RANDOM DISCOUNTED EXPECTED UTILITY

JOSE APESTEGUIA[†], MIGUEL A. BALLESTER[‡], AND ANGELO GUTIERREZ-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 that it has solid comparative statics and also demonstrate its computational convenience. Finally, we illustrate the empirical implementation of this model using two distinct experimental datasets.

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 attitude 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 datasets. Overall, we provide a well-founded and convenient novel framework for the analysis of heterogeneous risk and time preferences.

Date: August, 2022.

[*] We thank Daniel Benjamin, Holger Gerhardt, and Jesse Shapiro for helpful comments, and Rafael Suchy for outstanding research assistance. Financial support by Spanish Ministry of Science (PGC2018-098949-B-I00), Balliol College, "la Caixa" Foundation Research Grants on Socioeconomic Well-being (CG-2017-01) and "la Caixa" Severo-Ochoa International Doctoral Fellowship is gratefully acknowledged.

[†] ICREA, Universitat Pompeu Fabra and Barcelona GSE. E-mail: `jose.apesteguia@upf.edu`.

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

[§] Universitat Pompeu Fabra and Barcelona GSE. E-mail: `angelo.gutierrez@upf.edu`.

Our stochastic model is based on a probability distribution over a given collection of ordinal 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 and convex budgets. 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 use of the random utility model with multiple preference parameters has often been criticized as being too complex, based on the argument that the computation of choice probabilities involves multiple integration. In the case of discounted expected utility, this would demand the integration over two variables: the discounting factor and the curvature of the monetary utility function. We show, however, that this curse of dimensionality does not apply. We prove that, given any curvature of the monetary function, 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, thereby rendering the model both 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 practice in the literature, we consider CRRA monetary utility functions, which are determined by the parameter r describing the curvature, as the family of utilities. Importantly, 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, since the literature contemplates a variety of formulations, which have indeed, on occasions, provoked perverse implications. Hence every utility is characterized by a pair of parameters (r, δ), where δ is the discount factor. In addition, we consider a bivariate normal distribution over r and a transformation of δ over the reals as the distributional source of variability. This added parametric structure further accentuates the convenience of the model. Basically, under the normality assumption, the choice computation requires the evaluation of conditional and marginal distributions of a bivariate normal which are themselves normal distributions, thus making the computational task routine. Moreover, shifts and spreads in the probability distribution are the result of variations in the first two moments of the distribution.

Once established the theoretical grounds of the model, we illustrate it implementing a structural estimation exercise using the two major exemplars of the type of elicitation mechanisms considered, Andersen et al. (2008) and Andreoni and Sprenger (2012).[1] Following practice in the literature, we consider our parametric version of the model. The estimation results neatly show that the predictions of the RDEU model fit the observed choice data well. The model succeeds, in both experimental settings, to capture the widely diverse menu-dependent choice patterns, which are very difficult to account for with other empirical strategies.

We close this introduction with a brief review of the literature. Most relevant are the two main empirical strategies used in the experimental literature on risk and time preferences. One approach, based upon non-linear least squares, assumes a unique discounted expected utility and introduces randomness by perturbing the first order condition of a constrained utility-maximization problem. Hence, in this approach, the randomness lacks a behavioral foundation, in the sense that it does not explicitly connect heterogeneity of choice with heterogeneity of preference. Another approach found

[1]All our estimation programs are available for use at https://github.com/agutieda/random-discounted-expected-utility.

in the literature uses an iid additive random utility model, assuming variability at the moment of choice by perturbing an underlying utility function. The implementation of this approach with standard representations of discounted expected utility could lead to paradoxical predictions and perverse comparative statics properties (Apesteguia and Ballester, 2018) and thus hinder a thorough understanding of risk and time preferences. In contrast to these approaches, the present paper studies the RDEU model which establishes an explicit linkage between preference and choice heterogeneity, and, for the first time, proves the robustness of this model's comparative statics. Its implementation is also shown to be computationally convenient, as illustrated empirically with two distinct datasets. Finally, 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. 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. In the settings that we analyze in this paper, decision problems are either binary menus, i.e., $A = \{0, 1\}$, or continuous intervals, i.e., $A = [0, 1]$. In both cases, every alternative $a \in A$ corresponds to a finite collection of dated lotteries that, in case a is selected, will be received by the decision-maker. That is, a is a sequence of the form $\{(l^j, t^j)\}_{j=1}^{J}$, where it is assumed that $t^j \neq t^{j'}$. In the binary setting, each alternative corresponds to a unique dated lottery, while in the continuous setting, each alternative corresponds to a sequence of two dated lotteries.[2]

We start by introducing the most commonly-used deterministic model of behavior for the study of risk and time preferences, discounted expected utility (DEU). We construct it using 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

[2] Sections 3 and 4 define formally the two main experimental settings.

strictly ordered by concavity, i.e., $r < r'$ means that u'_r is strictly more concave than u_r. 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 $\delta \in (0,1)$. Hence, given parameters $(r,\delta) \in \mathbb{R} \times (0,1)$, the DEU evaluation of alternative $a \in A$ is:

$$DEU_{r,\delta}(a) = \sum_{j=1}^{J} \delta^{t^j} \sum_{n=1}^{N_j} p_n^j u_r(\omega + x_n^j).$$

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 $\mathbb{R} \times (0,1)$, capturing the prevalence of each possible DEU preference. At the moment of choice, parameters (r,δ) are realized with probability $f(r,\delta)$, and the alternative a that maximizes $DEU_{r,\delta}$ within menu A is selected.[3] We denote by $\Gamma(a,A) \subseteq \mathbb{R} \times (0,1)$ the collection of all parameter combinations (r,δ) for which the maximizer of $DEU_{r,\delta}$ within menu A is smaller or equal than a. The f-measure of $\Gamma(a,A)$ describes the cumulative choice probability of alternatives in $[0,a]$:

$$\mathcal{P}_f(a,A) = \int_{\Gamma(a,A)} f(r,\delta)\mathrm{d}(r,\delta).$$

In the binary case $\mathcal{P}_f(0,A)$ is the only relevant probability, corresponding to the choice probability of option 0. In the continuous case, $\mathcal{P}_f(a,A)$ corresponds to the CDF of choices at alternative a, i.e. the probability of choosing an alternative equal or smaller than a.

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 and (bivariate) normal distributions. We also use these specifications in our empirical exercise. Let us elaborate.

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

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.[4] 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 $-\frac{\omega^{1-r}}{1-r}$ is required. This is a crucial feature for the solid treatment of time preferences. Notice that, by always working with positive utilities, we can consider discounting in standard terms, i.e., as a parameter in $(0, 1)$. If some utility functions were negative, the use of discount factors in $(0, 1)$ would create for them a non-sensical systematic preference for delay. One could potentially separate the analysis of positive and negative utility functions, but this would complicate the analysis and the interpretation of results. Finally, the chosen constant also guarantees that $u_r^{crra}(\omega + 0) = 0$ holds, which facilitates the analysis of lotteries involving null prizes, and that the standard continuity property $\lim_{r \to 1} u_r^{crra}(\omega + x) = u_1^{crra}(\omega + x)$ is also satisfied.[5]

With respect to f, we consider bivariate normals. To do so, we work with parameter $r \in \mathbb{R}$ and with the bijection of δ into the reals given by $\gamma = \log \frac{\delta}{1-\delta} \in \mathbb{R}$.[6] Hence, distribution f is fully described by the set of distributional parameters $(\mu_r, \sigma_r, \mu_\gamma, \sigma_\gamma, \rho)$, where μ_x and σ_x are the corresponding median and standard deviation of parameter x, and ρ is the covariance.[7]

[4] Often times, the used CRRA formulations are undefined for either certain payoff or parameter ranges. In addition, they may imply negative evaluations, implying illogical interactions with the discount factor. For clarity, we expand on these issues in the text.

[5] A discussion on the role of wealth ω in CRRA utilities can be read in Appendix C. In Appendix F we also consider the CARA family of monetary utilities, that is independent of wealth considerations.

[6] That is, $\delta = \frac{1}{1+e^{-\gamma}}$. Given the strictly increasing bijection between δ and γ, we often times present the effects of changes in the underlying parameters of γ as *changes in* δ.

[7] Note that given the normality assumption the mean coincides with the median.

3. Double Multiple Price Lists

One of the most prominent settings in the experimental literature involves binary menus. The simplest version involves the study of risk and time decision problems separately, in the form of risk multiple price lists (MPLs) and time MPLs, i.e., the so-called double MPLs (see Andersen et al. (2008), Burks et al. (2009), Dohmen et al. (2010), Tanaka et al. (2010), Benjamin et al. (2013), Falk et al. (2018) or Jagelka (2021)).[8] As we now discuss, the structure of these decision problems facilities (i) the computation of choice probabilities, (ii) the understanding of the comparative statics and (iii) the parameter estimation.

3.1. **Risk Menus.** In these decision problems, each of the two alternatives corresponds to a single two-state contingent lottery, with prizes awarded in the present. Formally, $A_{\mathcal{R}} = \{0, 1\}$, with $0 = ([p, 1-p; x_1^0, x_2^0], 0)$, $1 = ([p, 1-p; x_1^1, x_2^1], 0)$ such that $x_1^1 > x_1^0 > x_2^0 > x_2^1$ and $p \in (0, 1)$.[9] Since 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; lottery 0, the safer lottery, is chosen if and only if parameter r is above a threshold $r^0_{A_{\mathcal{R}}} \in \mathbb{R}$, i.e., $\Gamma(0, A_{\mathcal{R}}) = \{(r, \delta) : r \geq r^0_{A_{\mathcal{R}}}\}$. The choice probability of alternative 0, being the f-measure of such rectangular set, can be conveniently computed by using the marginal CDF of r, that we denote by 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.[10]

[8] Ahlbrecht and Weber (1997), Coble and Lusk (2010), Baucells and Heukamp (2012) and Cheung (2015) use MPLs involving a third, hybrid, type of binary menus with both risk and time considerations active. We discuss this type of menus in Appendix B.

[9] 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.

[10] 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 $r^0_{A_{\mathcal{R}}}$.

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 x, with domain in an open interval, and denote by $\text{med}(F)$ the median of distribution F. Then, we say that: (i) F dominates G if $F(x) < G(x)$ holds for all values of x and (ii) F expands G if $\text{med}(F) = \text{med}(G)$, $F(x) > G(x)$ whenever $x < \text{med}(F)$ and $F(x) < G(x)$ whenever $x > \text{med}(F)$.

