Universitat Pompeu Fabra Barcelona
Department of Economics and Business

Economics Working Paper

Series Working Paper No. 1719

Random utility models with ordered types and domains

José Apesteguia and Miguel A. Ballester

April 2020

RANDOM UTILITY MODELS
WITH ORDERED TYPES AND DOMAINS

JOSE APESTEGUIA[†] AND MIGUEL A. BALLESTER[‡]

Abstract. We study random utility models in which heterogeneity of preferences is modeled using an ordered collection of utilities, or types. The paper shows that these models are particularly amenable when combined with domains in which the alternatives of each decision problem are ordered by the structure of the types. We enhance their applicability by: (i) working with arbitrary domains composed of such decision problems, i.e., we do not need to assume any particularly rich data domain, and (ii) making no parametric assumption, i.e., we do not need to formulate any particular assumption on the distribution over the collection of types. We characterize the model by way of two simple properties and show the applicability of our result in settings involving decisions under risk. We also propose a goodness-of-fit measure for the model and prove the strong consistency of extremum estimators defined upon it. We conclude by applying the model to a dataset on lottery choices.

Keywords: Random utility model; Ordered type-dependent utilities; Arbitrary domains; Non-parametric; Goodness-of-fit; Extremum estimators; Decision under risk.

JEL classification numbers: C00; D00

Date: April, 2020.

[*]We thank Christoph Semken for helpful comments and Angelo Gutierrez for outstanding research assistance. Financial support by the Spanish Ministry of Science (PGC2018-098949-B-I00) and Balliol College 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`.

1. Introduction

Consider an ordered collection of type-dependent utilities describing a behavioral trait, such as risk aversion, delay aversion or altruism, where types higher in the order represent more intense expressions of the behavioral trait. In turn, the order of types enables some alternatives to be considered *higher than* simply because they are seen to be preferred by higher types, as implied in the notions of a safer lottery, an earlier stream of payoffs, or a more altruistic distribution. Decision problems in which all maximal alternatives can be ordered by the *higher than* principle are very telling, because, in such situations the lowest types prefer one alternative, and successively higher types prefer successively higher alternatives. Unsurprisingly, most estimation exercises use domains, which we call ordered, composed exclusively of such decision problems. We then introduce heterogeneity by using the random utility model defined over the ordered collection of type-dependent utilities; a model which we call the random type model. The main objective of our paper is to provide characterization and estimation results for the random type model under ordered domains. Notably, and in order to further enhance the applicability of our results, no stage of our analysis requires us to assume domain richness and heterogeneity is fully non-parametrically treated throughout.

The first result of the paper provides a characterization of the choice frequencies that can be generated by the random type model in ordered domains, using two simple properties. The first of these, which we call extremeness after Gul and Pesendorfer (2006), simply states that only those alternatives that are maximal for at least one type in the population can be chosen with strictly positive probability. The second property, which we call monotonicity, is novel and emanates directly from the *higher than* principle. In a nutshell, suppose that the choice of alternative x_1 in decision problem 1 can be rationalized by a type higher than any type rationalizing the choice of alternative x_2 in decision problem 2. Then, the cumulated choice frequency of all alternatives lower than x_1 in decision problem 1 must be larger than the cumulated choice frequency of all alternatives lower than x_2 in decision problem 2.

The simplicity of these two properties proves useful beyond enabling the characterization result. We begin by illustrating the applicability of the model by particularizing it to decisions under risk. In this setting, the most prominent type-dependent utilities are expected utilities, ordered by increasing levels of risk aversion, and these generate a riskiness order on lotteries. We then present a variety of ordered decision problems, representative of the experimental elicitation procedures existing in the literature. Examples include menus in which all lotteries are: (i) defined on the same Marschak-Machina triangle, i.e., involve the same three monetary payoffs (ii) a combination of a prize and a probability, (iii) composed by a good and a bad state of nature, with fixed probabilities q and $1 - q$. Therefore, our characterization result applies to any domain containing any combination of these decision problems, and extremeness and monotonicity are the key tools for scrutinizing whether the observed behavior responds to the random (expected utility) type model.

We then lay down an intuitive goodness-of-fit measure based on the assumed possibility of menu-dependent perturbations of the distribution of types, provided that the magnitude of these perturbations is bounded above. Our goodness-of-fit measure is the minimum perturbation required to explain all the observed choice frequencies, and implicitly defines an extremum estimator. Since our measure can use literally any distance function on the space of probability distributions, we are indeed providing a class of goodness-of-fit measures and estimators and, more importantly, our subsequent analysis shows that any estimator in this class is strongly consistent. That is, as the number of observations per choice problem increases, the estimator converges to the true probability distribution over types. We then expand on the most common estimator, the maximum likelihood estimator.

In the final part of our theoretical analysis, we show that our model is flexible enough to be straightforwardly extended to more general settings. We consider two such exercises. First, we introduce a tremble version of the model, in which alternatives that are never maximal can be selected. We show that the structure of our original results is basically maintained in this richer model. Next, we consider the case in which the analyst may have choice information that varies across subpopulations, as occurs with control-treatment studies or gender/age-specific data. Under these conditions, the question arises as to whether some subpopulations present higher trait levels than others, e.g., whether the type-dependent distribution of the treatment group dominates (or is dominated by) that of the control group. We show how to use extremeness and monotonicity, together with an extra property, to test for these patterns.

We conclude with an empirical illustration of our theoretical results, using an existing experimental dataset on lottery choices. We use the ordered collection of types formed by CRRA expected utilities and argue that the experimental domain is ordered. We then estimate the random type model using maximum likelihood, and briefly discuss some findings.

2. Related Literature

Apesteguia, Ballester and Lu (2017) propose a random utility model built upon an ordered collection of utilities. They analyze the case in which utilities satisfy the well-known single-crossing condition; crucially, with respect to the grand set of alternatives.[1] The current paper requires only a local version of the single-crossing notion by imposing its logic only on alternatives that are available in the same decision problem. This is indeed a major change that substantially enhances the applicability of random models based on ordered collections of utilities. By adopting this weaker notion, we are able to provide characterization and estimation results for arbitrary domains of choice problems, responding directly to the actual requirements of virtually every empirical dataset.[2]

There is a handful of recent empirical papers focused on trying to exploit the single-crossing condition. Barseghyan, Molinari and Thirkettle (2019) use random utility models satisfying the single-crossing condition to provide semi-parametric identification of attention models under risk taking. Chiappori, Salanié, Salanié and Gandhi (2019) also use the single-crossing condition on individual risk preferences in a parimutuel horse-racing setting to establish the equilibrium conditions and ultimately identify the model. Our paper complements this literature, with some differences, such that our model applies to general settings beyond that of decisions under risk as well as to arbitrary ordered domains, and that we provide foundations for the model and its estimation.

[1]The single-crossing property has been intensively studied in economics at least since Mirrlees (1971) and Spence (1974), and has had a large impact in the social, biological and health sciences (see, e.g., Greene and Hensher, 2010). Recently, Filiz-Ozbay and Masatlioglu (2020) study a random choice model using a version of the single-crossing condition in order to model boundedly rational stochastic choice.

