

RANDOM UTILITY MODELS WITH ORDERED TYPES AND DOMAINS

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

ABSTRACT. Random utility models in which heterogeneity of preferences is modeled by means of an ordered collection of utilities, or types, provide a powerful framework for the understanding of a variety of economic behaviors. This paper studies the micro-foundations of ordered random utility models with the objective of meeting standard empirical requirements. This is done by working with arbitrary collections of ordered decision problems, and by making no parametric assumptions about the type distribution. The model is characterized by means of a simple monotonicity axiom. A set of extensions is studied, and goodness-of-fit measures are proposed, with proof provided of the strong consistency of extremum estimators defined upon them. A statistical test for the model is provided. The paper concludes with detailed guidelines for the practical implementation of the model.

Keywords: Random utility model; Ordered type-dependent utilities; Arbitrary domains.

JEL classification numbers: C00; D00

Date: April, 2022.

*We thank Jay Lu and Christoph Semken for helpful comments and Angelo Gutierrez for outstanding research assistance. Financial support by the Spanish Ministry of Economics and Competitiveness through Grant PGC2018-098949-B-I00 and through the Severo Ochoa Program for Centers of Excellence in R&D (CEX2019-000915-S) and Balliol College is gratefully acknowledged.

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

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

1. INTRODUCTION

Settings involving ordered utilities, or types, and ordered alternatives are common in economics. These are settings in which the analyst assumes an ordered collection of types describing some behavioral trait, with higher types selecting higher alternatives. Consider for example the case of decision under risk, where it is typical to assume a collection of types ordered by the notion of risk aversion and lotteries ordered by riskiness, such that the lowest type chooses the riskiest lottery, and successively higher types choose successively safer lotteries. As elaborated later in the paper, many other economic applications possess this structure, such as those involving quasi-linear or Cobb-Douglas utilities. In these settings, heterogeneity is often modeled as a probability distribution over the set of types, i.e., as an ordered random utility model.

The main objective of this paper is to bridge the gap between the theory of random utility models with ordered types and the standard empirical requirements for their actual implementation. Accordingly, we obtain our results without making particular assumptions neither on which ordered decision problems are available nor on which is the structure of the probability distribution over the type space.

The first contribution of the paper is to provide a characterization of the choice frequencies that can be generated by any random utility model over a given space of ordered types \mathcal{T} , which we call \mathcal{T} -RUM. We use the following novel property, \mathcal{T} -Monotonicity. Suppose that the set of types leading to alternatives B_1 within decision problem A_1 is a subset of the set of types yielding choices from B_2 within decision problem A_2 . Then, \mathcal{T} -Monotonicity states that the cumulated choice frequency of alternatives in B_1 , within decision problem A_1 , must be smaller than that of alternatives in B_2 , within decision problem A_2 . Interestingly, we show that, when decision problems are ordered, \mathcal{T} -Monotonicity is not only necessary but also sufficient. That is, \mathcal{T} -Monotonicity provides an exact test for the rationalizability of the data. To guarantee the applicability of our results, we provide an intuitive linear algorithm for the analysis of the property and, since the proof of the characterization theorem is fully constructive, we can determine the underlying type distribution which explains all choices when the property is satisfied by the data. Moreover, we show that the model is uniquely identified on a set of types that we characterize.

We then provide two generalizations of the ordered domain assumption. First, we note that our analysis goes through even if the alternatives of a decision problem are

not necessarily ordered, provided only that every alternative in every menu is chosen by an interval of types. Secondly, we show that any unordered decision problem can belong to the domain, as long as a simple richness condition is met. The first part of the paper concludes with an extension of the model, where we allow for the possibility of observing choices of dominated alternatives, that is, alternatives that are never maximal and are hence predicted to have zero choice probabilities. Our motivation for this stems from having often observed these patterns in actual datasets. Our approach is to minimally extend our main model, by introducing the possibility of mistakes being made at the time of choosing.

Given our interest in connecting theory with empirics, we then present some results of econometric interest dealing with finite data. First, observed choice frequencies, due to sampling issues, may violate \mathcal{T} -Monotonicity and, accordingly, we lay down choice-based goodness-of-fit measures. These are constructed based on allowing perturbation of the underlying distribution of types at each decision problem. Our approach considers the minimum perturbation required to explain all observed choice frequencies, and implicitly defines a class of extremum estimators. Most importantly, subsequent analysis shows that any estimator within this class is strongly consistent. That is, as the number of observations per choice problem increases, the estimator converges to the true distribution of probabilities over the types. Second, we show how the model can be statistically tested. We exploit the i.i.d. nature of \mathcal{T} -RUM which enables us to interpret the model as a collection of independent multinomial distributions with parameters connected by the distribution over the types. We then propose an aggregated Pearson statistic to statistically test the model.