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

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

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

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

3.2. **Time Menus.** In time decision problems, each of the two alternatives is composed by a unique dated degenerate lottery. Formally, $A_{\mathcal{T}} = \{0, 1\}$, with $0 = ([1; x^0], t^0)$ and $1 = ([1; x^1], t^1)$ such that $t^0 < t^1$ and $x^0 < x^1$.[11] 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 $\omega = 100$. Consider two DEU–CRRA individuals with parameters $(r_1, \delta_1) = (.95, .91)$, $(r_2, \delta_2) = (0, .9)$, and $A_{\mathcal{T}} = \{([1; 71.5], 0), ([1; 80], 1)\}$. Although $\delta_1 > \delta_2$, which may suggest that individual 1 is more patient, it is immediate to see that she is indeed the only one that prefers the earlier prize.*

The joint consideration of both parameters is then required to fully understand the predictions of DEU, and consequently, RDEU behavior. 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. Having fixed the value of r, the earlier alternative 0 is selected if and only if δ is below a threshold $\delta^0_{A_{\mathcal{T}}}(r) \in (0, 1)$, i.e., $\Gamma(0, A_{\mathcal{T}}) = \{(r, \delta) : \delta \leq \delta^0_{A_{\mathcal{T}}}(r)\}$. As a result, the choice probability of alternative 0 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 $\delta^0_{A_{\mathcal{T}}}(r)$, 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 $\{\delta^0_{A_{\mathcal{T}}}(r)\}_{r \in \mathbb{R}}$ are strictly increasing in r, and constitute a bijection from $\mathbb{R}$

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

to $(0,1)$, which can thus be inverted.[12] Hence, comparative statics of shifts are immediate, as keeping constant the marginal distribution of r (respectively, δ), and shifting downwards (respectively, upwards) the conditional distributions of δ (respectively, r) guarantee an increase in the choice probability of the earlier alternative 0. Second, with respect to spreads, we can again show that keeping constant the marginal distribution of one parameter, an expansion of the conditional distributions of the other always creates more stochasticity.[13]

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

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

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

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

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.

[12] In other words, conditioning on δ also renders ordered choices over parameter r. We denote the inverted map as $\{r^0_{A_{\mathcal{T}}}(\delta)\}_{\delta\in(0,1)}$.

[13] As in the case of risk, the median of each conditional distribution $F_{\delta|r}$ must be different to the corresponding threshold $\delta^r_{A_{\mathcal{T}}}(0)$ when expansions of δ are considered, with an analogous expression for the case of r.

Similarly, we can read Proposition 2 from the parametric point of view. With CRRA, the threshold map can be written as $\delta^0_{A_\mathcal{T}}(r) = \left[\frac{(\omega+x^0)^{1-r}-\omega^{1-r}}{(\omega+x^1)^{1-r}-\omega^{1-r}}\right]^{\frac{1}{t^1-t^0}}$. With the bivariate normal, all conditionals $F_{\gamma|r}$ are also normal, with mean $\mu_\gamma + \frac{\sigma_r}{\sigma_\gamma}\rho(r-\mu_r)$ and standard deviation $\sqrt{1-\rho^2}\cdot\sigma_\gamma$. This, combined with the already-mentioned normality of f^r makes the computation of probabilities a straightforward exercise. Moreover, considering $x, y \in \{r, \gamma\}$ with $x \neq y$, an increase of μ_x leaves unaffected the marginal F^y while generating a dominating shift in all conditionals $F_{x|y}$. Hence, Claim 2 states that, by reducing the mean of δ or, alternatively, by increasing the mean of r, we generate a larger choice probability for the earlier alternative. Third, increasing σ_x leaves unaffected the marginal F^y and, under the appropriate correction of the covariance, it generates the expansion of all conditionals $F_{x|y}$. 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.

3.4. **An Estimation Exercise.** We now empirically implement the above theoretical model using the influential dataset compiled by Andersen et al. (2008). Our experiment involves a set of 253 individuals, which we index by i and a total of 100 menus, belonging either to type $A_\mathcal{R}$ or type $A_\mathcal{T}$, which we index by $\{A_m\}_{m=1}^M$. Each individual i makes choices from a subset of these 100 menus. Thus, we have to deal with a collection of 22,096 observations, i.e. pairs of menus with the associated choices, which we denote by $\mathcal{O}$.[14]

For purposes of illustration, we adopt a representative agent approach and estimate a single set of parameters of the bivariate normal, $(\mu_r, \sigma_r, \mu_\gamma, \sigma_\gamma, \rho)$.[15] We use the same integrated average daily wealth value as Andersen et al. (2008, p.600) i.e., 118 Danish kroner (DKK). Denoting by $I_{a,m}$ the number of individuals who choose alternative

[14]Not all the individuals face every menu. Appendix D contains further details about the experimental datasets used in the paper. Moreover, in this exercise we discard the dominated alternatives. Appendix E incorporates these alternatives into the analysis using a simple trembling version of the model.

[15]Notice that exactly the same methodology could be implemented for the estimation of individual preferences. Also, our methodology would also allow for the combined treatment of individual and population variability. The purpose of the exercise is to discuss how to practically implement our framework, and to show that already the simple representative agent approach can accommodate the data.

Table 1. Estimation results Andersen et al. (2008).

μ_r	σ_r	μ_γ	σ_γ	ρ	log-likelihood
0.77	1.00	5.30	1.00	0.71	11,809.09
(0.04)	(0.06)	(0.07)	(0.04)	(0.00)	

$a \in \{0,1\}$ in menu A_m, the log-likelihood function is:

$$\log \mathcal{L}\left(f|\mathcal{O}\right) = \sum_{m\in M}\Big[I_{0,m}\log \mathcal{P}_f(0,A_m) + I_{1,m}\log(1-\mathcal{P}_f(0,A_m))\Big],$$

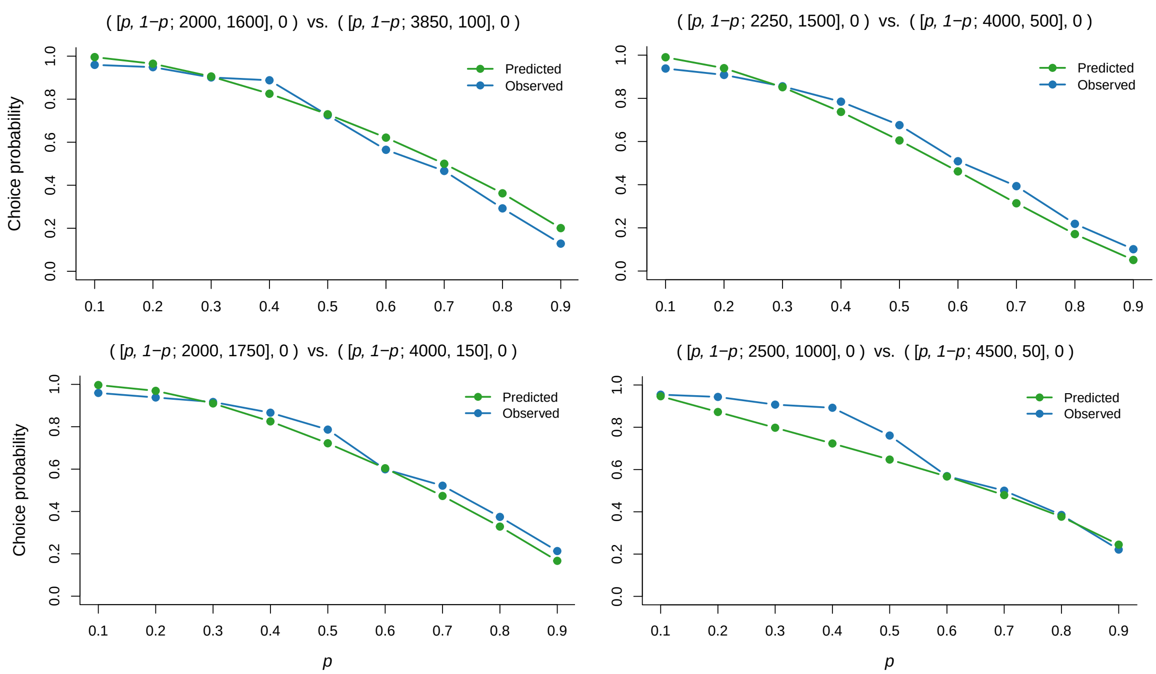
where the choice probabilities $\mathcal{P}_f(0,A_m)$ are the same as those presented in Propositions 1 and 2. Consistent estimates of the parameters of the bivariate normal can be achieved via maximization of the log-likelihood, with robust standard errors computed using the delta method and clustered at the individual level. Table 1 shows the estimated parameters.

We obtain a relatively large median curvature of $\mu_r = .77$, which is consistent with the observed marked tendency to choose safer over riskier lotteries in the risk menus. To facilitate interpretation of this curvature, we describe here the risk menu which yields the more stochastic choice pattern; that is, each one of the two lotteries is chosen by approximately half of the individuals. The safe and the risky lotteries are, respectively, $[.7,.3;2500,1000]$ and $[.7,.3;4500,50]$, with expected payments of 2050 and 3165 DKK. The CRRA curvature that makes these two lotteries indifferent is .83, in consonance with the median parameter estimate of .77. One standard deviation below (above) the median curvature corresponds to a curvature of $-.23$ (1.77). This is consistent with the high heterogeneity observed in risk menus, with a sizable amount of risk-neutral and moderately risk-loving behaviors on one side of the spectrum, and extremely risk-averse behaviors on the other.

The estimated μ_γ corresponds, using the transformation given in footnote 6, to a median monthly discount factor of .995. As previously discussed, for a proper interpretation of this value one needs to consider the estimated curvature. Given the relatively high curvature estimate, the monthly discount factor, .995, is a sign of impatience, resulting in a sizable fraction of early choices. To put this in perspective, a monthly discount factor of .995 with a curvature of .77 generates indifference between early and late payouts in the monetary range of the experimental menus for a risk-neutral individual with a monthly discount factor of around .978. That is, we observe behaviors in line with indifference towards 97.8 dollars today versus 100 dollars in a month's time.[16] One standard deviation below (above) the monthly discount factor corresponds to .987 (.998).[17]

Figure 1 (respectively, Figure 2) plots the observed and predicted choice probabilities of the safe (respectively, early) option across the different risk (respectively, time) menus. It is visually apparent that, across all risk and time menus, the concordance between observed and predicted choices is remarkably high. The average absolute values of the distance between observed and predicted choice probabilities for the risk and time menus are 0.05 and 0.08, respectively.

FIGURE 1. Observed and predicted choice probabilities in risk menus.