[2]Given its importance for the empirical analysis, the literature on the stochastic choice theory is turning to the issue of data requirements. For example, Dardanoni, Manzini, Mariotti and Tyson (2020) study limited-attention stochastic choice models, where the choice domain involves a single menu of alternatives.

A series of applied papers have implemented parametric versions of the random type model in order to estimate specific behavioral traits; most frequently, risk aversion.[3] Barsky, Juster, Kimball and Shapiro (1997) is one of the first examples of the use of this methodology, where the ordered structure of a menu of lotteries is exploited to obtain estimates and perform covariate analysis of risk aversion in the population. Cohen and Einav (2007) use data on deductible choices in auto insurance contracts. As the authors show, any given probability of accident leads to an ordered structure of the menu of deductibles and premiums, thereby facilitating the estimation of risk aversion. Andersson, Holm, Tyran and Wengström (2018) using a balanced experimental design show that cognitive ability affects choice variability rather than risk aversion. Our paper contributes to this literature by providing foundations for a more general, non-parametric, version of the model.

3. A Characterization of Random Type Models

Let $\mathcal{T} = \{1, 2, \ldots, T\}$ be an ordered set of types, with utility functions $\{U_t\}_{t\in\mathcal{T}}$ defined over a space of alternatives X.[4] All our results are ordinal in nature, therefore no parametric assumption is required here; we could equivalently work with the corresponding collection of ordinal preferences. We say that alternative x_h is *higher than* alternative x_l, and write $x_l \lhd x_h$, whenever there exists $t^* \in \mathcal{T} \setminus \{T\}$ such that $U_t(x_l) > U_t(x_h) \Leftrightarrow t \leq t^*$. In words, the relevant economic trait captured by the order of types induces a notion of higher alternatives, i.e., those preferred by high types (with at least type T expressing this preference) but not by low types (with at least type 1 expressing the opposite preference). For instance, types are ordered by risk aversion, delay aversion or altruism and these orders induce the notions of a safer lottery, a less delayed stream of payoffs, or a more altruistic distribution, respectively.

Menus are finite subsets of X and a domain is defined as a finite collection of menus, $\{M_j\}_{j\in\mathcal{J}=\{1,2,\ldots,J\}}$. The alternatives in menu $j \in \mathcal{J}$ that are maximal for at least one

[3]See Coller and Williams (1999) and Warner and Pleeter (2001) for similar estimation exercises within the context of time preferences, or Apesteguia, Ballester and Gutierrez (2020) and Jagelka (2020) for joint estimations of risk and time preferences.

[4]We present the results with finite domains in mind; therefore we can, without loss of generality, assume a finite number of types. The extension to infinite decision problems and sets of types is discussed in Remark 4. For ease of exposition, we also assume no role for indifferences, i.e., utility functions are strict over the domain for which data exist.

type are called *essential*, and denoted by $E_j = \cup_{t=1}^{T} \arg\max_{x \in M_j} U_t(x) \subseteq M_j$. The only assumption in the paper relates to the domain structure. Formally, we assume that the domain is ordered, i.e. for every menu $j \in \mathcal{J}$, $\lhd$ is complete on E_j. Given the structure of the *higher than* notion, the completeness assumption is indeed equivalent to assuming that E_j is linearly ordered by $\lhd$. To see that completeness implies transitivity, let e, e' and e'' all belong to E_j, such that $e \lhd e'$ and $e' \lhd e''$. Then it must be that $U_1(e) > U_1(e') > U_1(e'')$ and, given completeness, it follows that $e \lhd e''$. As a result, we can write $E_j = \{e_{j,1}, e_{j,2}, \ldots, e_{j,\kappa_j}\}$, where $e_{j,1} \lhd e_{j,2} \lhd \cdots \lhd e_{j,\kappa_j}$, with types $\{1, 2, \ldots, t_{j,1}\}$ preferring $e_{j,1}$, types $\{t_{j,1}+1, t_{j,1}+2, \ldots, t_{j,2}\}$ preferring $e_{j,2}$ and so on, ending with types $\{t_{j,\kappa_j-1}+1, t_{j,\kappa_j-1}+2, \ldots, t_{j,\kappa_j} = T\}$ preferring e_{j,κ_j}. In words, in any decision problem, alternatives are ordered such as to reflect the economic trait at hand. In applications, the identification of domains satisfying this property is immediate, as illustrated in Section 4 for the case of decisions under risk.

Let Ψ denote the set of all probability distributions over $\mathcal{T}$. In the random type model (RTM), type $t \in \mathcal{T}$ is realized according to a menu-independent distribution $\psi \in \Psi$, which leads to the choice of alternative $\arg\max_{y \in M_j} U_t(y)$. Our first result works with ideal (infinitely repeated) data and provides necessary and sufficient conditions for these data to be generated by the RTM or, as we put it more simply, to be RTM-rationalizable. Formally, data are modeled by a stochastic choice function, which is a map $p : X \times \mathcal{J} \to [0,1]$ such that, for every $j \in \mathcal{J}$, $p(x,j) > 0$ implies $x \in M_j$, and $\sum_{x \in M_j} p(x,j) = 1$.

Our first property, which we adapt from Gul and Pesendorfer (2006), is an immediate consequence of the optimizing nature of RTMs. It states that only essential alternatives can be observed with strictly positive probability.

Extremeness (EXT): p satisfies EXT if, for every $j \in \mathcal{J}$, $p(x,j) > 0$ implies $x \in E_j$.

Suppose that alternative e in menu j is maximal for a type that is higher than any type for which e' is maximal in menu j'. Our second property states that, under such conditions, the cumulative choice probability of all alternatives lower than e in menu j must be higher than the cumulative choice probability of all alternatives lower than e' in menu j'.

Monotonicity (MON): p satisfies MON if, for every $j, j' \in \mathcal{J}$, $e_{j,k} \in E_j$ and $e_{j',k'} \in E_{j'}$, $t_{j,k} \geq t_{j',k'}$ implies $\sum_{x \in M_j, x \trianglelefteq e_{j,k}} p(x,j) \geq \sum_{x' \in M_{j'}, x' \trianglelefteq e_{j',k'}} p(x',j')$.[5]

These two simple properties characterize RTMs.

Theorem 1. *p is RTM-rationalizable if, and only if, p satisfies EXT and MON.*

Proof of Theorem 1: Note that, under the infinite data assumption, the observed choice frequency of any alternative in any menu must correspond to the sum of masses of all the types for which this alternative is maximal in the menu. Now, first suppose that p is RTM-rationalizable, i.e., there exists an RTM, with distribution ψ, generating the choice probabilities. EXT is a version of the property used by Gul and Pesendorfer (2006) to characterize the random expected utility model and, by the same logic, it is satisfied by RTMs. We now see that MON also holds. Notice that we just showed that $\sum_{x \in M_j, x \trianglelefteq e_{j,k}} p(x,j) = \sum_{x \in E_j, x \trianglelefteq e_{j,k}} p(x,j)$ must hold. Now, the ordered nature of E_j guarantees that the latter is equal to $\sum_{l=1}^{k} p(e_{j,l}, j)$ and, given our notation, this simply corresponds to $\sum_{t=1}^{t_{j,k}} \psi(t)$. Then, MON immediately follows.