We then proceed to provide a detailed and exhaustive guide for implementing the model, and illustrate each step in our guide using an existing experimental dataset involving decision problems over lotteries. Finally, we relate our paper to the existing literature and finish with some concluding remarks.

2. ORDERED RANDOM UTILITY MODELS AND ORDERED DOMAINS

Let X be the set of all alternatives. We fix an ordered collection of type-dependent utilities $\{U_t\}_{t \in \mathcal{T}}$, where $\mathcal{T} = \{1, 2, \dots, T\}$.¹ Given the ordinal nature of all our results,

¹For purposes of exposition, we assume no role for indifferences, i.e., utility functions are strict over the domain for which data exist. As it is customary in the literature, one can interpret that the

we could equivalently work with the corresponding collection of ordinal preferences. A random utility model over \mathcal{T} , or \mathcal{T} -RUM, is defined by a probability distribution ψ over \mathcal{T} , which describes the probability mass with which each type is realized. In each decision problem, or menu, one of the utility functions is independently realized according to ψ and maximized, thus determining the choice. Menus are finite subsets of alternatives. We work with an arbitrary collection of menus, $\{A_j\}_{j \in \mathcal{J}}$, where $\mathcal{J} = \{1, 2, \dots, J\}$. Given the \mathcal{T} -RUM with distribution ψ , the probability of choosing alternative x in menu j is $\psi(\mathcal{T}(x, j))$, where $\mathcal{T}(x, j)$ denotes the set of types for which alternative x is the utility maximizer in menu j .

Ordered collections of utilities induce an order over some pairs of alternatives. We say that alternative x_h is *higher than* alternative x_l whenever there exists $t^* \in \mathcal{T} \setminus \{T\}$ such that $U_t(x_l) > U_t(x_h) \Leftrightarrow t \leq t^*$. In this case we write $x_l \triangleleft x_h$, and, as usual, $x \trianglelefteq y$ whenever $x \triangleleft y$ or $x = y$. In words, x_h is higher than x_l if x_h is the preferred alternative of high types (with at least type T expressing this preference) and x_l is the preferred alternative of low types (with at least type 1 expressing this opposite preference). For instance, types can be ordered by risk aversion or altruism, and hence the notion of a higher alternative corresponds either to the notion of a safer lottery or to that of a more altruistic distribution. We now introduce the only relevant assumption in the paper: i.e., that every menu in the domain is ordered, in the sense that its maximal alternatives can be ordered by \trianglelefteq .

Domain of Ordered Menus. For every $j \in \mathcal{J}$, \trianglelefteq is complete over $\{x : \mathcal{T}(x, j) \neq \emptyset\}$.

Domains composed by ordered menus appear naturally when studying a particular behavioral trait. We illustrate the generality of our setting with three economic applications.

2.1. Expected Utility. Consider the case of finite monetary lotteries and expected utility. Let $\{EU_t\}_{t \in \mathcal{T} = \{1, 2, \dots, T\}}$ be a collection of expected utilities ordered by increasing risk aversion, i.e., by increasing concavity of their respective monetary utilities. Thus, in this case, the induced relation \triangleleft represents the notion of a safer lottery.

We now show that the most standard domains of menus of lotteries used in the study of risk aversion are ordered. We consider the following three well-known classes

set of types corresponds to a population of heterogenous deterministic individuals, or equivalently to a single individual with internal heterogeneity.

of menus. Class (1) encompasses menus in which all lotteries contemplate a good-state payoff with some probability p , and a bad-state payoff with probability $1 - p$, where p is constant within each menu, but not necessarily across all menus. This formulation is a generalization of widely-used experimental decision problems, such as multiple price lists or convex budget sets, and it includes the deductible-premium insurance setting of Barseghyan, Molinari and Thirkettle (2021).² Class (2) involves any menu in which all lotteries belong to the same Marschak-Machina triangle; i.e., they are all composed of the same three monetary payoffs. A prominent example of the use of different Marschak-Machina triangles for different menus is Camerer (1989). Class (3) contemplates any menu with lotteries offering a possibly different prize with a possibly different probability, and a constant payoff otherwise, such as in the betting example of Chiappori, Salanié, Salanié and Gandhi (2019).