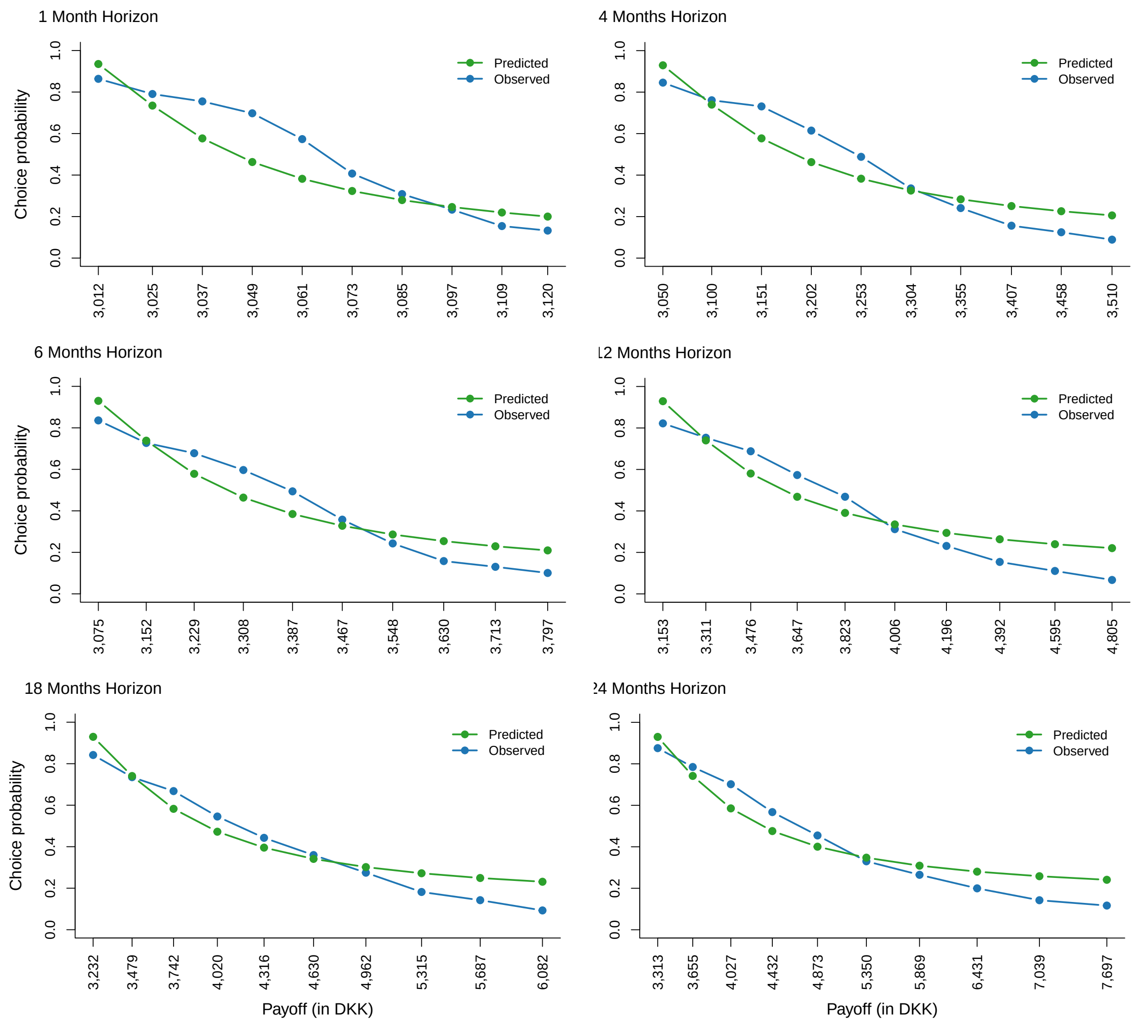


We conclude by noting that the preference heterogeneity underlying choice heterogeneity in experiments involving double multiple price lists is adequately captured

[16]Notice that this corresponds to an annual indifference between 76 dollars today versus 100 dollars in a year's time.

[17]These estimates must be read in combination with the positive covariance between the parameters, which is equal to .71. Overall, choice heterogeneity in time menus is sizable but slightly moderated by the fact that low (high) curvatures realize together with low (high) discount factors, thereby producing conditional choice probabilities consistently favoring one of the alternatives in the menu.

FIGURE 2. Observed and predicted choice probabilities in time menus.



thanks to the adoption of the appropriate random model. This is all the more remarkable because we have entertained the simplest, rational, model and used a representative agent approach. There is obviously room for improvement in these results either by incorporating other deterministic models of behavior or by allowing for heterogeneity to vary across individuals or groups. For an illustration, see Appendix F which reports the estimates obtained with slight variations in our main approach: (i) varying the deterministic model of behavior by replacing CRRA by CARA or exponential discounting by hyperbolic discounting, and (ii) using group-specific distributions of the parameters; in particular, age-dependent or gender-dependent parameters.

4. Convex Menus

In an alternative setting pioneered by Andreoni and Sprenger (2012), subjects are faced with the so-called convex menus.[18] The individual must select a value a in $A_{\mathcal{C}} = [0,1]$, with alternative a granting the individual the following two dated lotteries: $(([p^0, 1-p^0; (1-a)x^0, 0], t^0), ([p^1, 1-p^1; ax^1, 0], t^1))$ with $x^0 \leq x^1$ and $t^0 < t^1$.[19]

Although convex menus are more convoluted, the analysis can be analogously simplified by conditioning again on parameter r, because this also creates here an intuitive, ordered, structure of choices for parameter δ. Having fixed r, each $a \in [0,1)$ has an associated threshold $\delta^a_{A_{\mathcal{C}}}(r) \in (0,1]$, such that the choice is below a if and only if the value of parameter δ falls below the threshold. That is, $\Gamma(a, A_{\mathcal{C}}) = \{(r,\delta) : \delta \leq \delta^a_{A_{\mathcal{C}}}(r)\}$. 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 $\delta^a_{A_{\mathcal{C}}}(r)$, whenever below 1, 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 $\{\delta^a_{A_{\mathcal{C}}}(r)\}_{r\in\mathbb{R}}$, whenever below 1, is strictly increasing in r if and only if $a > e = \frac{x^0}{x^0+x^1}$, and strictly decreasing whenever $a < e$. The value 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 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 versus alternative 1. In the current continuous case, this trade-off has

[18] 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).

[19] The DEU evaluation of alternative a is $DEU_{r,\delta}(a) = \delta^{t^0} p^0 u_r(\omega + (1-a)x^0) + \delta^{t^1} p^1 u_r(\omega + ax^1)$.

alternative e as the critical value. Alternatives below (respectively, above) e allocate a larger potential prize to the earlier period (respectively, 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 $\mathcal{P}_f(e, A_\mathcal{C})$ closer to 1/2. That is, the probabilities of choices below and above e become closer, implying that behavior is now more stochastic.[20]

Proposition 3. *For every pair of RDEUs, f and g, and every menu $A_\mathcal{C}$:*

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

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

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

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

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

4.1. **Implications for the Parametric Version.** The implications of Proposition 3 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 separates the choice of 0 and 1 and it corresponds to $\min\left\{1, \left[\frac{p^0}{p^1}\right]^{\frac{1}{t^1-t^0}} \left[\frac{(\omega+x^0)^{1-r} - \omega^{1-r}}{(\omega+x^1)^{1-r} - \omega^{1-r}}\right]^{\frac{1}{t^1-t^0}}\right\}$. In the concave part, the threshold determining a choice below a can be obtained from the standard first-order condition, and corresponds to $\delta^a_{A_\mathcal{C}}(r) = \min\left\{1, \left[\frac{p^0}{p^1}\frac{x^0}{x^1}\left(\frac{\omega+ax^1}{\omega+(1-a)x^0}\right)^r\right]^{\frac{1}{t^1-t^0}}\right\}$. Since the choice probabilities are again built on the basis of the conditional probabilities that are normally distributed, computation is routine. Ceteris paribus, a decrease in the median of δ generates larger choice probabilities for alternatives allocating more resources to the earlier period. Given the 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.

[20]As in the case with $A_\mathcal{T}$, the median of any conditional must be different to $\delta^e_{A_\mathcal{C}}(r)$ or $r^e_{A_\mathcal{C}}(\delta)$. The thresholds $\{\delta^e_{A_\mathcal{C}}(r)\}_{r\in\mathbb{R}}$ are now strictly increasing only up to some value of the parameter r. As it will become clearer in the proof, the condition is only needed where the map is strictly increasing and we prefer to maintain notation for simplicity.

4.2. **An Estimation Exercise.** We use Andreoni and Sprenger's (2012) original 80-subject dataset, where individuals indexed by i choose options from a set of 84 convex menus denoted by $\{A_m\}_{m=1}^M$. The experimental implementation uses discretized versions of the continuous share problem, allowing for 101 equidistant shares, including the two corners. For practical reasons, we further discretize the choice set to a set of S possible shares, including the two corners, i.e., $0 = a_1 < a_2 < \cdots < a_S = 1$.[21] There is a total of 6,523 undominated observations, which we denote by $\mathcal{O}$.[22]

As in Section 3, we assume the bivariate normal and CRRA using the same integrated average daily wealth value as in Andreoni and Sprenger (2012) i.e., $\omega = \$4.05$. Denoting by $I_{a_s,m}$ the number of individuals who choose option a_s in menu A_m, the log-likelihood of the representative agent can be written as

$$\log \mathcal{L}\,(f|\mathcal{O}) = \sum_{m=1}^{M}\sum_{s=1}^{S} I_{a_s,m} \log[\mathcal{P}_f(a_s, A_m) - \mathcal{P}_f(a_{s-1}, A_m)],$$

where the choice probabilities $\{\mathcal{P}_f(a_s, A_m)\}_{s=1}^S$ are the same as those studied in Proposition 3 which, for notational convenience, we write as $\mathcal{P}_f(a_{-1}, A_m) = 0$. Table 2 shows the parameter estimates.

Table 2. Estimation results Andreoni and Sprenger (2012).

μ_r	σ_r	μ_γ	σ_γ	ρ	log-likelihood
0.14	0.96	4.83	3.69	−0.71	11,173.92
(0.08)	(0.09)	(0.30)	(0.27)	(0.00)	

Among the findings, we stress the following. The median curvature is equal to .14, with relatively large variability, as reflected by the one standard deviations around the median curvature, which are −.82 and 1.1. These estimation results are explained by our theoretical findings; although the subjects are risk averse, we know that only negative curvatures cause DEU to predict corner solutions, and these are the most common choices in the data. The estimates are therefore the result of balancing both factors.