Now suppose that p satisfies both EXT and MON. We define a mapping F on a sub-collection $\mathcal{T}' \subseteq \mathcal{T}$ by setting, for every menu $j \in \mathcal{J}$, and $k \in \{1, 2, \ldots, \kappa_j\}$, $F(t_{j,k}) = \sum_{x \in M_j, x \trianglelefteq e_{j,k}} p(x,j) = \sum_{l=1}^{k} p(e_{j,l}, j)$. Notice that MON guarantees that $t_{j,k} = t_{j',k'}$ implies $\sum_{x \in M_j, x \trianglelefteq e_{j,k}} p(x,j) = \sum_{x \in M_{j'}, x \trianglelefteq e_{j',k'}} p(x,j')$, which makes F a single-valued map. Similarly, MON also guarantees that this map is weakly increasing on $\mathcal{T}'$. Notice, moreover, that T must belong to $\mathcal{T}'$, because, for every menu $j \in \mathcal{J}$, $T = t_{j,\kappa_j}$. Notice also that, since E_j is totally ordered by $\triangleleft$, it must be that $1 = \sum_{x \in M_j} p(x,j) \geq \sum_{x \in M_j, x \trianglelefteq e_{j,\kappa_j}} p(x,j) = F(t_{j,\kappa_j}) = F(T) \geq \sum_{x \in E_j} p(x,j)$. EXT guarantees that the latter summand is also 1 and, hence, it must be that $F(T) = 1$. It is then evident that F can be extended to a CDF over the entire set of types $\mathcal{T}$, which, for simplicity, we denote by F. Letting $F(0) = 0$, it is immediate to see that the induced probability distribution $\psi(t) = F(t) - F(t-1)$ rationalizes p. ■

The proof is constructive. Intuitively, type $t_{j,k}$ is the last type to select $e_{j,k}$ in menu j. The order structure guarantees that any lower type must choose $e_{j,l}$, with $l \leq k$ and any higher type must choose $e_{j,h}$ with $h > k$. Hence, we can define the CDF over the types as $F(t_{j,k}) = \sum_{x \in M_j, x \trianglelefteq e_{j,k}} p(x,j) = \sum_{l=1}^{k} p(e_{j,l}, j)$, and MON guarantees that it is

[5] As usual, $x \trianglelefteq e$ whenever $x \triangleleft e$ or $x = e$.

a weakly increasing single-valued map. Also, since EXT guarantees that only essential alternatives have strictly positive probability, it must be that $F(T) = 1$.

Remark 1. If data are available only on a small number of menus, the CDF will be identified only for a number of types, with an upper bound given by $\sum_{j=1}^{J} |E_j| - J$. This upper bound is tight if the threshold types $t_{j,k}$ of essential alternatives all differ across menus. In an experimental setting, the analyst may obviously design menus that help to identify F wherever desired. It is obvious that the model is fully identified whenever, for every $t \in \mathcal{T}$, there exists a menu $j \in \mathcal{J}$ and $k \in \{1, 2, \dots, \kappa_j\}$ such that $t_{j,k} = t$. For instance, the simplest way of doing this is to observe the binary menu $M_j = \{x, y\}$, with $x \lhd y$ and such that $t^* = t$.

Remark 2. The assumption on the completeness of $\lhd$ on each E_j could be substantially relaxed. The logic of our results only requires the following weaker version, where a domain is said to be an interval-domain if, for every $j \in \mathcal{J}$ and $e \in M_j$: $e = \arg\max_{x \in M_j} U_{t_l}(x) = \arg\max_{x \in M_j} U_{t_h}(x)$ implies that $e = \arg\max_{x \in M_j} U_{t_m}(x)$ for every $t_l < t_m < t_h$. That is, if an alternative is maximal for two types, it is also maximal for the intermediate types. Under this interval condition, we can order the essential alternatives of a menu as $e_{j,1}, e_{j,2}, \dots, e_{j,\kappa_j}$, with types $\{1, 2, \dots, t_{j,1}\}$ preferring $e_{j,1}$, types $\{t_{j,1} + 1, t_{j,1} + 2, \dots, t_{j,2}\}$ preferring $e_{j,2}$ and so on, ending with types $\{t_{j,\kappa_j - 1} + 1, t_{j,\kappa_j - 1} + 2, \dots, t_{j,\kappa_j} = T\}$ preferring e_{j,κ_j}. Importantly, notice that this order does not necessarily require alternatives to be related by $\lhd$, thus showing that this is a weaker assumption.

Remark 3. We have opted for the practical approach, in which an arbitrary collection of ordered menus is considered. Interestingly, our results are also informative about domains in which the order property does not necessarily hold, but where a richness condition is met instead. Suppose a menu that is not ordered. That is, the set of types is partitioned into the largest possible intervals, where each interval has the same maximal alternative, but contrary to the structure in ordered domains, the same alternative corresponds to two of these intervals. The richness condition required for our results calls for the existence of a $\lhd$-*replica* of this menu, i.e., an ordered menu with a partition of types coinciding exactly with that of the non-ordered menu. In domains containing $\lhd$-*replicas*, a stochastic choice function is RTM-rationalizable if, and only if, (i) EXT and MON hold for ordered menus, and (ii) for any non-ordered menu $j \in \mathcal{J}$

and every e in menu j, $p(e, j)$ is equal to the choice probability of all the alternatives that replace e in the $\triangleleft$-replica of menu j.

Remark 4. The entire analysis immediately extends to RTMs defined on infinite type-spaces $\{U_\omega\}_{\omega\in\Omega}$ and infinite ordered menus, under the usual continuity and measurability assumptions. We can reformulate EXT by stating that only subsets of alternatives that are maximal for a measurable set of types can have strictly positive choice probability. For MON, denote by $\omega_{j,e}$ the largest type for which essential alternative e is maximal in menu j. The continuous version of MON requires larger values of $\omega_{j,e}$ to be associated with larger choice probabilities for the measurable set of alternatives that are lower than e.

4. Applications to Decision under Risk

In this section, we discuss the case in which X is a set of monetary lotteries. We consider the most natural ordered set of types in this setting; namely, a collection of expected utilities, $\{EU_t\}_{t\in\mathcal{T}=\{1,2,\dots,T\}}$, in increasing order of risk aversion, i.e., ordered by increasing concavity of monetary utility. Thus, the induced relation $\triangleleft$ represents the notion of a safer lottery. Theorem 2 below proves that the most standard menus of lotteries in experimental studies are ordered. Then, whenever the domain is formed by any combination of such menus, we can use MON and EXT to determine whether behavior corresponds to the RTM based on expected utility.

We consider three possible sets of menus. The first are Marschak-Machina menus, in which all the lotteries belong to the same Marschak-Machina triangle; i.e., they are all composed of the same three monetary payoffs. The second are basic menus, in which all lotteries are basic, i.e., they all comprise a (possibly different) strictly positive prize and a (possibly different) probability of receiving this prize. The third are binary-state menus, in which all lotteries entertain a good-state payoff and a bad-state payoff with fixed probabilities.[6] We can now establish the following result.