Lemma 1 uses classical results on second order stochastic dominance, such as Hammond (1974), to show that every domain composed of such menus is ordered.³

Lemma 1. *Let $\{EU_t\}_{t \in \mathcal{T}=\{1,2,\dots,T\}}$ be ordered by increasing risk aversion. Any domain composed of menus in classes (1), (2) and (3) is ordered.*

2.2. Quasi-Linear Utility. Consider now the case where alternatives are pairs of the form (q, w) , with q belonging to an ordered set Q and w representing money. For example, q may describe the quality of a product or the level of leisure time in a labor application, and w the income after purchase or after salary. Consider a collection of quasi-linear utilities, $\{QL_t\}_{t \in \mathcal{T}=\{1,2,\dots,T\}}$, with $QL_t(q, w) = v_t(q) + w$, such that the family $\{v_t\}_{t \in \mathcal{T}}$ satisfies the well-known increasing differences condition (see, e.g., Topkis (1978) or Milgrom and Shannon (1994)). In this context, \triangleleft represents the notion of alternatives with higher quality or more leisure time, and it is immediate to see that every domain of menus of such objects is ordered.

2.3. Cobb-Douglas Utility. Consider a consumption setting where alternatives are described by means of pairs (x_1, x_2) of positive real values. Another standard use of this setting is for the study of other-regarding preferences where the pairs describe

²Another influential case within this class is Tanaka, Camerer and Nguyen (2010) who study a domain in which it is the good-state payoff, and not the probability of the state, that varies across menus. Also, notice that the limit case, in which the good-state and bad-state payoffs are equal, represents the degenerate lotteries often used to elicit certainty equivalents, for example.

³All the proofs are given in Appendix A.

one's own income and the income of another individual. Consider a collection of Cobb-Douglas utilities, $\{CD_t\}_{t \in \mathcal{T}=\{1,2,\dots,T\}}$, with $CD_t(x_1, x_2) = x_1^{\alpha_t} x_2^{1-\alpha_t}$ such that $0 \leq \alpha_1 < \alpha_2 < \dots < \alpha_T \leq 1$. The induced relation \triangleleft trivially corresponds to the idea that: $(x_1, x_2) \triangleleft (y_1, y_2)$ if and only if $x_1 > y_1$ and $x_2 < y_2$. It is again trivial to see that, in this setting, every domain of menus of bundles or income distributions is ordered.

3. A CHARACTERIZATION OF \mathcal{T} -RUMS

Suppose that the analyst has access to a stochastic choice function p over an ordered domain of menus \mathcal{J} . Formally, p is a map from $X \times \mathcal{J}$ to $[0, 1]$ such that, for every $j \in \mathcal{J}$, $p(x, j) > 0$ implies that $x \in A_j$, and $\sum_{x \in A_j} p(x, j) = 1$. Consider the following property.

$$\mathcal{T}\text{-Monotonicity: } \bigcup_{x \in B} \mathcal{T}(x, j) \subseteq \bigcup_{x' \in B'} \mathcal{T}(x', j') \Rightarrow \sum_{x \in B} p(x, j) \leq \sum_{x' \in B'} p(x', j').$$

\mathcal{T} -Monotonicity captures the intuition that more support leads to larger choice probabilities. Whenever the set of types leading to alternatives B in menu j is contained in the set of types leading to alternatives B' in menu j' , the cumulated probability of alternatives in B must be lower than that of the alternatives in B' . \mathcal{T} -Monotonicity is satisfied by \mathcal{T} -RUMs because of the menu-independent structure of choice. We now show that, when the domain is composed of ordered menus, this property is not only necessary but also sufficient for \mathcal{T} -RUMs.

Theorem 1. *In a domain of ordered menus, p satisfies \mathcal{T} -Monotonicity if, and only if, p is a \mathcal{T} -RUM.*

The following discussion is instrumental in understanding the strategy of the proof, and hence the result. First, we show that the ordered structure of menus guarantees that choices are ordered, that is, for any menu j and alternative $x \in A_j$, the set of types $\mathcal{T}(x, j)$ is always an interval and the set of types $\bigcup_{y \in A_j, y \triangleleft x} \mathcal{T}(y, j)$ is of the form $\{1, 2, \dots, \max \mathcal{T}(x, j)\}$. That is, in every menu, choice is ordered by the set of types, with lower types selecting lower alternatives. This suggests, further, that the most relevant types are those that are the largest maximal types for the different alternatives and menus. We denote the set of such types by \mathcal{T}^I . The proof then constructs a correspondence F on \mathcal{T}^I , that will provide the basis for the construction of the CDF over the relevant set of types that rationalizes choice. Given the ordered choices, F