[21] Since the vast majority of participants tend to choose multiples of 10, we use $S = 11$ options comprising $[0, .04]$, $[.05, .14]$, $\ldots$, $[.95, 1]$.

[22] As in the double MPLs, there are some dominated options for all parameter values, such as $a \in [e, 1]$ whenever $x^0 = x^1$ and $p^0 \geq p^1$, or $a \in [e, 1)$ whenever $p^0x^0 \geq p^1x^1$. In Appendix E, we again incorporate all observations into the estimation by applying a trembling version of the model.

The monthly discount factor implied by the estimated median μ_γ is .992. The one standard deviations around the monthly discount factor are .76 and approximately 1, again revealing a high degree of heterogeneity, motivated, as before, by the large fraction of corner choices.[23]

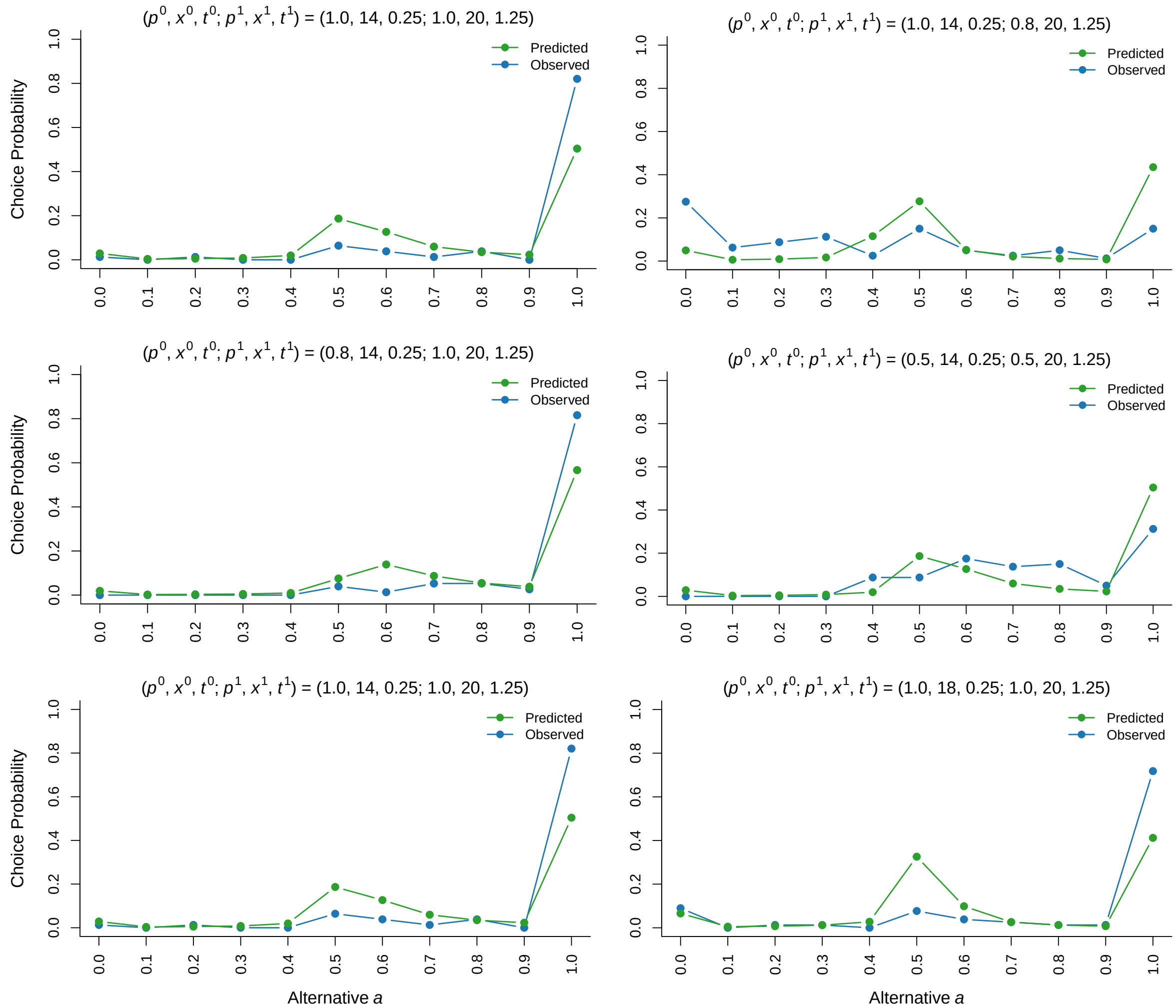
Figure 3 shows the predicted and observed choice frequencies for a sample of menus. We selected menu $(p^0, x^0, t^0; p^1, x^1, t^1) = (1, 14, 0.25; 1, 20, 1.25)$ and performed variations in the probability for late payment, the probability for early payment, the probabilities for both types of payment keeping their ratios constant, the timing for late payment, and the amount allocated for early payment. Appendix G reports the predicted and observed choice frequencies for all the menus. The figure neatly shows that the predictions of the simple model we have adopted appear to perform remarkably well. The model is able to capture the menu-dependent tri-modal choice patterns, namely, the large menu-dependent fraction of corner choices, followed by the large menu-dependent fraction of central choices. The average absolute distance between observed and predicted choice probabilities is 0.07. This result is remarkable for the following reasons: (i) we are using the simplest possible rational model, discounted expected utility, with a simple account of heterogeneity via the random utility model, (ii) the apparent complexity of observed choice behavior, and (iii) the difficulties encountered in previous empirical attempts to comprehensively account for these choice patterns.

5. Final Remarks

In this paper we have studied the random utility model in the context of the most standard treatment of risk and time preferences, namely discounted expected utility. 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 statistics. In addition, we have applied the model to two very different datasets, and

[23] The covariance is now negative, a result that can be interpreted based on our theoretical results. Notice that, as the threshold for low values of a is decreasing in the curvature, the selection of these options can only be explained by concave utility functions with very low discount factors. Now, given the low median estimated curvature and high median estimated discount factor, only a negative correlation between the parameters can explain these observations. Interestingly, this result is driven by a very small fraction of choices. When the trembling parameter is incorporated in Appendix E, these choices are accounted for by trembling and the estimates replicate the high positive correlation seen in binary menus.

FIGURE 3. Observed and predicted frequencies in a sample of menus.



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 $A_{\mathcal{R}} = \{0, 1\} = \{([p, 1-p; x_1^0, x_2^0], 0), ([p, 1-p; x_1^1, x_2^1], 0)\}$, 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; therefore, 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 $DEU_{r,\delta}(0) \geq DEU_{r,\delta}(1)$, then $DEU_{r',\delta'}(0) > DEU_{r',\delta'}(1)$. To see this, suppose that $DEU_{r,\delta}(0) \geq DEU_{r,\delta}(1)$. 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 $DEU_{r',\delta'}(0) > DEU_{r',\delta'}(1)$ and concludes the argument. In other words, $\Gamma(0, A_{\mathcal{R}})$ must correspond to the cartesian product of an interval in $\mathbb{R}$ and the interval $(0, 1)$. 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 and 1 are preferred, and there must be a unique $r^0_{A_{\mathcal{R}}} \in \mathbb{R}$ such that alternative 0 is preferred if and only if $r \geq r^0_{A_{\mathcal{R}}}$ which leads to Claim 1 in the proposition. For Claim 2, note that whenever F^r dominates G^r, it must be $\mathcal{P}_f(0, A_{\mathcal{R}}) = 1 - F^r(r^0_{A_{\mathcal{R}}}) > 1 - G^r(r^0_{A_{\mathcal{R}}}) = \mathcal{P}_g(0, A_{\mathcal{R}})$. For Claim 3, notice that the assumption requires us to consider two cases, $\text{med}(F^r) > r^0_{A_{\mathcal{R}}}$ or $\text{med}(F^r) < r^0_{A_{\mathcal{R}}}$. In the first case, since F^r expands G^r, it must be $F^r(r^0_{A_{\mathcal{R}}}) > G^r(r^0_{A_{\mathcal{R}}}) > 1/2$, while in the second case, it must be that $F^r(r^0_{A_{\mathcal{R}}}) < G^r(r^0_{A_{\mathcal{R}}}) < 1/2$, concluding the proof. ∎

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

$$DEU_{r,\delta}(0) \geq DEU_{r,\delta}(1) \Leftrightarrow \delta \leq \delta^0_{A_{\mathcal{T}}}(r) = \left[\frac{u_r(\omega + x^0)}{u_r(\omega + x^1)}\right]^{\frac{1}{t^1 - t^0}}.$$

Strict monotonicity of u_r guarantees that this threshold is always a value in $(0, 1)$ and, hence, the first expression in Claim 1, and Claim 2a, follow immediately.

We now claim that the threshold map $\{\delta^0_{A_{\mathcal{T}}}(r)\}_{r\in\mathbb{R}}$ is strictly increasing 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 $\left[\frac{v_r(\omega+x^0)}{1}\right]^{\frac{1}{t^1-t^0}}$ is strictly increasing 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 increasing.

The unbounded curvature assumption also proves that the map is onto and hence, a bijection. It can be inverted to obtain the strictly increasing thresholds $\{r^0_{A_{\mathcal{T}}}(\delta)\}_{\delta\in(0,1)}$, such that, for a given δ, alternative 0 is chosen if and only if r is above this threshold. Hence, $\mathcal{P}_f(0, A_{\mathcal{T}}) = \int_r F_{\delta|r}(\delta^0_{A_{\mathcal{T}}}(r)) f^r(r)\mathrm{d}r = \int_\delta (1 - F_{r|\delta}(r^0_{A_{\mathcal{T}}}(\delta))) f^\delta(\delta)\mathrm{d}\delta$, 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 $\{\delta^0_{A_{\mathcal{T}}}(r)\}_{r\in\mathbb{R}}$ guarantees that all conditional medians of δ lie on the same side of the threshold map. As a result, the same alternative, either

0 or 1, 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: Consider a menu $A_{\mathcal{C}}$ defined by values $(p^0, x^0, t^0; p^1, x^1, t^1)$. We first claim that, for every $r \in \mathbb{R}$, the argument that maximizes $DEU_{r,\delta}$ is increasing 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 $a^* = 0$, we are done. Consider then the case of $a^* > 0$ and any alternative $0 \leq a < a^*$. Given the optimality of a^*, we know that $DEU_{\delta,r}(a) \leq DEU_{\delta,r}(a^*)$, i.e., $\delta^{t^0}p^0u_r(\omega + (1-a)x^0) + \delta^{t^1}p^1u_r(\omega + ax^1) \leq \delta^{t^0}p^0u_r(\omega + (1-a^*)x^0) + \delta^{t^1}p^1u_r(\omega + a^*x^1)$ holds. The latter inequality is equivalent to $p^0u_r(\omega + (1-a)x^0) + \delta^{t^1-t^0}p^1u_r(\omega+ax^1) \leq p^0u_r(\omega+(1-a^*)x^0)+\delta^{t^1-t^0}p^1u_r(\omega+a^*x^1)$. Now, it is evident that an increase in the discount factor leaves unaffected the first term in both the left and the right hand side but increases more significantly the second term of the the right hand side, because the function u_r is strictly increasing. Hence, alternative a^* must be preferred to alternative a for the larger discount factor, and the argument maximizing DEU must be a^* or larger. We have proved our claim. Hence, given $r \in \mathbb{R}$, we can define $\delta^a_{A_{\mathcal{C}}}(r) \in (0, 1]$ as the supremum of the discount factors for which any alternative in $[0, a]$ is the DEU maximizer. Trivially, the maximizer of $DEU_{r,\delta}$ is below a if and only if δ lies below $\delta^a_{A_{\mathcal{C}}}(r)$, i.e., $\Gamma(a, A) = \{(r, \delta) : \delta \leq \delta^a_{A_{\mathcal{C}}}(r)\}$. Claim 1 follows immediately.