Theorem 2. *Let $\{EU_t\}_{t\in\mathcal{T}=\{1,2,\dots,T\}}$ be ordered by increasing risk aversion. The essential lotteries of any Marschak-Machina menu, basic menu or binary-state menu are*

[6] Notice that this formulation is a generalization of widely used experimental decision problems, such as those in Holt-Laury (2002) and Choi, Fisman, Gale, and Kariv (2007).

completely ordered by $\triangleleft$*. Hence, in any domain formed by a combination of Marschak-Machina menus, basic menus or binary-state menus, EXT and MON are necessary and sufficient conditions for RTM-rationalizability.*

Proof of Theorem 2: First consider a menu of lotteries in a Marschak-Machina triangle and let $e = (q_e^B, 1 - q_e^B - q_e^W, q_e^W)$ and $e' = (q_{e'}^B, 1 - q_{e'}^B - q_{e'}^W, q_{e'}^W)$ be two essential lotteries in this menu, where lotteries are referred to by their probabilties on the different prizes $B > I > W$. Since the lotteries are essential, it must be the case that none of them first-order stochastically dominates the other and hence we can assume, w.l.o.g., that $q_e^B > q_{e'}^B$ and $q_e^W < q_{e'}^W$. It is immediately evident that e will be preferred to e' if, and only if, the slope of the linear indifference curve of the EU type is below $\frac{q_e^B - q_{e'}^B}{q_{e'}^W - q_e^W}$.[7] Since types are ordered by increasing slopes and both lotteries are essential, it must be that $e \triangleleft e'$ and so we are done. Similarly, consider a menu of basic lotteries and let $e = (q, n)$ and $e' = (q', n')$ be two essential lotteries, where lotteries are referred to by the probability and monetary size of the strictly positive prize. Again, since the lotteries are essential, they cannot be related by first-order stochastic dominance, thus guaranteeing that one of the lotteries, say (q, n), has a strictly higher payoff, while the other has a strictly higher probability. By normalizing $u_t(0) = 0$, it is then evident that (q, n) is preferred to (q', n') if, and only if, $\frac{u_t(n)}{u_t(n')} \geq \frac{q'}{q}$, which is true if, and only if, the curvature of u_t is sufficiently low. Since types are ordered by increasing curvature and both lotteries are essential, it must be that $(q, n) \triangleleft (q', n')$, thus proving that essential alternatives are ordered by $\triangleleft$. Finally, consider a binary-state menu and let $e = (n_1, n_2; q)$ and $e' = (n_1', n_2'; q)$ be two essential alternatives, where notation is used to describe the good-state payoff, that of state 1; the bad-state payoff, that of state 2; and the (common) probability of state 1. Absence of first-order stochastic dominance guarantees that $q \neq \{0, 1\}$ and that payoffs can be ordered, w.l.o.g, as $n_1 > n_1' > n_2' > n_2$. We can normalize the monetary utilities to $u_t(n_1) = 1$ and $u_t(n_2) = 0$. Suppose then that e' is preferred to e by some type $t \in \mathcal{T}$. Then $q \leq q u_t(n_1') + (1 - q) u_t(n_2')$, i.e., $\frac{1 - u_t(n_1')}{u_t(n_2')} \leq \frac{1-q}{q}$. Let type $s > t$ be more risk averse than type t. Since type t is indifferent between receiving n_2' with certainty and receiving n_1 and n_2 with probabilities $u_t(n_1')$ and $1 - u_t(n_2')$, the more risk-averse type must strictly prefer to receive n_2' with certainty, and hence $u_s(n_2') > u_t(n_2')$. Similar

[7]Notice that we can normalize $u_t(W) = 0$ and $u_t(B) = 1$ and hence, any variation in expected utility can be related to the value $u_t(I) = \omega \in (0, 1)$. The slope of the linear indifference curves is a monotone transformation of ω.

reasoning proves that $u_s(n_1') > u_t(n_1')$. Hence, $\frac{1-u_s(n_1')}{u_s(n_2')} < \frac{1-u_t(n_1')}{u_t(n_2')} \leq \frac{1-q}{q}$ and e' must also be preferred to e by the more risk-averse type s. Essential alternatives in binary-state menus are thus completely ordered by $\triangleleft$. Having therefore proved that the three classes of menus are ordered, the second claim follows immediately from Theorem 1.■

Remark 5. Gul and Pesendorfer's (2006) random expected utility assumes the existence of a probability distribution over the entire family of expected utilities. In our approach, we assume an ordered family of expected utilities. Notice that in some spaces, such as the one in which X corresponds to a Marschak-Machina triangle, all expected utilities can be ordered by risk aversion; hence, EXT and MON provide an alternative characterization of random expected utility. In other spaces, not all expected utilities can be ordered by risk aversion; but the standard estimation procedure is to consider an ordered family, such as those given by CARA and CRRA utilities. In these cases, we can use EXT and MON to characterize the corresponding RTMs.

Remark 6. In relation to Remark 3, it is interesting to note that we can use any Marschak-Machina triangle to construct $\triangleleft$-replicas of any menu of lotteries. To see this, consider the relevant partition of types in a menu $j \in \mathcal{J}$, $\{1, 2, \ldots, t_{j,1}\}, \{t_{j,1}+1, t_{j,1}+2, \ldots, t_{j,2}\}, \ldots, \{t_{j,\kappa_j-1}+1, t_{j,\kappa_j-1}+2, \ldots, t_{j,\kappa_j} = T\}$, where the corresponding maximal alternatives $e_{j,k}$ are not necessarily ordered. Hence, for a given Machina-Marschak triangle, denote by $m(t)$ the (strictly positive) slope of the indifference curves of type t. Notice that $m(t)$ must be strictly increasing in t, because types are ordered by risk aversion. Now, construct the $\triangleleft$-replica of menu j as follows. Let $x_1 = (\frac{1}{2}, 0, \frac{1}{2})$ and, for $k \in \{2, \ldots, \kappa_j\}$, let $x_k = x_{k-1} + (-\alpha_{k-1}, \alpha_{k-1} + \beta_{k-1}, -\beta_{k-1})$, with $m(t_{j,k-1}) < \frac{\alpha_{k-1}}{\beta_{k-1}} < m(t_{j,k-1}+1)$ and such that $\sum_{k=2}^{\kappa} \alpha_k$ and $\sum_{k=2}^{\kappa} \beta_k$ are both smaller than $\frac{1}{2}$, thus guaranteeing that all vectors correspond to lotteries in the triangle. It is then evident that lottery x_k is maximal for, and only for, types $\{t_{j,k-1}+1, \ldots, t_{j,k}\}$, as desired. Therefore, a complete characterization of any RTM based on expected utilities can be obtained by assuming EXT and MON on menus ordered by $\triangleleft$ (say, those of one Machina-Marschak triangle), together with the replica property of Remark 3.

