Risk and Time Preferences at the Individual Level: AHLR

	Risk Aversion Coefficient σ_r				Discount Rate σ_δ			
Moment	**RDEU**	**LUCE**	**WILCOX**	**SPE**	**RDEU**	**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) by 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 5. Estimated Risk and Time Preferences when $\omega \approx 0$

	DMPL-AHLR		**CB-AS**	
	Baseline	$\omega \approx 0$	*Baseline*	$\omega \approx 0$
μ_r	0.781	0.681	0.207	0.095
	[0.053]	[0.032]	[0.062]	[0.045]
σ_r	0.895	0.768	0.752	0.562
	[0.049]	[0.039]	[0.079]	[0.054]
μ_δ	0.125	0.102	0.339	0.383
	[0.008]	[0.007]	[0.108]	[0.110]
σ_δ	0.125	0.116	1.805	1.821
	[0.010]	[0.007]	[0.124]	[0.125]
ρ	−0.958	−0.999	−0.164	−0.202
	[0.016]	[0.001]	[0.053]	[0.052]

NOTES.- The table reports the risk aversion coefficient and annual discount rate at the population level estimated by the RDEU model under two different assumptions about the value of integrated wealth ω.

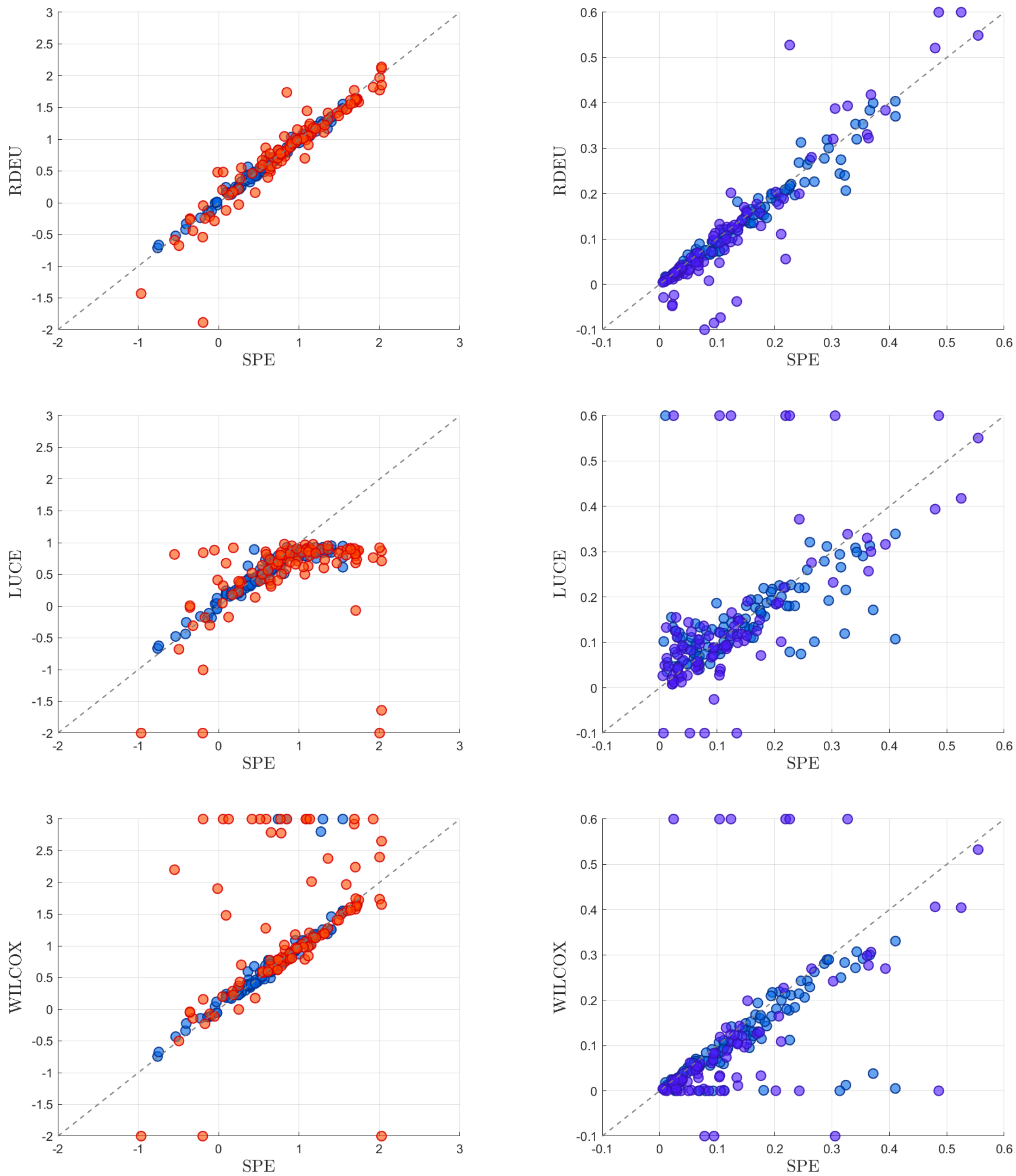
Table 6. Estimated Risk and Time Preferences in Andreoni and Sprenger (2012a)