For the case of convex monetary utilities, i.e., $r \leq 0$, convexity and the fact that $u_r(\omega) = 0$ guarantee that $\delta^{t^0}p^0u_r(\omega+(1-a)x^0)+\delta^{t^1}p^1u_r(\omega+ax^1) \leq \delta^{t^0}p^0[au_r(\omega)+(1-a)u_r(\omega+x^0)]+\delta^{t^1}p^1[(1-a)u_r(\omega)+au_r(\omega+x^1)] = \delta^{t^0}p^0(1-a)u_r(\omega+x^0)+\delta^{t^1}p^1au_r(\omega+x^1) \leq \max\{\delta^{t^0}p^0u_r(\omega + x^0), \delta^{t^1}p^1u_r(\omega + x^1)\}$. Hence, only alternatives 0 or 1 can be the maximizers of $DEU_{r,\delta}$. Thus, the thresholds $\{\delta^a_{A_{\mathcal{C}}}(r)\}_{r\leq 0}$ are indeed independent of a. Notice that, whenever the threshold is below 1, it corresponds to the discount factor that, given r, equalizes the DEU value of 0 and 1. That is, whenever $r \leq 0$, the threshold for every $a \in [0, 1)$ is $\min\left\{1, \left[\frac{EU_r(0)}{EU_r(1)}\right]^{\frac{1}{t^1-t^0}}\right\} = \min\left\{1, \left[\frac{p^0u_r(\omega+x^0)}{p^1u_r(\omega+x^1)}\right]^{\frac{1}{t^1-t^0}}\right\}$, where EU_r refers to the expected utility of the corresponding lottery using the monetary utility u_r. This can also be written as $\min\left\{1, \left[\frac{p^0}{p^1}\right]^{\frac{1}{t^1-t^0}}\delta^0_{A_{\mathcal{T}}}(r)\right\}$, with $\delta^0_{A_{\mathcal{T}}}(r)$ 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. Since we know that $\delta^0_{A_{\mathcal{T}}}(r)$ converges to 0 whenever r approaches $-\infty$, the threshold must belong to $(0, 1)$ for at least an interval of values of r on the negative reals. We can then use this fact to prove Claim 2a. To do so, consider any value $a \in [0, 1)$. We can reproduce the analysis in Proposition 2 using the thresholds $\{\delta^a_{A_C}(r)\}_{r\in\mathbb{R}}$. The only significant novelty is that, whenever $\delta^a_{A_C}(r) = 1$, the conditional probability of choosing an alternative below a is already equal to 1 and, hence, it cannot further increase. However, given that $\delta^a_{A_C}(r) \in (0, 1)$ for at least an interval of the negative reals, the conditional probability of choosing an alternative below a strictly increases with downward shifts of δ whenever r belongs to such interval. Given the measurability of f, Claim 2a follows.

We now analyze strictly concave utilities, $r > 0$. We start by claiming that, whenever $r > 0$, the threshold $\delta^a_{A_C}(r)$ is decreasing for every $a < e$, and increasing for every $a > e$. We start with the former, assuming by contradiction that $0 < r < r'$ but $\delta^a_{A_C}(r) < \delta^a_{A_C}(r')$ for some $a < e$. Using continuity and the definition of the thresholds, there must exist δ^* with $\delta^a_{A_C}(r) < \delta^* < \delta^a_{A_C}(r')$ such that the maximizer for DEU_{r,δ^*} is a^*, with $e > a^* > a$. Consider any $a' < a^*$. It must be $DEU_{r,\delta^*}(a^*) \geq DEU_{r,\delta^*}(a')$, i.e., $(\delta^*)^{t^0} p^0 u_r(\omega + (1 - a^*)x^0) + (\delta^*)^{t^1} p^1 u_r(\omega + a^* x^1) \geq (\delta^*)^{t^0} p^0 u_r(\omega + (1 - a')x^0) + (\delta^*)^{t^1} p^1 u_r(\omega + a' x^1)$. Dividing both terms by the positive constant $p^0(\delta^*)^{t^0} + p^1(\delta^*)^{t^1}$ and denoting $p = \frac{p^0(\delta^*)^{t^0}}{p^0(\delta^*)^{t^0} + p^1(\delta^*)^{t^1}}$, 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 and a to alternative 1. Hence, since a^* is preferred at (r, δ^*), we know from Proposition 1 that a^* will also be preferred at (r', δ^*) because $r' > r$.[24] Thus, the maximizer of DEU_{r',δ^*} cannot be below a^*. This contradicts the definition of $\delta^a_{A_C}(r')$. That is, the threshold must be decreasing whenever $a < e$. The proof that the threshold is increasing whenever $a > e$ is analogous and thus omitted.

Consider now $0 < \underline{a} < e < \overline{a} < 1$. From the previous reasoning, the set of values of δ for which the solution belongs to $(\underline{a}, \overline{a})$ is (an) increasing (interval) in r. Thus, for any given δ, a dominating shift in the conditional distribution of r will create an increase in the conditional mass of the set of values of δ in $(\delta^{\underline{a}}_{A_C}(r), \delta^{\overline{a}}_{A_C}(r))$. Now, consider any value of δ below $\delta^{\underline{a}}_{A_C}(0)$. We know that alternative 0 is the maximizer of $DEU_{0,\delta}$. For these values of δ, and given the unbounded curvature assumption, alternative $\underline{a}$ will be better than any other alternative below it for r sufficiently large, and hence we know that the conditional shift guarantees strictly more choice probability for at least an

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

interval of discount factors. The measurability assumption of f guarantees that Claim 2b holds.

We now prove Claim 3a. We already know that the choice belongs to $[0, e)$ if and only if $\delta < \delta^e_{A_\mathcal{C}}(r)$. We can then reproduce the analysis of Proposition 2 for the case of expansions in the conditional distribution of δ. Like in previous cases, there may be a range of values of r for which $\delta^e_{A_\mathcal{C}}(r) = 1$ but we already proved that this no longer is true for an interval of negative reals. Thus, Claim 3a follows. To show Claim 3b, note that the map $\delta^e_{A_\mathcal{C}}(r)$ must be always strictly increasing up to some value $r^* \leq 0$ and then constant. There are two cases to consider: (i) $\delta^e_{A_\mathcal{C}}(0) = 1$ and (ii) $0 < \delta^e_{A_\mathcal{C}}(0) < 1$. In the former case, every δ has an associated, strictly increasing threshold $r^e_{A_\mathcal{C}}(\delta)$ such that $\delta^e_{A_\mathcal{C}}(r^e_{A_\mathcal{C}}(\delta)) = \delta$. This corresponds to a negative value of r where the choice of options 0 and 1 is separated exactly for discount factor δ. We can then reproduce the proof of Proposition 2. In the latter case, notice that for any $\delta > \delta^e_{A_\mathcal{C}}(0)$, choices always correspond to alternatives that satisfy $a > e$ and hence the conditional choice probability that we are analyzing is always equal to 0, and trivially unaffected by the expansion of parameter r. Whenever $\delta < \delta^e_{A_\mathcal{C}}(0)$, we can again find the associated strictly increasing negative threshold $r^e_{A_\mathcal{C}}(\delta)$ such that $\delta^e_{A_\mathcal{C}}(r^e_{A_\mathcal{C}}(\delta)) = \delta$. We can then apply again the same logic than in Proposition 2 for Claim 3b, concluding the proof.■

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, $A_\mathcal{H} = \{0, 1\}$ with $0 = ([p, 1-p; x^0_1, x^0_2], t^0)$ and $1 = ([p, 1-p; x^1_1, x^1_2], t^1)$ such that $x^1_1 > x^0_1 > x^0_2 > x^1_2$, $p \in (0, 1)$ and $t^0 < t^1$. Denote by $r^0_\mathcal{H}$ the value of r, that we know exists because of Proposition 1, such that the two lotteries are equally attractive if paid in the present.[25] Since $\delta < 1$, it is then evident that for every $r \geq r^0_\mathcal{H}$, the safer alternative 0 is chosen no matter the discount factor. Whenever $r < r^0_\mathcal{H}$, it is similarly evident that the earlier alternative 0 will be selected if and only if δ lies below the threshold $\delta^0_{A_\mathcal{H}} = \left[\frac{EU_r(0)}{EU_r(1)}\right]^{\frac{1}{t^1 - t^0}}$. We can then straightforwardly compute the choice probability of alternative 0 as:

$$\mathcal{P}_f(0, A_\mathcal{T}) = 1 - F^r(r^0_\mathcal{H}) + \int_{r < r^0_\mathcal{H}} F_{\delta|r}(\delta^0_{A_\mathcal{H}}(r)) f^r(r) \mathrm{d}r.$$

[25]This corresponds to $\left[\frac{p(\omega+x^0_1)^{1-r}+(1-p)(\omega+x^0_2)^{1-r}-\omega^{1-r}}{p(\omega+x^1_1)^{1-r}+(1-p)(\omega+x^1_2)^{1-r}-\omega^{1-r}}\right]^{\frac{1}{t^1-t^0}}$ in the case of CRRA.

The effect of shifts and spreads of δ are trivially understood from this structure, by applying the logic of Proposition 2 to the measurable interval $(-\infty, r^0_{\mathcal{H}})$. 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 ratio is indeed strictly increasing in r as long as $r < r^0_{\mathcal{H}}$, which is all that is needed for the analysis of a hybrid menu. Using RDEU and these standard families of monetary utilities, the effect of shifts and spreads of r can be intuitively understood by merely reproducing the logic of Proposition 2.