5. ϵ-rationalizability and Strongly Consistent Estimators

In this section, we present several results relating to the estimation of RTMs. We start by providing an intuitive generalization of the RTM-rationalizability notion, which takes into account menu-dependent perturbations. More concretely, we contemplate the possibility that the choice probabilities are generated by menu-dependent distributions of types ψ_j. We say that data are d_ϵ-rationalizable if there exists a distribution $\psi \in \Psi$ such that $\max_{j\in\mathcal{J}} d(\psi, \Psi_j) \leq \epsilon$, where d is a distance function on Ψ, and $\Psi_j \subseteq \Psi$ is the set of all distributions that generate the choice probabilities of menu $j \in \mathcal{J}$.[8] Trivially:

Corollary 1. *p is RTM-rationalizable if, and only if, p is d_0-rationalizable. Furthermore, every p satisfying EXT is d_ϵ-rationalizable for some value of ϵ.*

The first part of the result shows that d_ϵ-rationalizability constitutes a generalization of RTM-rationalizability in which there exist no menu-dependent perturbations. The second part of the result simply states that, with ϵ large enough, any behavior that satisfies EXT can be explained by the menu-dependent variant of the model.[9] Hence, the question arises as to which minimal value of ϵ guarantees d_ϵ-rationalizability, thus providing a natural goodness-of fit-measure for the model. This is in line with Afriat's (1973) goodness-of-fit measure in deterministic consumer settings.[10]

We now consider the practical case of finite data and the estimation exercise. Formally, data are represented by way of a map $z : X \times \mathcal{J} \rightrightarrows \mathbb{Z}_+$, with $z(x, j) = 0$ whenever $x \notin M_j$, describing the number of times that alternative x has been chosen from menu $j \in \mathcal{J}$. For every menu $j \in \mathcal{J}$, we denote by z_j the vector describing the observed choices in this menu, and by $Z_j = \sum_{x\in M_j} z(x, j) > 0$ the total number of observations regarding this menu. Observed choice frequencies in each menu $j \in \mathcal{J}$ are therefore $\frac{z_j}{Z_j}$. When the choice frequencies satisfy EXT and MON, the first part of Corollary 1 guarantees the existence of an RTM generating a stochastic choice map equal to the observed choice frequencies. With data not satisfying MON, the second part of Corollary 1 guarantees the existence of an RTM that minimizes the ϵ-perturbation required to accommodate all the data. In essence, this analysis provides an intuitive estimator of the model, denoted as $\widehat{\psi}_d$, which can be shown to be strongly consistent.

Theorem 3. *$\widehat{\psi}_d$ is strongly consistent.*

[8]We define, as usual, the distance of ψ to Ψ_j as the minimum distance of ψ to any element of Ψ_j. In the case of Ψ_j being empty, we say that the distance is infinite.

[9]Section 7 studies the case where EXT is not required.

[10]See Apesteguia and Ballester (2020) for other goodness-of-fit measures in stochastic settings.

Proof of Theorem 3: Consider an RTM with probability distribution $\psi \in \Psi$ and the domain of menus $\{M_j\}_{j \in \mathcal{J}=\{1,2,\ldots,J\}}$. Trivially, the stochastic choice function implied by ψ is equivalent to the one implied by the restricted RTM defined on the support of ψ, ordered as in $\mathcal{T}$. Also, since $x \in M_j \setminus E_j$ will never be chosen, the choice probabilities generated by the restricted RTM are equivalent to those generated by the same restricted RTM on the domain formed by the corresponding menus of essential alternatives $\{E_j\}_{j \in \mathcal{J}=\{1,2,\ldots,J\}}$. Hence, we can assume, w.l.o.g., that ψ has full support and that $M_j = E_j$ for all $j \in \mathcal{J}$.

Let Θ_j be the space of all strictly positive (multinomial) probability distributions on menu j and $\Theta = \Theta_1 \times \Theta_2 \times \cdots \times \Theta_J$.[11] Any RTM distribution of types ψ corresponds to one such multinomial distribution $\theta = (\theta_1, \theta_2, \ldots, \theta_J) \in \Theta$, where $\theta_j(e)$ is equal to the mass of all types, according to ψ, for which essential alternative e is maximal in menu j. Notice that not all the elements of Θ correspond to an RTM, since some multinomial distributions may, obviously, fail to satisfy the additional structure imposed by MON. Formally, taking into consideration the full support, the multinomial distributions associated with the RTM must be those for which $t_{j,k} \geq t_{j',k'} \Leftrightarrow \sum_{l=1}^{k} \theta_j(e_{j,l}) \geq \sum_{l'=1}^{k'} \theta_{j'}(e_{j',l'})$.

Let us now consider an increasing number of observations for each menu. The multinomial structure guarantees, almost surely, that the strict part of MON must hold across observed frequencies, guaranteeing that the smallest frequencies will correspond to all pairs (j,k) such that $t_{j,k} = 1$, the next larger frequencies will correspond to all pairs (j,k) such that $t_{j,k} = 2$, etc, with the largest frequencies corresponding to all pairs (j,k) such that $t_{j,k} = T - 1$.[12] Importantly, notice that these frequencies may not be RTM-rationalizable, because two pairs associated with the same type t may, due to sampling, have slightly different frequencies. Given the structure of the d-extremum estimator, it is evident that, in all these cases, the estimator will construct the CDF of type t by selecting a probability value within the interval determined by all the type t frequencies.[13] As the intervals shrink and, almost surely, converge to the corresponding value, the estimator converges almost surely to ψ. ■

[11]Notice that, with a slight modification of the structure, this simply corresponds to the space of all stochastic choice functions with strictly positive probabilities defined on $\{M_j\}_{j \in \mathcal{J}=\{1,2,\ldots,J\}}$.

[12]Note that, when $t_{j,k} = T$, the associated probabilities are always 1.

[13]The exact value will obviously depend on the distance function d being used.

Theorem 3 shows that every d-extremum estimator is strongly consistent. The proof is based on the fact that stochastic choice functions are collections of multinomials (one for each menu), and that the collections generated by RTMs have a clear structure determined by MON. We then argue that, when the number of observations increases, the choice data generated by an RTM must almost surely satisfy the strict part of MON. The estimator will reflect this, and simply equalize the remaining frequencies, which should be equal. Thus, the result follows from the known fact that all these frequencies must converge to the values determined by the RTM.

6. Maximum Likelihood Estimation

We now briefly comment on the most common estimator, the maximum likelihood (ML) estimator, denoted by ψ^{ML}.[14]

Theorem 4. *Suppose that the likelihood of ψ^{ML} is strictly positive. Then, for every pair of types t, s in the support of ψ^{ML}, it is the case that:*

$$\sum_{(j,k):t_{j,k-1}<t\leq t_{j,k}} \frac{z(e_{j,k},j)}{\sum_{t=t_{j,k-1}+1}^{t_{j,k}} \psi^{ML}(t)} = \sum_{(j,k):t_{j,k-1}<s\leq t_{j,k}} \frac{z(e_{j,k},j)}{\sum_{t=t_{j,k-1}+1}^{t_{j,k}} \psi^{ML}(t)}.$$