	RDEU			**Hyperbolic Discounting**	
	Baseline	*QMC*	*Constrained*	*Unconstrained*	*Constrained*
Median (r)	-0.051 [0.084]	-0.050 [0.080]	-0.037 [0.039]	-0.039 [0.071]	-0.038 [0.038]
IQR (r)	0.718 [0.166]	0.715 [0.152]	0.347 [0.072]	0.665 [0.122]	0.348 [0.070]
Median (δ)	0.043 [0.157]	0.037 [0.151]	0.504 [0.058]	0.072 [0.148]	0.490 [0.052]
IQR (δ)	2.317 [0.335]	2.310 [0.305]	0.784 [0.087]	2.217 [0.271]	0.764 [0.079]
Median (β)	–	–	–	0.991 [0.004]	0.898 [0.008]
IQR (β)	–	–	–	0.004 [0.005]	0.247 [0.056]
Cor (r, δ)	-0.306 [0.071]	-0.302 [0.068]	0.020 [0.056]	-0.284 [0.067]	0.017 [0.060]
Cor (r, β)	–	–	–	-0.002 [2.605]	-0.078 [0.459]
Cor (δ, β)	–	–	–	-0.001 [0.479]	0.134 [0.356]
$\mathcal{L}$	-1.709	-1.709	-1.835	-1.709	-1.833

NOTES.- The table reports estimated moments of the distributions of risk aversion (r), discounting (δ), and present bias (β) estimated at the population level. The second column reports the results obtained using the RDEU model with same methodology and assumptions as in Section 6. The third column shows the results using Quasi-Monte Carlo methods (QMC), as discussed in Appendix C. The fourth column shows the results using QMC and assumes that δ follows a normal distribution truncated at zero. The fourth column shows the results for a model extended to allow for present bias under the assumption that the (r, δ, β) follow a multivariate normal distribution. The last column shows the results assuming that δ follows a normal distribution truncated at zero and β a beta distribution. Standard errors for each MLE are shown in brackets and are clustered at the individual level.