assigns to each type $t \in \mathcal{T}^I$ with $t = \max \mathcal{T}(x, j)$ the sum of the choice probabilities of all the alternatives lower than x in menu A_j . Then, the proof uses \mathcal{T} -Monotonicity to show that F satisfies the conditions to be a CDF, that is, it is a single-valued increasing map, with $F(T) = 1$. Since types outside \mathcal{T}^I are inconsequential for choice, we can then construct a monotone extension, G , of F over the entire collection of types. By a monotone extension we mean that $F(t) = G(t)$ for every $t \in \mathcal{T}^I$ and that, whenever $t_1 < t_2$, $G(t_1) \leq G(t_2)$. Then, the probability distribution ψ derived from G is shown to rationalize the data.

3.1. Identification. The proof of Theorem 1 is based upon the construction of the model on the set of types \mathcal{T}^I , which are, in fact, the ones at which the CDF of the \mathcal{T} -RUM is fully identified. The above discussion shows that every $t \in \mathcal{T}^I$ is fully identified. Whenever $t \notin \mathcal{T}^I$, the CDF at t cannot generally be fully identified. Consider e.g. $t_1 < t < t_2$, where t_1 and t_2 are two consecutive types in \mathcal{T}^I . If the value of the CDF at t_2 is strictly larger than its value at t_1 , the value of the CDF at t can be constructed by assigning any value between these two. The reason for this is that, given $G(t_1) < G(t_2)$, the value $G(t) \in [G(t_1), G(t_2)]$ is inessential for choice, since every $t \in (t_1, t_2]$ has the same maximal alternative in every menu in the domain. It can, therefore, be deduced that the CDF at type t can only be fully identified in the extreme case in which $G(t_1) = G(t_2)$, implying $G(t) = G(t_1) = G(t_2)$, and $\psi(t) = 0$. Obviously, data on more menus may expand the set of identifiable types. For example, in an experimental setting, the analyst may use these results in order to select the domain that allows identification of the most important components of the CDF.

3.2. Complexity. \mathcal{T} -Monotonicity provides intuition for understanding the model, since it basically claims that more support on the type space comes with increased choice probability. We now discuss the computational complexity of verifying it. To do this, we use a weaker version of \mathcal{T} -Monotonicity and argue that it is computationally easy to check.

$$\mathcal{T}\text{-Monotonicity}^*: \max \mathcal{T}(x, j) \leq \max \mathcal{T}(x', j') \Rightarrow \sum_{z \in A_j, z \leq x} p(z, j) \leq \sum_{z \in A_{j'}, z \leq x'} p(z, j').$$

\mathcal{T} -Monotonicity*, as opposed to \mathcal{T} -Monotonicity, directly uses the ordered structure of the setting in order to establish the implication on choice behavior. In essence,

\mathcal{T} -Monotonicity* applies the logic of \mathcal{T} -Monotonicity to subsets that are lower contour sets. Claim 2 in the proof of Theorem 1 makes this apparent, and hence \mathcal{T} -Monotonicity* is a weaker version of \mathcal{T} -Monotonicity. The construction of the CDF in the remainder of the proof uses only the lower contour sets involved in \mathcal{T} -Monotonicity*, showing that these two properties are in fact equivalent within ordered domains.⁴

Now, we argue that testing for \mathcal{T} -Monotonicity* is a simple task. Let $k_{\mathcal{J}} = \sum_{j \in \mathcal{J}} |A_j|$, which corresponds to the total number of possible pairs of alternatives and menus, (x, A_j) with $x \in A_j$. A brute force algorithm enables simple checking of all such pairs, requiring a total number of $k_{\mathcal{J}}(k_{\mathcal{J}} - 1)$ checks, which is already polynomial in the input $k_{\mathcal{J}}$. This can be significantly improved, moreover, by using the following recursive argument. Let all pairs of alternatives and menus be ordered in some way. Suppose that the property holds for the first n pairs. When considering pair $n + 1$, one needs to relate this pair with, at most, two previously-considered pairs. For the case where there are previously considered pairs with the same largest maximal type, one can select any of these to check for the equality of their (cumulative) choice probability with that of pair $n + 1$. Otherwise, one needs to select the pairs with largest maximal types closest, above and below, to the current one, and check that the (cumulative) choice probability of pair $n + 1$ lies between the choice probabilities of these two. Hence, this algorithm involves at most $2k_{\mathcal{J}}$ comparisons, and is therefore very low in complexity.