Proof of Theorem 4: It is immediately evident that the likelihood can be strictly positive if, and only if, z satisfies EXT; hence, we can assume, w.l.o.g., that z is restricted to menus E_j. In an RTM with distribution ψ, the probability of observing $e_{j,k}$ in menu j is given simply by the mass of types for which $e_{j,k}$ is maximal in menu j. We know that this is $\psi(\{t_{j,k-1}+1, t_{j,k-1}+2, \dots, t_{j,k}\}) = \psi(t_{j,k-1}+1)+\cdots+\psi(t_{j,k})$. We can write the ML estimator as the real-vector ψ that maximizes the total log-likelihood, subject to the constraint that ψ is a probability distribution. That is, ψ^{ML} must coincide with the real vector that maximizes $\sum_{(j,k)} z(e_{j,k},j)\log(\psi(t_{j,k-1}+1)+\cdots+\psi(t_{j,k}))$ subject to the constraint that $\sum_{t\in\mathcal{T}}\psi(t) = 1$ and to positivity constraints. Hence, for any two types t, s in the support of ψ^{ML}, the first-order condition of the corresponding Lagrangian must take the form $\sum_{(j,k):t_{j,k-1}<t\leq t_{j,k}} \frac{z(e_{j,k},j)}{\psi^{ML}(t_{j,k-1}+1)+\cdots+\psi^{ML}(t_{j,k})} = \sum_{(j,k):t_{j,k-1}<s\leq t_{j,k}} \frac{z(e_{j,k},j)}{\psi^{ML}(t_{j,k-1}+1)+\cdots+\psi^{ML}(t_{j,k})}$, thus concluding the proof. ■

[14] It is well-known that this estimator converges to the one that minimizes the Kullback-Leibler divergence. Since the result in the previous section applies equally to both divergence and distance measures, the ML estimator ψ^{ML} is strongly consistent.

Any observation that could potentially be explained by type t is relevant for determining the mass of t, and hence must be counted using the conditional probability formula, i.e., taking into account the fact that the type must cover the entire interval revealed by the observation. For instance, an observation showing that the type is exactly t should be weighted much more than an observation showing that the parameter matches a large superset of t. The ML estimator equalizes this value across all relevant types in the support.

Example 1. Consider $T = 5$, and $X = \{a, b, c, d\}$ with $a \lhd b \lhd c \lhd d$. The domain is composed of menus of (essential) alternatives $E_1 = \{a, b, d\}$, $E_2 = \{b, c, d\}$ and $E_3 = \{b, d\}$ for which we know that $t_{1,1} = 1$, $t_{2,1} = 2$, $t_{1,2} = t_{3,1} = 3$, $t_{2,2} = 4$ and $t_{1,3} = t_{2,3} = t_{3,2} = 5$. This example corresponds to a subset of lotteries and CRRA expected utilities analyzed in our empirical study of Section 8.[15] The choice of a from menu 1 is the only choice revealing exclusively type $\{1\}$. Similarly, $(b, 1), (b, 2), (c, 2), (d, 2)$ and $(b, 3)$ are the only choices associated with the intervals of types $\{2, 3\}, \{1, 2\}, \{3, 4\}, \{5\}$ and $\{1, 3\}$, respectively. The choices of d from menus 1 and 3 reveal the same information; namely, that the type belongs to $\{4, 5\}$. Hence, we can write the log-likelihood as $z(a, 1) \log(\psi(1)) + z(b, 1) \log(\psi(2) + \psi(3)) + z(b, 2) \log(\psi(1) + \psi(2)) + z(c, 2) \log(\psi(3) + \psi(4)) + z(d, 2) \log(\psi(5)) + z(b, 3) \log(\psi(1) + \psi(2) + \psi(3)) + (z(d, 1) + z(d, 3)) \log(\psi(4) + \psi(5))$, with the extra condition that ψ is a probability distribution. Consider, for instance, the actual data from the experiment, $z(a, 1) = 12$, $z(b, 1) = 5$, $z(b, 2) = 14$, $z(c, 2) = 8$, $z(d, 2) = 15$, $z(b, 3) = 41$, $z(d, 1) = 19$ and $z(d, 3) = 46$. Theorem 4 can be applied immediately to obtain the ML estimator as $\psi^{ML}(1) = .33$, $\psi^{ML}(2) = .05$, $\psi^{ML}(3) = .09$, $\psi^{ML}(4) = .12$ and $\psi^{ML}(5) = .41$.

Remark 7. In many experimental studies, the data exclusively comprise choices over binary menus, $\{M_j = E_j = \{e_{j,1}, e_{j,2}\}\}_{j=1}^{J}$.[16] Based on our previous discussions, we can assume, w.l.o.g., that the ordered collection of types is composed of types $\{t_{1,1}, t_{2,1}, \dots, t_{J,1}, T\}$. We can then provide an immediate algorithm for the computation of the ML estimator. Denote $\rho(\underline{j}, \overline{j}) = \frac{\sum_{j=\underline{j}}^{\overline{j}} z(e_{j,1}, j)}{\sum_{j=\underline{j}}^{\overline{j}} Z_j}$. Then, define $j_0 = 0$ and $F_0 = 0$ and, recursively, j_n (with $j_N = J$ obviously being the last element of the sequence) as

[15] Namely, alternatives a, b, c and d correspond to lotteries l_3, l_4, l_5 and l_1, respectively, and the types correspond to the CRRA relevant types.

[16] We assume that menus are formed by essential alternatives; as, otherwise, only one alternative would be chosen with probability one.

the largest integer j, with $j_{n-1} < j \leq J$ minimizing $\rho(j_{n-1}+1, j)$, and let F_n be equal to the corresponding minimized argument. Then, it can be seen that the ML estimator is:

$$\psi^{ML}(t) = \begin{cases} F_n - F_{n-1} & \text{if } t = t_{1,j_{n-1}} + 1 \\ 1 - F_N & \text{if } t = T \\ 0 & \text{otherwise.} \end{cases}$$

For example, consider $T = 4$, $X = \{a, b, d\}$ with $a \triangleleft b \triangleleft d$, and the menus of (essential) alternatives $\bar{E}_1 = \{a, b\}$, $\bar{E}_2 = \{a, d\}$ and $E_3 = \{b, d\}$, with $t_{1,1} = 1$, $t_{2,1} = 2$, $t_{3,1} = 3$. This, again, corresponds to the case analyzed in Section 8, with data $z(a, 1) = 37$, $z(b, 1) = 50$, $z(a, 2) = 35$, $z(c, 2) = 52$, $z(b, 3) = 41$ and $z(d, 3) = 46$. Since $\rho(1, 1) = .43$, $\rho(1, 2) = .41$ and $\rho(1, 3) = .43$, it must be that $j_1 = 2$ and $F_1 = .41$. Now, $\rho(3, 3) = .47$; hence, $j_2 = 3$ and $F_2 = .47$. Thus, the ML estimator is simply $\psi^{ML}(1) = .41$, $\psi^{ML}(2) = 0$, $\psi^{ML}(3) = .06$ and $\psi^{ML}(4) = .53$.

7. Extensions

7.1. **Tremble.** Random type models cannot explain the choice of non-essential alternatives, such as first-order stochastically dominated lotteries. In some cases, however, a significant number of non-essential alternatives are chosen and the analyst may therefore wish to extend RTMs to accommodate this behavioral regularity. A convenient way of doing this is to extend RTMs to include trembling behavior, thus giving rise to what we call the random type model with tremble (RTMT). We say that p is RTMT-rationalizable if there exists a probability distribution over the set of types, ψ, and a tremble value, $\lambda \in [0, 1)$, such that: (i) whenever $M_j = E_j$, choices are determined according to ψ and (ii) whenever $M_j \neq E_j$, choices are determined according to ψ with probability $(1-\lambda)$; and uniformly randomly on $M_j \setminus E_j$, otherwise.[17] RTMT is a simple model, and its characterization follows immediately from the analysis in Theorem 1.