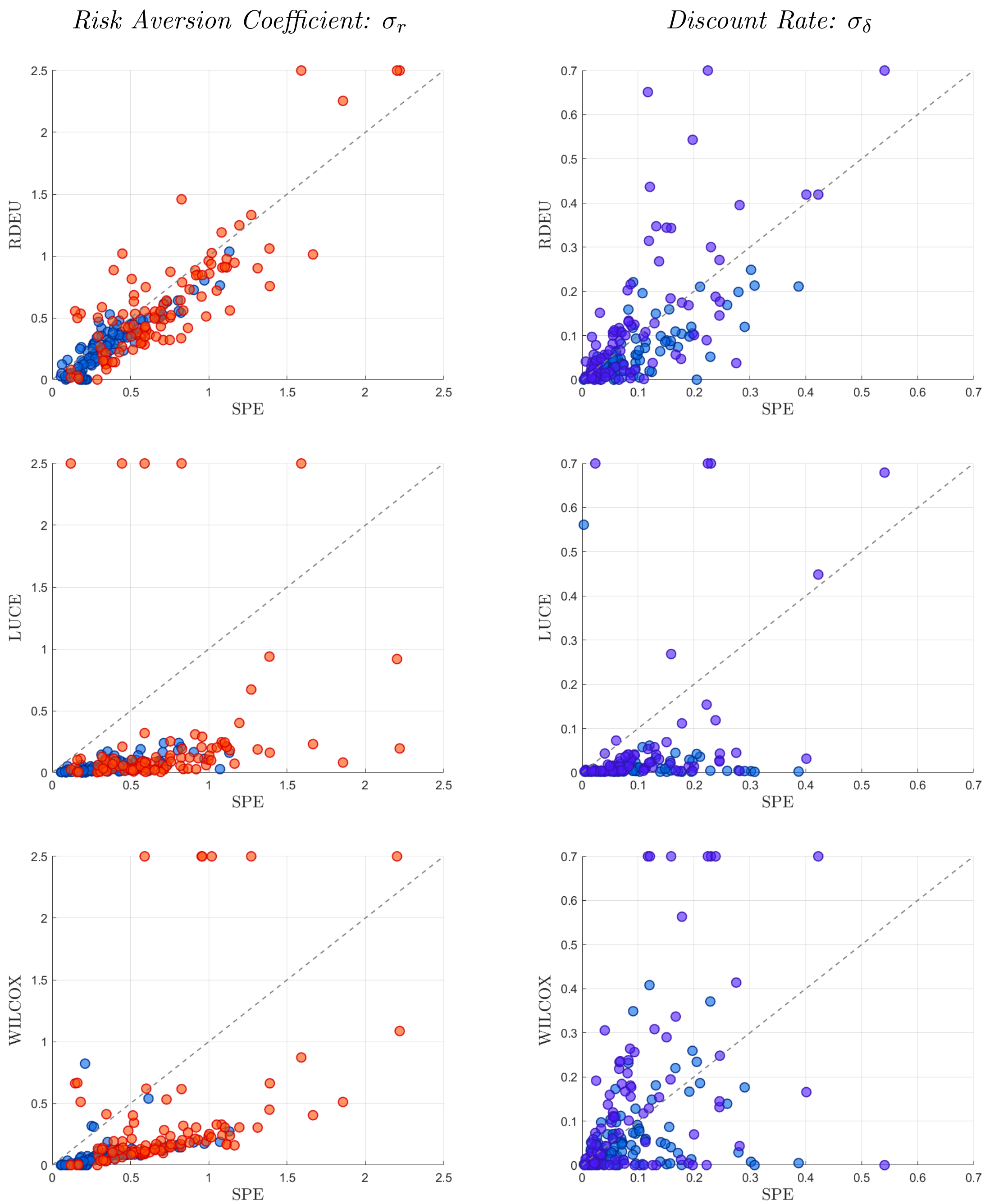
Figure 1. Risk and Time Preferences at the Individual Level AHLR