3.3. Endogenization of the set of types. The implementation of our characterization result requires us to start by fixing a family of types. Although this is common practice in applied work, the theorist may wonder whether it is possible to endogenize the set of types, providing characterization results that do not explicitly assume a family ex-ante. The answer is affirmative; in many cases, we can establish our results only in terms of the economic application at hand. Intuitively, this only requires us to replace the key term in \mathcal{T} -Monotonicity*, $\max \mathcal{T}(x, j)$, with its specific meaning in the application. In the following, we provide two examples.

⁴Formally, \mathcal{T} -Monotonicity* is equivalent to \mathcal{T} -Monotonicity provided that the domain includes at least one menu with no dominated options. This is a very weak assumption, but even when it is not met, it could be shown that \mathcal{T} -Monotonicity* together with \mathcal{T} -Extremeness, presented in the proof of the theorem, are equivalent to \mathcal{T} -Monotonicity. Note that \mathcal{T} -Extremeness is also a trivial property to be checked.

Consider three monetary payouts $a > b > c$ and the space of all lotteries contained in the associated Marschak-Machina triangle, i.e., vectors of the form $q = (q_a, q_b, q_c)$ with $q_a, q_b, q_c \geq 0$ and $q_a + q_b + q_c = 1$. Let \mathcal{D}_{MM} be a finite domain composed of finite menus constructed using lotteries from this Marschak-Machina triangle. Then, given $A_j \in \mathcal{D}_{MM}$, denote by $\bar{A}_j = \{q^{j,1}, q^{j,2}, \dots, q^{j,K_j}\}$ the collection of vertices of the convex hull of A_j that are not first-order stochastically dominated by any other vertex.⁵ The notation is chosen to satisfy $q_a^{j,1} > q_a^{j,2} > \dots > q_a^{j,K_j}$ and hence also, $q_c^{j,1} > q_c^{j,2} > \dots > q_c^{j,K_j}$. We now describe two properties that arise immediately from the consideration of \mathcal{T} -Extremeness and \mathcal{T} -Monotonicity*. For the latter, for every menu j and $k < K_j$, denote the ratio value $r(q^{j,k}, j) = \frac{q_a^{j,k} - q_a^{j,k+1}}{q_c^{j,k} - q_c^{j,k+1}}$.

Property 1: $p(q, j) > 0$ implies $q \in \bar{A}_j$.

Property 2: $r(q^{j,k}, j) \leq r(q^{j',k'}, j')$ implies $\sum_{h \leq k} p(q^{j,h}, j) \leq \sum_{h' \leq k'} p(q^{j',h'}, j')$.

We say that a Random Expected Utility Model (REUM) is a random utility model over expected utilities.⁶ The following result is an immediate consequence of our main characterization result.

Corollary 1. *In \mathcal{D}_{MM} , p satisfies properties 1 and 2 if, and only if, p is a REUM.*

A similar result can be obtained for what we call the Random Cobb-Douglas Model (RCDM), that is a random utility model over Cobb-Douglas utilities.⁷ Let \mathcal{D} be a finite domain composed of finite menus of two-dimensional bundles. Given $A_j \in \mathcal{D}$, denote by I_j all interior bundles of A_j . For any $x \in I_j$ we can write the bundle in logarithmic terms, i.e., $y = (y_1, y_2) = (\log x_1, \log x_2)$, and denote by Y_j the set of the logarithmic transformations of bundles in I_j . Denote by \bar{Y}_j the subset of vertices of the convex hull of Y_j that are not dominated by any other different vertex in both dimensions. Finally, denote by $\bar{I}_j = \{x^{j,1}, \dots, x^{j,L_j}\}$ the interior bundles corresponding

⁵In the present case, q first-order stochastically dominates q' if $q_a \geq q'_a$ and $q_c \leq q'_c$. \bar{A}_j is, as we argue in the proof of Corollary 1, the subset of lotteries in menu A_j that can be uniquely selected by some expected utility.

⁶See Gul and Pesendorfer (2006) for a full treatment of the REUM model. Importantly, notice that the finiteness of the domain guarantees that the masses can be always concentrated over a finite number of expected utilities and, hence, we avoid a formal treatment of measurability. We assume that none of these expected utilities creates indifferences over the lotteries in the domain.