Corollary 2. *p is RTMT-rationalizable if, and only if, p satisfies*

(1) *EXT*: For every $j, j' \in \mathcal{J}$, $\sum_{x \notin E_j} p(x, j) = \sum_{x \notin E_{j'}} p(x, j')$, and*
(2) *MON*: For every $j, j' \in \mathcal{J}$, $e_{j,k} \in E_j$ and $e_{j',k'} \in E_{j'}$, $t_{j,k} \geq t_{j',k'}$ implies that $\frac{\sum_{x \in M_j, x \trianglelefteq e_{j,k}} p(x,j)}{\sum_{x \in E_j} p(x,j)} \geq \frac{\sum_{x' \in M_{j'}, x' \trianglelefteq e_{j',k'}} p(x',j')}{\sum_{x \in E_{j'}} p(x,j')}$.*

[17] We select the simplest tremble model for purposes of illustration. Similar techniques can be used with many other trembling mechanisms.

Corollary 2 provides a simple test for the RTMT model. The probability of a mistake is constant across menus and hence a simple reformulation of EXT must hold. Similarly, once the observed probabilities have been normalized by the possible trembling in the menu, MON must hold. The practical implementation of this simple model is illustrated in our empirical application, described below.

7.2. **Subpopulations.** In many applications, the analyst envisions a model in which different subpopulations behave differently, and wishes to establish a relationship across the RTMs of these subpopulations. We illustrate this idea with the intuitive case in which an ordered characteristic (say age or income, or a gender dummy or one of a set of treatment dummies) is such that higher subpopulations have RTM distributions that are first-order stochastically dominating, i.e., they are skewed towards higher types. Formally, let $\mathcal{G} = \{1, \ldots, G\}$ be a partition of the population into G groups. We say that the collection of RTMs $\{\psi_g\}_{g\in\mathcal{G}}$ is monotone-in-characteristics if, for every $g, g' \in \mathcal{G}$, with $g < g'$, and every $t \in \mathcal{T}$, $\sum_{s=1}^{t} \psi_g(s) \geq \sum_{s=1}^{t} \psi_{g'}(s)$. Denote by p_g the stochastic choice function of subpopulation g. Then,

Corollary 3. *The collection of stochastic choice functions $\{p_1, p_2, \ldots, p_G\}$ can be represented by a monotone-in-characteristics collection of RTMs if, and only if, $\{p_1, p_2, \ldots, p_G\}$ satisfies*

1. *For every $g \in \mathcal{G}$, p_g satisfies EXT and MON, and*
2. *For every $g, g' \in \mathcal{G}$, with $g < g'$, every $j \in \mathcal{J}$ and every $e \in E_j$, we have $\sum_{x\in M_j, x\trianglelefteq e} p_g(x, j) \geq \sum_{x\in M_j, x\trianglelefteq e} p_{g'}(x, j)$.*

Clearly, the stochastic choice function of each of the subgroups must, by Theorem 1, satisfy EXT and MON. Then, a distribution with more mass on higher types will generate higher choice probabilities for alternatives in the menu that are higher.

8. Empirical Illustration

We now illustrate our framework and results using the experimental dataset analyzed in Apesteguia and Ballester (2020). This dataset involves 87 UCL undergraduates choosing lotteries from menus of lotteries of sizes 2, 3, and 5. Concretely, there were nine equiprobable monetary lotteries, described in Table 1. Each of the participants faced a total of 108 different menus of lotteries, including all 36 binary menus; 36 menus with 3 alternatives out of the possible 84; and another 36 menus with 5 alternatives out of the possible 126. Random individual processes, without replacement, were used to select the menus of 3 and 5 alternatives to determine the order of presentation, and to locate the lotteries on the screen and the monetary prizes within a lottery. There were two treatments, NTL and TL. Treatment NTL was a standard implementation, that is, the choice was not time-constrained. In treatment TL, subjects had to select a lottery within 5, 7 and 9 seconds from the menus with 2, 3, and 5 alternatives, respectively.[18]

Table 1. Lotteries

$l_1 = (17)$	$l_4 = (30, 10)$	$l_7 = (40, 12, 5)$
$l_2 = (50, 0)$	$l_5 = (20, 15)$	$l_8 = (30, 12, 10)$
$l_3 = (40, 5)$	$l_6 = (50, 12, 0)$	$l_9 = (20, 12, 15)$

The set of types is defined by adopting expected utility and CRRA, i.e., we use monetary utilities $\frac{x^{1-\omega}}{1-\omega}$, whenever $\omega \neq 1$, and $\log x$ for $\omega = 1$, where ω represents the risk-aversion coefficient.[19] Given the decision problems considered in the experiment, there are exactly 30 relevant types, which are ordered by increasing risk aversion and described in Columns 1-3 of Table 2. Column 1 reports the type number and Column 2 reports the upper bound of ω corresponding to that type. Finally, Column 3 reports the preference over the lotteries of the respective type. We describe the preference of the first type and then specify the pair(s) of alternatives that flip from the previous type. Given the types, it is easy to see that all menus of lotteries are ordered by $\triangleleft$ and, hence, our basic framework can be used.[20]

There are several menus with non-essential alternatives, such as, for example, the binary menu $\{l_5, l_9\}$, where l_5 first-order stochastically dominates l_9. We observe significant choice probabilities for non-essential alternatives (that is, EXT clearly fails), and thus prefer to use the tremble version of the model, RTMT. We then analyze the ML estimator of the data. In Columns 4-7 in Table 2, we report the estimated densities of the RTMT using all the data (Column 4); the treated data (Column 5); the

[18]Experimental payoffs were determined by randomly selecting one menu, and the subject was paid according to his or her choice from that menu.

[19]Since lotteries l_2 and l_6 involve 0 payoffs, we assume a small fixed positive background consumption.

[20]Note that some of these menus do not fall within the classes covered in Theorem 2, thus reinforcing our claim regarding the wide applicability of the setting studied in this paper.

TABLE 2. Preferences and Estimation Results

Preferences			Estimated RTMTs			
ID	ω		All Obs	TL	NTL	Only Binary
1	-4.148	2-6-3-7-4-8-5-9-1	0.1166	0.1173	0.1158	0.2197
2	-0.518	$(1, 9)$	0	0	0	0
3	-0.313	$(3, 6)$	0.1009	0.0937	0.1081	0.0673
4	-0.083	$(4, 7)$	0	0	0	0
5	0.065	$(5, 8)$	0	0	0	0
6	0.154	$(4, 6)$	0	0	0	0
7	0.209	$(1, 8)$	0	0	0	0
8	0.229	$(3, 2)$ and $(7, 6)$	0.0561	0.0713	0.0289	0
9	0.258	$(5, 6)$	0	0	0.031	0
10	0.262	$(1, 6)$	0.0274	0.009	0.0263	0
11	0.273	$(5, 7)$	0	0	0	0
12	0.339	$(4, 2)$ and $(8, 6)$	0	0	0	0
13	0.342	$(1, 7)$	0	0	0	0
14	0.358	$(5, 2)$ and $(9, 6)$	0	0	0	0
15	0.363	$(1, 2)$	0	0	0	0
16	0.374	$(7, 2)$	0	0	0	0
17	0.408	$(8, 2)$	0	0	0	0
18	0.443	$(9, 2)$	0.0915	0.067	0.1172	0.1545
19	0.516	$(4, 3)$ and $(8, 7)$	0.0079	0.013	0	0
20	0.607	$(6, 2)$	0	0	0	0
21	0.652	$(5, 3)$ and $(9, 7)$	0	0	0	0
22	0.808	$(1, 3)$	0	0	0	0
23	0.844	$(8, 3)$	0	0	0.009	0
24	1.000	$(9, 3)$	0.1388	0.1482	0.1182	0.0645
25	1.124	$(5, 4)$ and $(9, 8)$	0	0	0	0
26	1.309	$(1, 4)$	0	0	0	0
27	2.000	$(7, 3)$	0.0497	0.0473	0.0594	0
28	2.826	$(9, 4)$	0.0489	0.025	0.0632	0
29	4.710	$(1, 5)$	0	0	0.0373	0.0573
30	∞	$(8, 4)$	0.3622	0.4082	0.2856	0.4368
			Estimated λ			
			0.2516	0.2829	0.2199	0.1744
			Log Likelihood			
			-10.7168	-5.3898	-5.2938	-27.1106