Risk Aversion Coefficient μ_r *Discount Rate* μ_δ



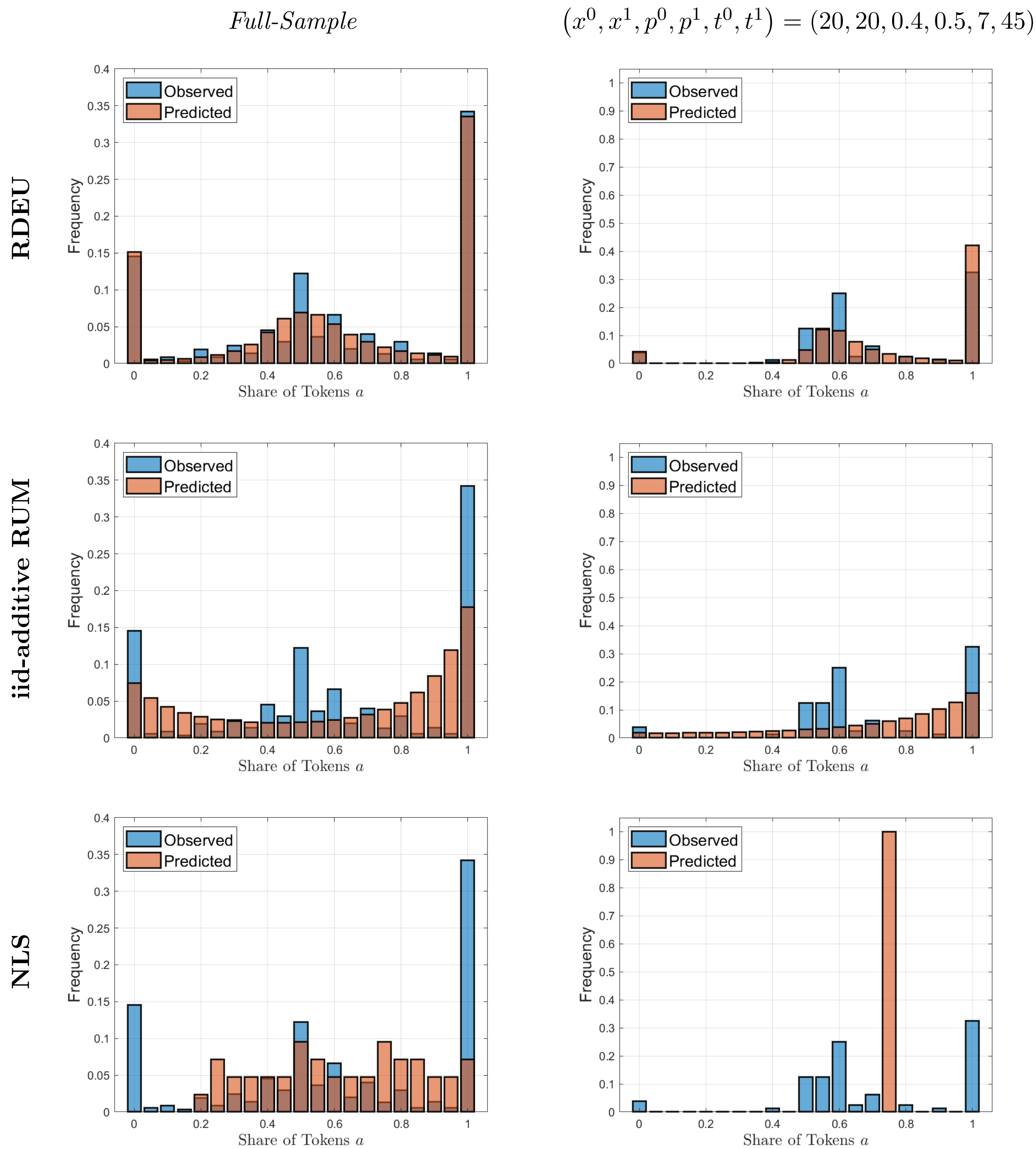
NOTES. The figure shows the estimated average risk aversion coefficient μ_r (first column) and annual discount rate μ_δ (second column) for each individual in the double-multiple price list data from Andersen et al. (2008). Each dot shows a subject's estimate using the corresponding structural model 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