⁷We assume again that indifferences play no role.

to the elements in \bar{Y}_j .⁸ The notation is chosen to satisfy $x_1^{j,1} < x_1^{j,2} < \dots < x_1^{j,L_j}$ and hence also, $x_2^{j,1} > x_2^{j,2} > \dots > x_2^{j,L_j}$. Menu j may also contain corner bundles. Whenever the latter is the case, denote by $z^{j,2}$ (respectively, $z^{j,1}$) the corner bundle with the largest amount of commodity two (respectively, commodity one) satisfying $z_2^{j,2} > x_2^{j,L_j}$ (respectively, $z_1^{j,1} > x_1^{j,1}$). We can then consider the ordered sub-menu $\bar{A}_j = \{a^{j,1}, \dots, a^{j,K_j}\} = \{z^{j,2}, x^{j,1}, \dots, x^{j,L_j}, z^{j,1}\}$ and the following two properties arise immediately from the consideration of \mathcal{T} -Extremeness and \mathcal{T} -Monotonicity*. For the latter, for every menu j and $l < L_j$, we define $\hat{r}(x^{j,l}, j) = \frac{y_2^{j,l+1} - y_2^{j,l}}{y_1^{j,l} - y_1^{j,l+1}}$. Let \hat{r} and \hat{R} be the smallest and the largest of all these values and define let $\hat{r}(z^{j,2}, j) = \hat{r} - \epsilon$, $\hat{r}(x^{j,L_j}, j) = \hat{R} + \epsilon$ and $\hat{r}(z^{j,1}, j) = \hat{R} + 2\epsilon$ for some $\epsilon > 0$. Then:

Property 3: $p(x, j) > 0$ implies $x \in \bar{A}_j$.

Property 4: $\hat{r}(a^{j,k}, j) \leq \hat{r}(a^{j',k'}, j')$ implies $\sum_{h \leq k} p(a^{j,h}, j) \leq \sum_{h' \leq k'} p(a^{j',h'}, j')$.

Corollary 2. *In \mathcal{D} , p satisfies properties 3 and 4 if, and only if, p is a RCDM.*

4. BEYOND ORDERED MENUS

Domains of ordered menus have the attractive feature of being built on the basis of the intuitive “higher than” relation, which is, arguably, the building block for the empirical study of any given behavioral trait. This section discusses two generalizations of the domain assumption.

4.1. Interval domain. In an ordered menu j , if $t_1 < t_3$ are types leading to the choice of x , $t_2 \in (t_1, t_3)$ must also result in the same choice, given that, otherwise, the alternative chosen by t_2 would be incomparable with x according to \triangleleft . Hence, the set of types $\mathcal{T}(x, j)$ is in fact an interval set. Formal proof of this is given in Claim 1 of Theorem 1. Here, we argue that this interval structure in choices can be used, instead of the ordered menu structure, to obtain the same characterization result. To see the latter, consider the simple case of three alternatives $\{x, y, z\}$, and the following types: $U_1(x) > U_1(y) > U_1(z)$, $U_2(y) > U_2(z) > U_2(x)$, and $U_3(z) > U_3(x) > U_3(y)$. It is immediate to see that $x \triangleleft z$ and $y \triangleleft z$, but x and y cannot be compared for \triangleleft , since x is maximal for types 1 and 3 and y is maximal for type 2. Hence, menus $\{x, y\}$ and

⁸Again, as we argue in the proof of Corollary 2, it is evident that \bar{I}_j corresponds to the set of interior bundles that can be uniquely selected by some Cobb-Douglas utility.

$\{x, y, z\}$ fail to be ordered and fall outside our initial domain assumption. However, $\{x, y, z\}$ satisfies the interval structure, since type 1 chooses x , type 2 chooses y and type 3 chooses z , and hence it can be incorporated into the analysis.

Domain of Interval Menus. For every $j \in \mathcal{J}$ and $x \in A_j$, $\mathcal{T}(x, j)$ is an interval.

Corollary 3. *In a domain of interval menus, p satisfies \mathcal{T} -Monotonicity if, and only if, p is a \mathcal{T} -RUM.*

4.2. Replica domain. We now show that the assumption of interval menus can be further generalized to the following domain condition.

Domain of Replica Menus. For every $j \in \mathcal{J}$ and every $x \in A_j$, there exists a replica menu $j' \in \mathcal{J}$ such that: (i) j' is an interval menu, and (ii) $\mathcal{T}(x, j) = \bigcup_{y \in B} \mathcal{T}(y, j')$ for some $B \subseteq A_{j'}$.