non-treated data (Column 6); and the binary data, that is, the aggregated treated and non-treated data (Column 7).

The estimated RTMT using all the data shows that a relatively small number of types, a third of the total, is sufficient to capture the behavior of the population in this experiment. The results depict a very significant fraction of highly risk-averse types. The fraction of types with curvature close to or higher than logarithmic (type 24 and

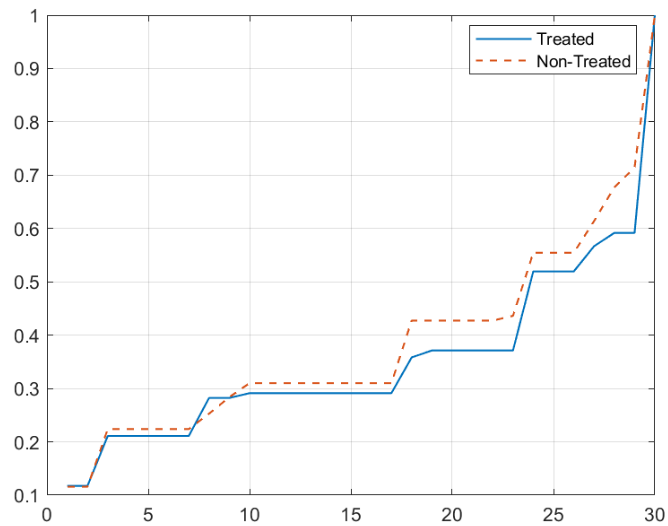

FIGURE 1.—CDFs of the estimated RTMTs of the treated and non-treated data.

above) is 60%, with more than a third of decisions corresponding to the highest type, type 30, which represents risk-aversion levels as high as 4.7 and above. Interestingly, the results also show that a relevant fraction of the decisions correspond to highly risk-seeking attitudes (22%), and even to extreme risk-seeking behavior, (with 12% of decisions corresponding to the lowest type, type 1, which represents risk-aversion coefficients below -4.8). The estimated probability of tremble is .25, thus reflecting the behavioral relevance of non-essential alternatives.

Comparison of the estimated RTMTs in the treatment-control data yields some interesting results. See Columns 5 and 6 in Table 2 for the densities, and Figure 1 for a representation of the corresponding CDFs. Both RTMTs allocate masses to similar preferences, although there is a slight tendency towards more risk aversion in the treated data. The masses allocated to types with curvature close to logarithmic or higher are 63% and 56% in the treated and non-treated data, respectively. This can also be seen in Figure 1, where starting from the 10th preference, the CDF of the treated RTMT dominates that of the non-treated. In addition, note that the estimated tremble with the treated data is about 22% larger than with the non-treated data. In sum, there are differences between treatment and control, suggesting a behavioral shift towards more risk aversion, and to large choice inconsistencies, but the latter do not appear dramatic. They may be due to the fact that the subject population was highly risk averse to begin with, thus leaving little room for the identification of larger effects in the treatment-control comparison, and also to the fact that the imposed time constraints were probably not strict enough. Finally, Column 7 reports the estimated RTMT using all the binary data. The estimation of the binary data can be computed exactly following the simple algorithm given in Remark 7. Interestingly, the binary data reveal risk-aversion levels analogous to those obtained previously.

References

[1] Afriat, S.N. (1973). "On a System of Inequalities in Demand Analysis: An Extension of the Classical Method," *International Economic Review*, 14:460–72.

[2] Andersson, O., H.J. Holm, J.R. Tyran and E. Wengström (2018). "Robust Inference in Risk Elicitation Tasks," mimeo.

[3] Apesteguia, J. and M.A. Ballester (2020). "Separating Predicted Randomness from Residual Behavior," *Journal of the European Economic Association*, forthcoming.

[4] Apesteguia, J., M.A. Ballester and A. Gutierrez (2020). "Random Models for the Joint Treatment of Risk and Time Preferences," mimeo.

[5] Apesteguia, J., M.A. Ballester and J. Lu (2017). "Single-Crossing Random Utility Models," *Econometrica*, 85(2):661–674.

[6] Barseghyan, L., F. Molinari and M. Thirkettle (2019). "Discrete Choice under Risk with Limited Consideration," mimeo.

[7] Barsky, R.B., F.T. Juster, M.S. Kimball and M.D. Shapiro (1997). "Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study," *Quarterly Journal of Economics*, 112(2):537–79.

[8] Chiappori, P-A, B. Salanié, F. Salanié and A. Gandhi (2019). "From Aggregate Betting Data to Individual Risk Preferences," *Econometrica*, 87(1):1–36.

[9] 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.

[10] Cohen, A. and L. Einav (2007). "Estimating Risk Preferences from Deductible Choice," *American Economic Review*, 97(3):745–788.

[11] Coller, M. and M.B. Williams (1999). "Eliciting Individual Discount Rates," *Experimental Economics*, 2(2):107–27.

[12] Dardanoni, V., P. Manzini, M. Mariotti and C.J. Tyson (2020). "Inferring Cognitive Heterogeneity from Aggregate Choices," *Econometrica*, forthcoming.

[13] Filiz-Ozbay, E. and Y. Masatlioglu (2020). "Progressive Random Choice," mimeo.

[14] Greene, W.H. and D.A. Hensher (2010). *Modeling Ordered Choices*, Cambridge University Press.

[15] Gul, F. and W. Pesendorfer (2006). "Random Expected Utility," *Econometrica*, 74(1):121–146.

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

[17] Jagelka, T. (2020). “Are Economists’ Preferences Psychologists’ Personality Traits? A Structural Approach,” mimeo.

[18] Mirrlees, J.A. (1971). “An Exploration in the Theory of Optimum Income Taxation,” *The Review of Economic Studies*, 38(2):175–208.

[19] Spence, A.M. (1974). *Market Signaling: Information Transfer in Hiring and Related Screening Processes*, Harvard University Press.

[20] Warner, J.T. and S. Pleeter (2001). “The Personal Discount Rate: Evidence from Military Downsizing Programs,” *American Economic Review*, 91(1):33–53.