Appendix C. Baseline wealth and CRRA

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 $r^0_{A_{\mathcal{R}}} \neq 0$, $r^0_{A_{\mathcal{R}}}$ is strictly increasing (respectively, decreasing) in ω whenever $r^0_{A_{\mathcal{R}}} > 0$ (respectively, $r^0_{A_{\mathcal{R}}} < 0$).[26] Consequently, ceteris paribus, the alternative with larger expected value will be chosen more often. In time menus $A_{\mathcal{T}}$, every threshold $\delta^0_{A_{\mathcal{T}}}(r)$ converges monotonically to the constant $\delta^0_{A_{\mathcal{T}}}(0)$ 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 (alternatively, lover) individuals will choose more often the present (future) option. Similarly, in convex menus $A_{\mathcal{C}}$, every threshold map $\delta^a_{A_{\mathcal{C}}}(r)$ converges monotonically to the constant map $\delta^0_{A_{\mathcal{C}}}(0)$ as ω increases. The conditional behavior of every r becomes more aligned with the conditional choices of $r = 0$ (with interior solutions vanishing).

We illustrate these theoretical findings by reproducing the estimation exercise reported in the main text, but now assuming a larger baseline wealth level. In order to make these consistent across the experimental settings, instead of using the daily baseline wealth level we use the corresponding monthly ones, that amount to $3,304$DKK and \$113.40 in Andersen et al. (2018) and Andreoni-Sprenger (2012), respectively. Tables 3 and 4 report the results. The results are consistent with our theoretical analysis. First, the estimation now needs higher levels of (median and variance) risk aversion parameters to explain the same revealed preference of safe options. Second, given the correction of the risk aversion parameters and the less intense correlation coefficients, the time parameters are remarkably similar to those obtained in the main specifications.

[26] In the degenerate case where the expected values of both lotteries coincide, we obviously have $r^0_{A_{\mathcal{R}}} = 0$ for all levels of ω.

TABLE 3. Estimation results Andersen et al. (2008), $\omega = 3,304$.

μ_r	σ_r	μ_γ	σ_γ	ρ	log-likelihood
2.60	3.11	5.21	0.76	0.14	11,699.59
(0.13)	(0.14)	(0.07)	(0.05)	(0.01)	

Third, notice that both likelihoods improve slightly with respect to those obtained in the main text.

TABLE 4. Estimation results Andreoni and Sprenger (2012), $\omega = 113.40$.

μ_r	σ_r	μ_γ	σ_γ	ρ	log-likelihood
1.77	8.29	5.00	2.73	-0.02	11,141.21
(0.69)	(0.69)	(0.38)	(0.23)	(0.01)	

C.1. **Negligible baseline wealth.** 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 risk aversion kicking in risk menus. In time menus $A_{\mathcal{T}}$, it can be seen that all thresholds $\delta^0_{A_{\mathcal{T}}}(r)$ converge to the value $\left(\frac{x^0}{x^1}\right)^{\frac{1-r}{t^1-t^0}}$, whenever $r < 1$. Hence, for parameters (r, δ), with $r < 1$, the earlier option 0 is preferred to the later option for such parameters if and only if $\delta^{\frac{1}{1-r}} \leq \left(\frac{x^0}{x^1}\right)^{\frac{1}{t^1-t^0}}$. That is, the expression $\delta^{\frac{1}{1-r}}$ represents a simple correction of the discount factor 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 $DEU_{0,\delta^{\frac{1}{1-r}}}$, 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 $\delta^{\frac{1}{1-r}}$ can be considered. Importantly, behavior for $r \geq 1$ becomes extreme when wealth is negligible, as alternative 0 is always preferred. Finally, in convex menus $A_{\mathcal{C}}$, negligible wealth simplifies the analysis of interior solutions since, for every $r > 0$, only interior solutions can exist; the thresholds $\delta^0_{A_{\mathcal{C}}}(r)$ and $\lim_{a \to 1} \delta^a_{A_{\mathcal{C}}}(r)$ converge respectively to 0 and 1.

APPENDIX D. DATASETS

D.1. **Andersen et al. (2018).** This study separately elicited risk and time preferences of a representative sample of 253 subjects from the adult Danish population. In the risk part, four multiple-price lists were implemented, each comprising ten pairs of present lotteries. The four tasks were (i) $([p, 1-p; 3850, 100], 0)$ and $([p, 1-p; 2000, 1600], 0)$, (ii) $([p, 1-p; 4000, 500], 0)$ and $([p, 1-p; 2250, 1500], 0)$, (iii) $([p, 1-p; 4000, 150], 0)$ and $([p, 1-p; 2000, 1750], 0)$, and (iv) $([p, 1-p; 4500, 50], 0)$ and $([p, 1-p; 2500, 1000], 0)$, with $p \in \{.1, .2, .3, .4, .5, .6, .7, .8, .9, 1\}$. All 253 subjects were confronted with the four tasks, with 116 individuals facing all pairs, 67 individuals facing pairs 3, 5, 7, 8, 9, and 10, and the remaining 70 subjects facing pairs 1, 2, 3, 5, 7, and 10, for a total of 7,928 choices in this part. In the time part, six multiple-price lists involving dated degenerate lotteries were implemented. The first degenerate lottery always paid 3.000 DKK after one month. The second degenerate lottery was designed by varying two parameters: (i) the awarding time, which could vary between 2, 5, 7, 13, 19 or 25 months, and (ii) the annual interest obtained by the subject, which increased by multiples of 5, from 5% to 50%.[27] All the subjects faced all 60 temporal binary choices, making a total of 15,180 choices. That is, $\mathcal{O}$ is formed by 23,108 observations with the number of individual observations varying between 84 and 100. For every pair of options, subjects could either choose one of them or express indifference between the two. In the latter case, they were told that the experimenter would settle indifferences by tossing a fair coin. Only 5% of the choices were indifferences; as in the original paper, we treat them assigning a half choice to each of the two gambles.

D.2. **Andreoni and Sprenger (2012).** Andreoni and Sprenger (2012) introduced risky considerations in the convex budget design of Andreoni and Sprenger (2012a). 80 undergraduate students took part in the experiment, each making 84 choices, for a total of 6,720 observations. The experimental parameters can be described, following our notation, as follows: $([p, 1-p; \alpha x, 0], t)$ and $([q, 1-q; (1-\alpha)y, 0], s)$, with $(p, q) \in \{(1, 1), (0.5, 0.5), (1, 0.8), (0.5, 0.4), (0.8, 0.1), (0.4, 0.5)\}$, $(t, s) \in \{(7, 28), (7, 56)\}$ in days, and $(x, y) \in \{(20, 20), (19, 20), (18, 20), (17, 20), (16, 20), (15, 20), (14, 20)\}$. The decision variable was discrete, with $\alpha \in \{\frac{0}{100}, \frac{1}{100}, \ldots, \frac{100}{100}\}$.

Appendix E. Tremble

We incorporate into the empirical estimations the choices of dominated alternatives by introducing a tremble parameter; with a large probability $1-\nu$, the individual chooses according to RDEU and, with a small probability ν, the individual trembles.

[27]The interest was compounded quarterly to be consistent with general Danish banking practices.

Table 5. Estimation results Andersen et al. (2008), tremble.

μ_r	σ_r	μ_γ	σ_γ	ρ	ν	log-likelihood
0.81	0.58	5.20	1.04	0.71	0.15	11,909.48
(0.04)	(0.03)	(0.08)	(0.05)	(0.00)		

Table 6. Estimation results Andreoni and Sprenger (2012), tremble.

μ_r	σ_r	μ_γ	σ_γ	ρ	ν	log-likelihood
0.06	0.98	4.56	2.84	0.71	0.15	11,713.28
(0.09)	(0.09)	(0.32)	(0.25)	(0.00)		

This guarantees positive probabilities over dominated alternatives, allowing a well-defined log-likelihood. Given a distribution f and trembling ν, the cumulative choice probabilities at a is $\mathcal{P}_{f,\nu}(a, A_\kappa) = (1-\nu)\mathcal{P}_f(a, A_\kappa) + \nu\,\mathbb{U}_\kappa$, where $\mathbb{U}_\kappa = 1/2$ for $A_\mathcal{R}$ and $A_\mathcal{T}$, and $\mathbb{U}_\kappa = a$ for $A_\mathcal{C}$.

Tables 5 and 6 report the estimated distributions of risk and time preferences with tremble, using CRRA with the same baseline wealth level consumption levels used in the main text. We see that levels of tremble found are relatively small, .15 in both cases. In Andersen et al. (2008) the tremble helps to explain some of the more extreme choices, and hence the estimations require a lower variance of the risk parameter. In Andreoni and Sprenger (2012) we observe a lower median risk parameter. This is so because the tremble parameter helps in explaining interior solutions, that are relatively less prevalent than corner ones, and hence lower levels of risk aversion are needed to explain interior choices.

Appendix F. Variations in the Empirical Analysis of Andersen et al.

We start by reporting the estimation results of substituting the CRRA monetary utility by the CARA one: for every $r \neq 0$, $u_r^{cara}(\omega + x) = \frac{e^{-r\omega} - e^{-r(\omega+x)}}{r} = e^{-r\omega}\left(\frac{1-e^{-rx}}{r}\right)$ and $u_0^{cara} = (\omega + x) - \omega = x$. The empirical strategy is identical to the one reported in the main text, leading to the estimation results reported in Tables 7. Notice that in terms of the likelihood, the CARA specification seems to be slightly worse than the corresponding CRRA one.

Table 7. Estimation results with CARA utility function.

μ_r^*	σ_r^*	μ_γ^*	σ_γ^*	ρ	log-likelihood
0.48	0.56	5.14	0.85	0.71	11,937.05
(0.03)	(0.02)	(0.07)	(0.05)	(0.00)	

Note: Parameter values with an asterisk are multiples of 1,000 e.g., $\mu_r^* = \mu_r \times 1,000$.

Table 8. Estimation results with hyperbolic discounting.