Consider a menu j for which there is an alternative x that is maximal for a set of types $\mathcal{T}(x, j)$ that is not an interval. This would leave the domain outside the interval order domain. However, we show here that, by using a replica menu, we can incorporate this sort of menus into the analysis. The replica menu for j is an interval menu j' where alternative x is replicated in the sense that there is a set of alternatives $B \subseteq A_{j'}$ such that x is maximal in j if, and only if, an element in B is maximal in j' .

Clearly, every domain of interval menus is also a domain of replica menus, but the opposite is not necessarily true. To see this, reconsider the example given in the previous subsection, where menu $\{x, y\}$ is neither an ordered menu nor an interval menu, since x is maximal for types 1 and 3 while y is maximal for type 2. We now claim that $\{x, y, z\}$ is a replica menu of $\{x, y\}$. We learnt above that $\{x, y, z\}$ is an interval menu, so (i) holds. Moreover, since each alternative in $\{x, y, z\}$ is chosen by a unique type, alternative x in $\{x, y\}$ can be replicated by alternatives x and z in $\{x, y, z\}$. Therefore, we can include the pair $\{x, y\}$ in the domain as long as menu $\{x, y, z\}$ is also added. Hence, in this particular case, the universal domain is a replica domain despite being neither an interval nor an ordered menu domain. The following result extends Theorem 1 to the case of domains of replica menus.⁹

⁹Intuitively, the result can be extended beyond replica domains by noticing that the replication of menu j could be obtained equivalently by considering a collection of interval menus rather than a unique replica menu.

Corollary 4. *In a domain of replica menus, p satisfies \mathcal{T} -Monotonicity if, and only if, p is a \mathcal{T} -RUM.*

We now illustrate the usefulness of domains of replica menus in the expected utility context, where it is immediate that menus are not necessarily interval ordered. We show that a single Marschak-Machina triangle is sufficient to produce replicas of any menu of lotteries, and hence allow the analysis of domains including such menus.

Proposition 1. *Let $\{EU_t\}_{t \in \mathcal{T}=\{1,2,\dots,T\}}$ be ordered by increasing risk aversion and consider any three monetary payoffs $a > b > c$. For every menu of lotteries j and every lottery $x \in A_j$, there exists a replica menu in the Marschak-Machina triangle defined by these three payoffs.*

The proof of Proposition 1 constructs elementary menus which exactly reproduce the choice intervals of any arbitrary, not-necessarily interval ordered, menu of lotteries. Thus, any domain of menus of lotteries can be extended into a replica domain simply by adding menus of lotteries from a given Marschak-Machina triangle.

5. TREMBLE

Given a menu j , denote by D_j the set of alternatives that are not maximal for any of the utility types, i.e., $D_j = \{x \in A_j : \mathcal{T}(x, j) = \emptyset\}$. We know that \mathcal{T} -RUMs assign zero probability to the choice of alternatives in D_j , and their choice, therefore, constitutes a mistake. In this section, we extend \mathcal{T} -RUMs in order to allow for the possibility of such mistakes.

The \mathcal{T} -RUM with tremble (\mathcal{T} -RUMT) is defined by means of a possibly menu-dependent tremble function $\lambda : \mathcal{J} \rightarrow [0, 1]$, such that $\lambda_j = 0$ whenever $D_j = \emptyset$, and a probability distribution ψ over the set of types. Then, for any menu j , the total mass of choices from D_j is given by λ_j , while the choice probability of $x \in A_j \setminus D_j$ is equal to $\psi(\mathcal{T}(x, j))(1 - \lambda_j)$. That is, mistakes occur with probability λ_j , and otherwise behavior is governed by the \mathcal{T} -RUM with distribution ψ . Notice that this tremble version of the model is agnostic about the size and distribution of mistakes in D_j , and allows different trembles in different menus. The only assumption in the trembling mechanism is that when choices are not mistakes, they follow a \mathcal{T} -RUM.

\mathcal{T} -RUMT is a simple model whose characterization follows immediately from the analysis in Theorem 1. Given the stochastic choice function p , denote by \bar{p} the conditional stochastic choice function:¹⁰

$$\bar{p}(x, j) = \begin{cases} \frac{p(x, j)}{1 - \sum_{y \in D_j} p(y, j)}, & \text{if } x \in A_j \setminus D_j; \\ 0, & \text{otherwise.} \end{cases}$$

We then have:

Corollary 5. *In a domain of ordered menus, p is a \mathcal{T} -RUMT if, and only if, \bar{p} satisfies \mathcal{T} -Monotonicity.*

Remark 1. The implementation of the tremble in \mathcal{T} -RUMT is flexible. For example, one could discard the observed choices of the non-maximal alternatives, and estimate a \mathcal{T} -RUM over the remaining data. Alternatively, one could adopt a particular structure for tremble. The simplest way would be to use a λ that is menu-independent and such that choices over non-maximal alternatives are uniformly random.¹¹

The main assumption of the trembling model is that the conditional choice probabilities over the maximal alternatives follow a \mathcal{T} -RUM. There are other plausible trembling mechanisms that would also lead to this property. For example, one could consider the possibility of mistakes occurring uniformly over the entire menu j , not just over D_j , and behaving a la \mathcal{T} -RUM otherwise.