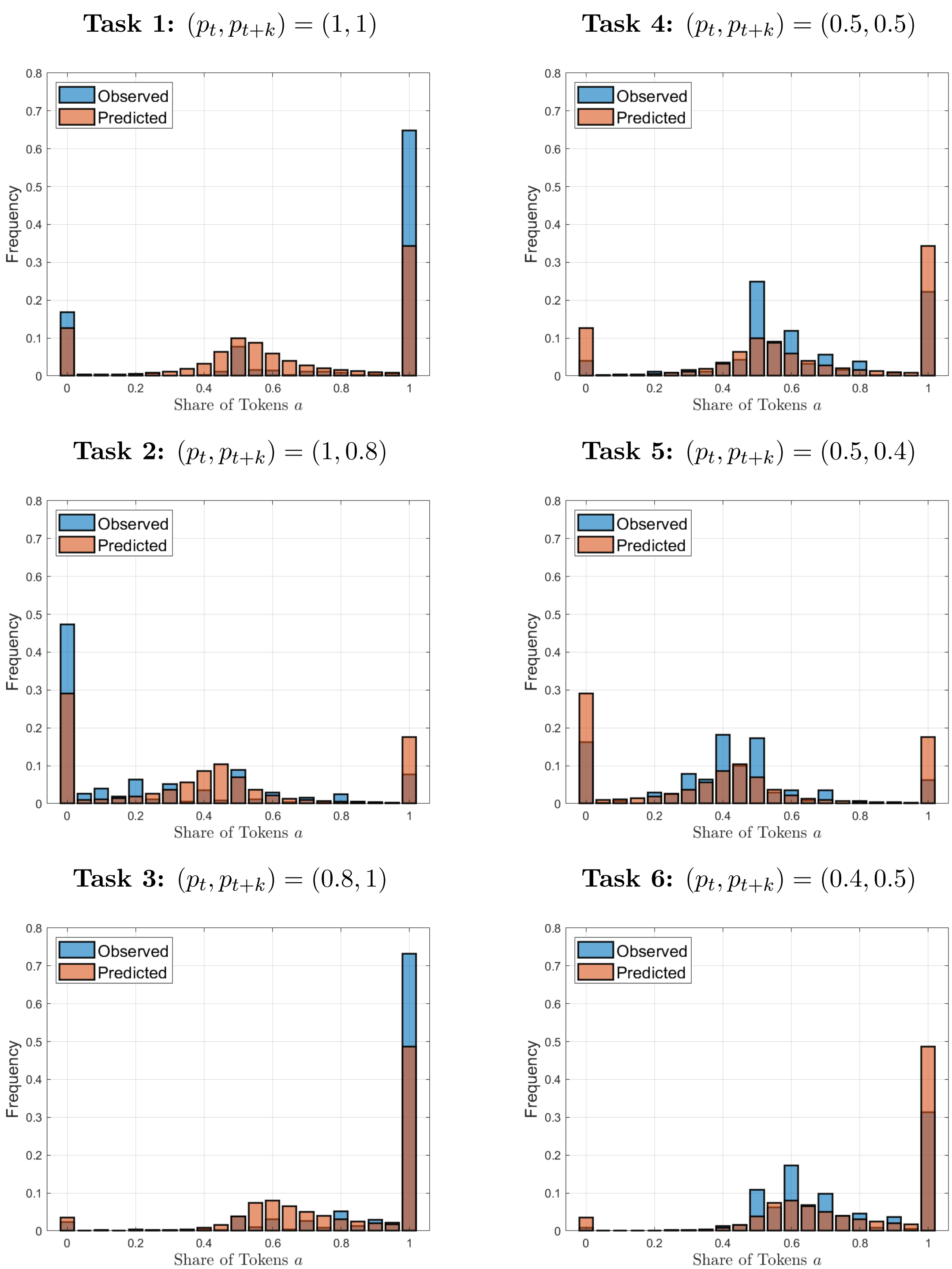
NOTES. The figure shows the estimated standard deviation of the risk aversion coefficient σ_r (first column) and annual discount rate σ_δ (second column) for each individual in the double-multiple price list data from Andersen et al. (2008). Each dot shows a subject's estimate using the corresponding 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 RDEU model using the population estimates reported in Table 3.