μ_r	σ_r	μ_g	σ_g	ρ	log-likelihood
0.77	0.98	-5.22	1.03	-0.71	11,808.78
(0.04)	(0.06)	(0.08)	(0.05)	(0.00)	

One may also wonder about the implications of substituting exponential discounting by the hyperbolic formulation. Namely, we now consider the hyperbolic expected utility

$$HEU_{r,d}(a) = \sum_{j=1}^{J} \frac{1}{1+dt^j} \sum_{n=1}^{N_j} p_n^j u_r^{crra}(\omega + x_n^j),$$

with $d \in \mathbb{R}_{++}$. Note that, given r, more impatience is now captured with higher values of d. Propositions 1 and 2 extend readily to this setting.[28] In the empirical estimation of the model we transform d into a parameter g in $\mathbb{R}$ by setting $g = \log d$. The results are reported in Table 8. The hyperbolic specification seems to perform as well as the exponential case used in the main text.

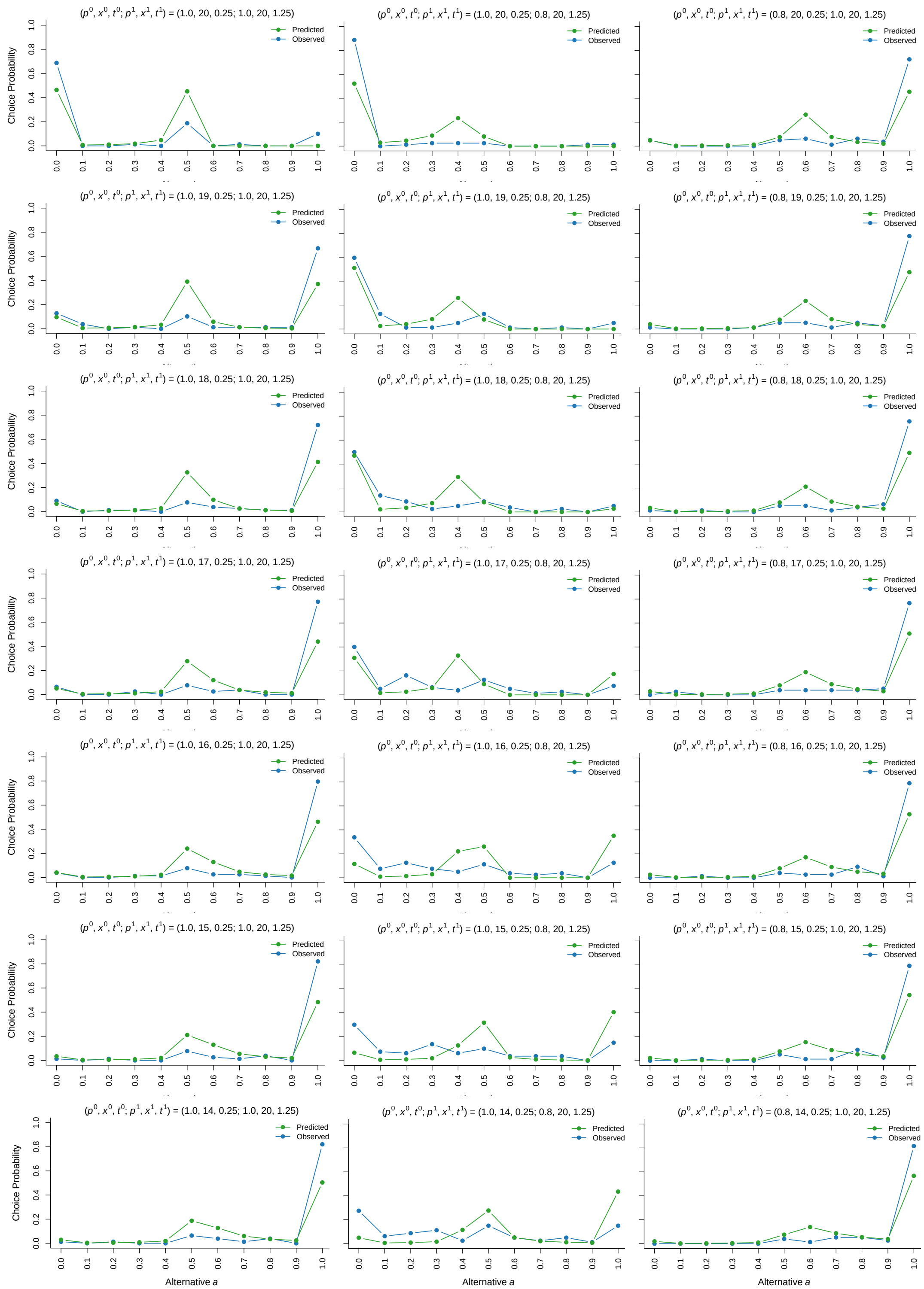
Finally, our empirical strategy can be directly applied to the study of the behavior of particularly relevant groups. Here we illustrate using our main specification for the study of gender and age. Table 9 reports the estimates for females, males, and four group ages. In both cases there is a slight improvement in the aggregated likelihoods with respect to the main specifications.

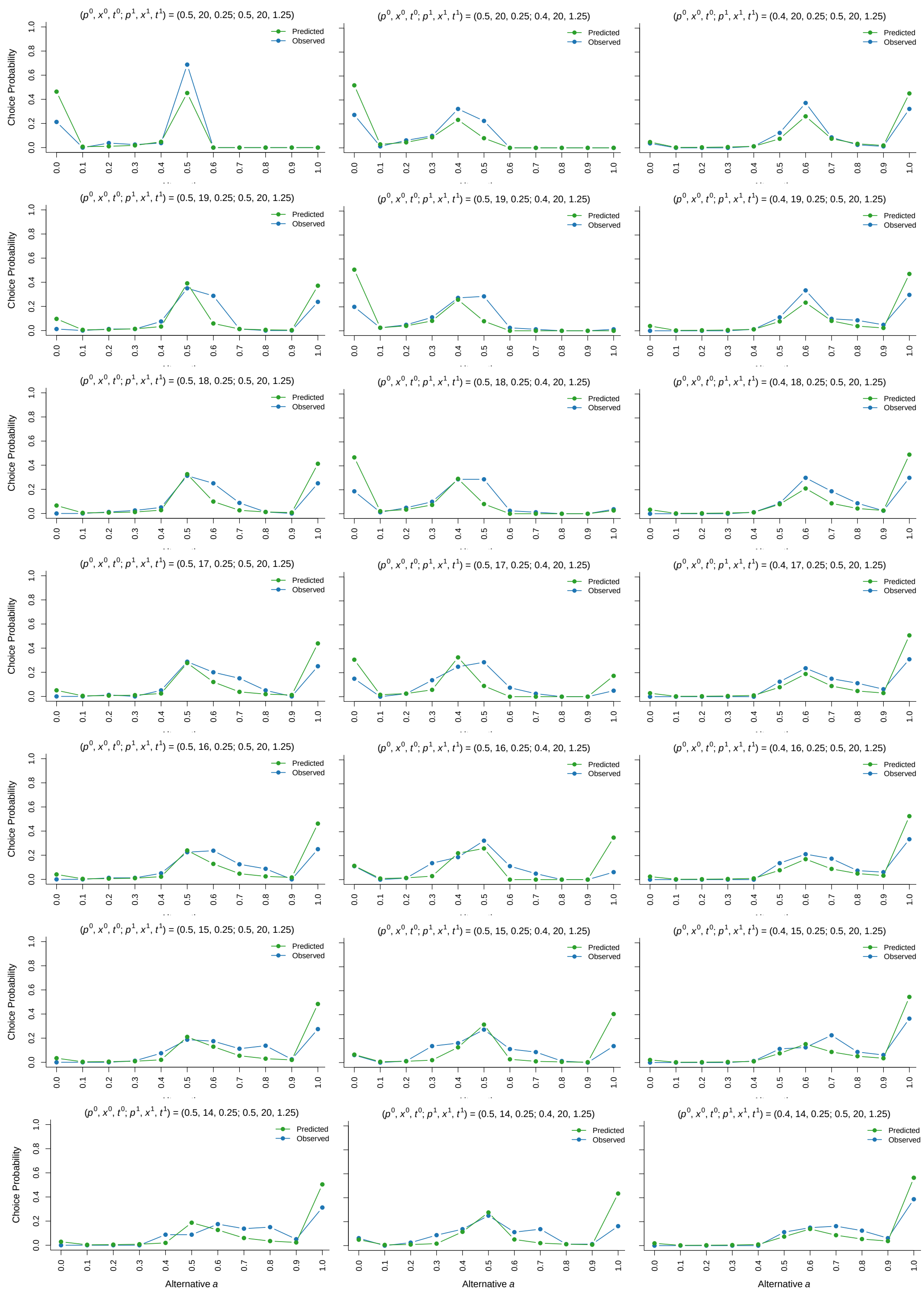
[28]The estimation program uses the corresponding integral, the derivation of which are available upon request.

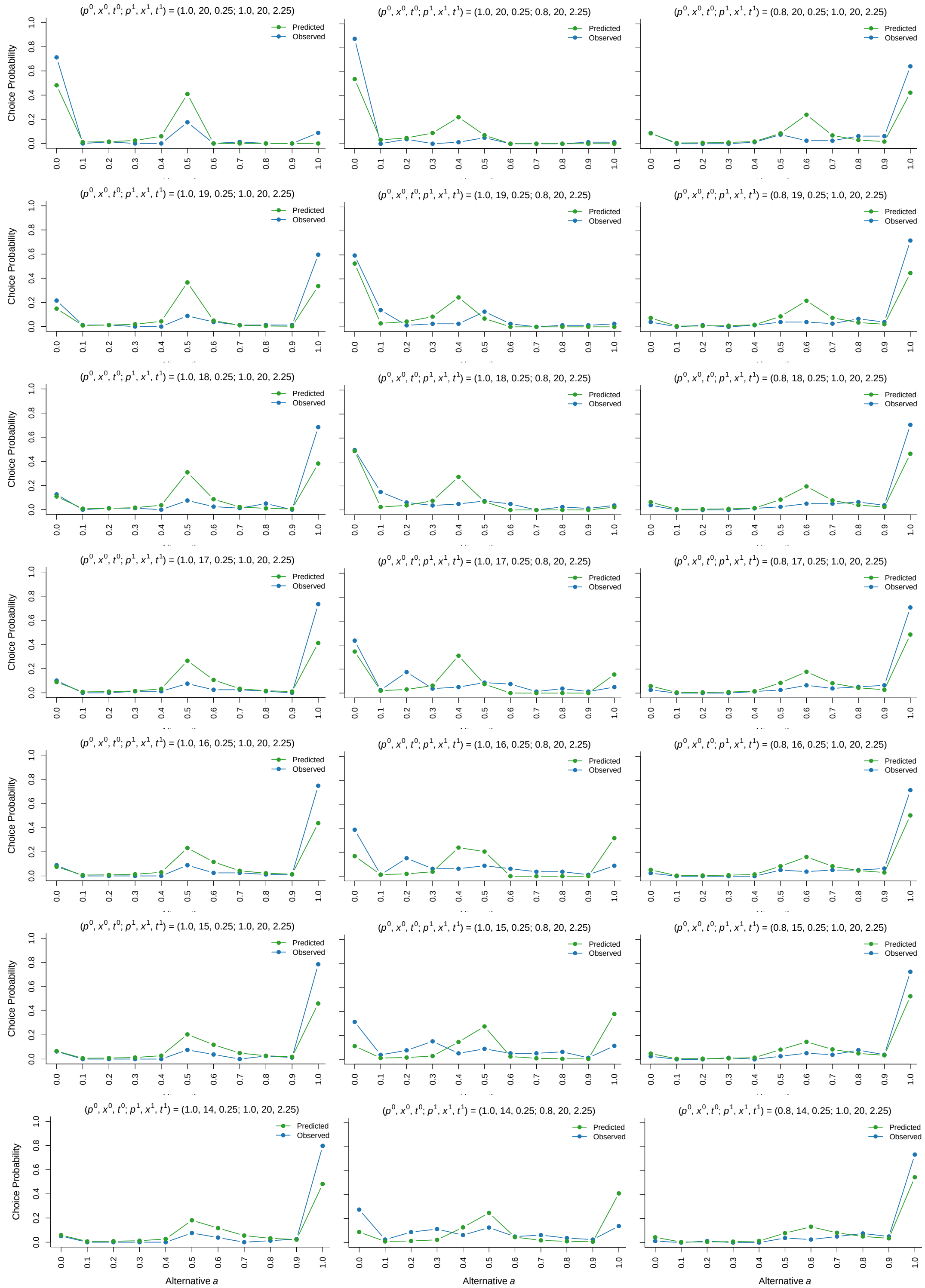
Table 9. Estimation results for certain subgroups.

Subgroup	N	μ_r	σ_r	μ_γ	σ_γ	ρ	log-likelihood
Female	130	0.81	1.15	5.41	0.99	0.71	6,164.00
		(0.05)	(0.09)	(0.08)	(0.04)	(0.00)	
Male	123	0.75	0.86	5.21	1.00	0.71	5,626.68
		(0.04)	(0.06)	(0.09)	(0.05)	(0.00)	
Young	43	1.09	0.86	5.88	0.99	0.71	1,952.58
($<$ 30 years)		(0.03)	(0.05)	(0.07)	(0.03)	(0.00)	
Semi-middle	45	0.91	0.61	5.49	1.17	0.71	1,987.48
(30 – 39 years)		(0.01)	(0.01)	(0.08)	(0.06)	(0.00)	
Middle	70	0.64	1.16	5.19	1.00	0.71	3,269.31
(40 – 50 years)		(0.05)	(0.09)	(0.06)	(0.03)	(0.00)	
Old	95	0.66	1.11	5.09	0.91	0.71	4,435.34
($>$ 50 years)		(0.04)	(0.05)	(0.07)	(0.05)	(0.00)	

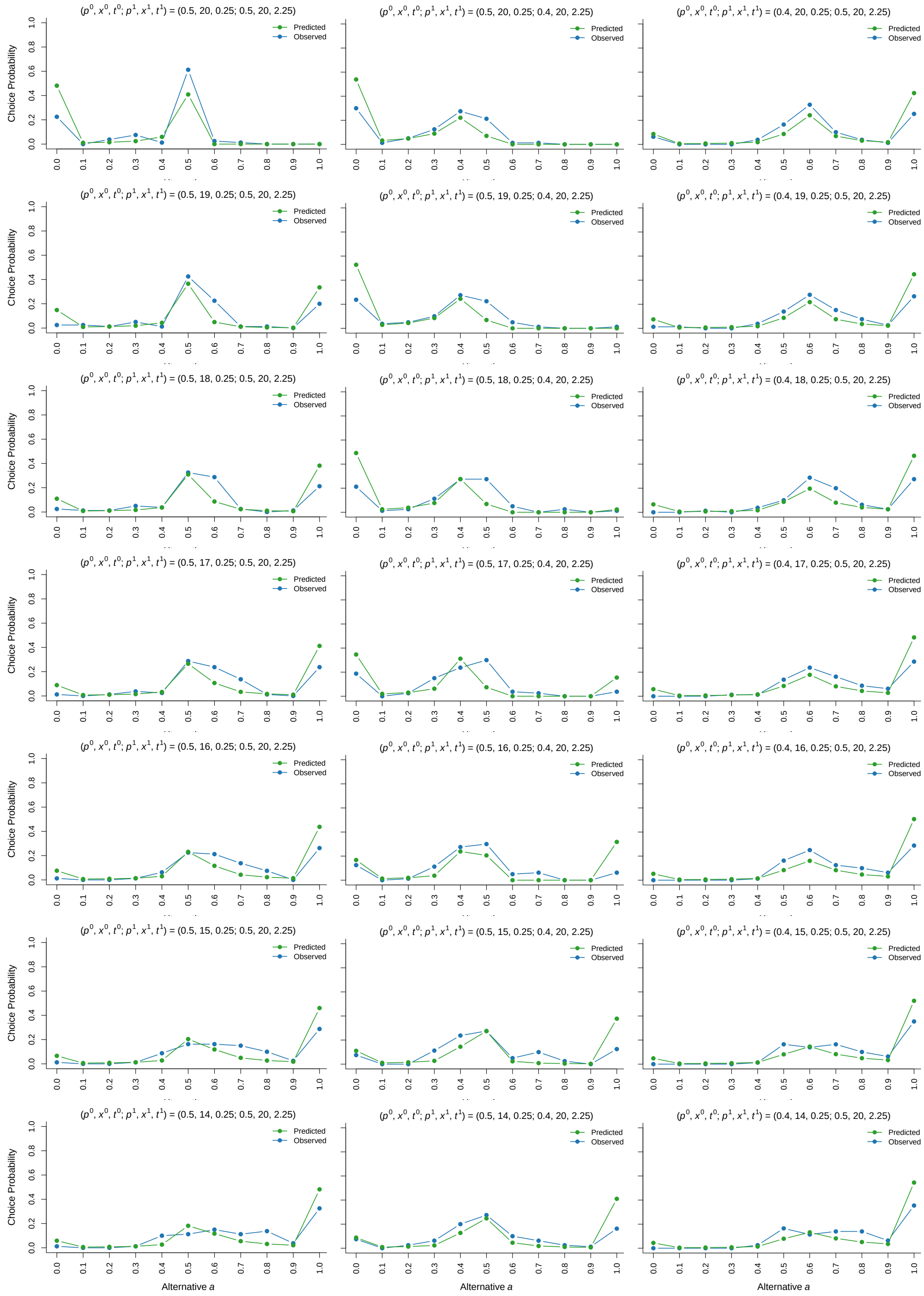
Appendix G. Predicted and observed choice frequencies for all convex menus




(p^0, x^0, t^0; p^1, x^1, t^1) = (0.5, 20, 0.25; 0.5, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.5, 20, 0.25; 0.4, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.4, 20, 0.25; 0.5, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.5, 19, 0.25; 0.5, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.5, 19, 0.25; 0.4, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.4, 19, 0.25; 0.5, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.5, 18, 0.25; 0.5, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.5, 18, 0.25; 0.4, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.4, 18, 0.25; 0.5, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.5, 17, 0.25; 0.5, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.5, 17, 0.25; 0.4, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.4, 17, 0.25; 0.5, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.5, 16, 0.25; 0.5, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.5, 16, 0.25; 0.4, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.4, 16, 0.25; 0.5, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.5, 15, 0.25; 0.5, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.5, 15, 0.25; 0.4, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.4, 15, 0.25; 0.5, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.5, 14, 0.25; 0.5, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.5, 14, 0.25; 0.4, 20, 1.25)
(p^0, x^0, t^0; p^1, x^1, t^1) = (0.4, 14, 0.25; 0.5, 20, 1.25)
Predicted
Observed
Choice Probability
Alternative a



$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (1.0, 20, 0.25; 1.0, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (1.0, 20, 0.25; 0.8, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (0.8, 20, 0.25; 1.0, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (1.0, 19, 0.25; 1.0, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (1.0, 19, 0.25; 0.8, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (0.8, 19, 0.25; 1.0, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (1.0, 18, 0.25; 1.0, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (1.0, 18, 0.25; 0.8, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (0.8, 18, 0.25; 1.0, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (1.0, 17, 0.25; 1.0, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (1.0, 17, 0.25; 0.8, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (0.8, 17, 0.25; 1.0, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (1.0, 16, 0.25; 1.0, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (1.0, 16, 0.25; 0.8, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (0.8, 16, 0.25; 1.0, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (1.0, 15, 0.25; 1.0, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (1.0, 15, 0.25; 0.8, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (0.8, 15, 0.25; 1.0, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (1.0, 14, 0.25; 1.0, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (1.0, 14, 0.25; 0.8, 20, 2.25)
$(p^0, x^0, t^0; p^1, x^1, t^1)$ = (0.8, 14, 0.25; 1.0, 20, 2.25)
Predicted
Observed
Choice Probability
Alternative a


(p0, x0, t0; p1, x1, t1) = (0.5, 20, 0.25; 0.5, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.5, 20, 0.25; 0.4, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.4, 20, 0.25; 0.5, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.5, 19, 0.25; 0.5, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.5, 19, 0.25; 0.4, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.4, 19, 0.25; 0.5, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.5, 18, 0.25; 0.5, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.5, 18, 0.25; 0.4, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.4, 18, 0.25; 0.5, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.5, 17, 0.25; 0.5, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.5, 17, 0.25; 0.4, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.4, 17, 0.25; 0.5, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.5, 16, 0.25; 0.5, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.5, 16, 0.25; 0.4, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.4, 16, 0.25; 0.5, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.5, 15, 0.25; 0.5, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.5, 15, 0.25; 0.4, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.4, 15, 0.25; 0.5, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.5, 14, 0.25; 0.5, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.5, 14, 0.25; 0.4, 20, 2.25)
(p0, x0, t0; p1, x1, t1) = (0.4, 14, 0.25; 0.5, 20, 2.25)
Predicted
Observed
Choice Probability
Alternative a

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

[6] Apesteguia, J. and M.A. Ballester (2017). "Monotone Stochastic Choice Models: The Case of Risk and Time Preferences." *Journal of Political Economy*, 126(1):74–106.

[7] 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.

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

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

[10] 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.

[11] 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.

[12] 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.

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

[14] 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.

[15] 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.

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

[17] 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.

[18] 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.

[19] 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.

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

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

[22] 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.

[23] 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.