6. ESTIMATION AND STATISTICAL TESTING

We now return to our base model described in Section 2, and present several results on the estimation and statistical testing of \mathcal{T} -RUMs when the data are finite. Formally, the data form a map $z : X \times \mathcal{J} \rightarrow \mathbb{Z}_+$, describing the number of observed instances in which each alternative is chosen in each menu, with $z(x, j) > 0$ implying that $x \in A_j$. For every $j \in \mathcal{J}$, we denote by $z(\cdot, j)$ the vector describing the observed choices in menu j , and by $Z_j = \sum_{x \in A_j} z(x, j) > 0$ the total number of observations for this menu. The observed choice frequencies in menu j are therefore $\tilde{z}_j(\cdot) = \frac{z(\cdot, j)}{Z_j}$.

¹⁰When $\sum_{x \in D_j} p(x, j) = 1$, define the (irrelevant) conditional probability as $\bar{p}(x, j) = \frac{|\mathcal{T}(x, j)|}{|\mathcal{T}|}$.

¹¹We illustrate the latter route in Section 7.

6.1. Estimation. Suppose that the data are generated by a \mathcal{T} -RUM but that, due to sampling issues, \mathcal{T} -Monotonicity is violated. In this section, we provide a class of estimators that are based on a notion of rationalizability and show that they are strongly consistent. For this, we assume that $\mathcal{T} = \mathcal{T}^I$, that is, for ease of exposition, we work directly with the set of types that are fully identified.

Consider the following generalization of the rationalizability notion embedded in \mathcal{T} -RUMs. Let Ψ denote the set of all probability distributions over the given set of types. Let $d : \Psi \times \Psi \rightarrow \mathbb{R}_+$ be a continuous function measuring the divergence between two probability distributions such that $d(\psi, \psi) = 0$ and $|\psi - \psi'| \preceq |\psi - \psi''|$ implies $d(\psi, \psi') < d(\psi, \psi'')$.¹² Now, let $f : \mathbb{R}_+^J \rightarrow \mathbb{R}_+$ be a continuous function aggregating all deviations across menus, such that $f(0, \dots, 0) = 0$ and $\gamma = (\gamma_1, \dots, \gamma_J) \preceq (\gamma'_1, \dots, \gamma'_J) = \gamma'$ implies $f(\gamma) < f(\gamma')$. We then say that the data z is ϵ -rationalizable if there exist distributions $(\psi, \{\psi_j\}_{j \in J})$ such that: (i) for every j , distribution ψ_j rationalizes the choice frequencies in menu j , \tilde{z}_j , and (ii) $f(d(\psi, \psi_1), \dots, d(\psi, \psi_J)) \leq \epsilon$. That is, there exists a fundamental probability distribution, ψ , over the set of types but, at the moment of choice, the menu-dependent distribution ψ_j determines choices in menu j . However, the aggregate measure of deviations between ψ and $\{\psi_j\}_{j \in J}$ is smaller than or equal to ϵ . We can now present the following straightforward result:

Corollary 6. *For every z there is a minimum ϵ such that z is ϵ -rationalizable. Moreover, z is 0-rationalizable if, and only if, the stochastic choice function defined by its choice frequencies is a \mathcal{T} -RUM.*

With ϵ large enough, any finite choice data generated by a \mathcal{T} -RUM can be ϵ -rationalized by allowing sufficiently large menu-dependent perturbations. The continuity of maps d and f guarantees a minimal rationalizability value. The second part of the result describes how ϵ -rationalizability constitutes a generalization of rationalizability by a \mathcal{T} -RUM, as the latter requires no perturbation whatsoever, i.e., $\epsilon = 0$. Hence, when finite data violate \mathcal{T} -Monotonicity, a natural goodness-of fit measure for

¹²Where \preceq refers to the component wise inequality, with at least one component being different. The continuity claims in this section involve always finite real simplices or finite dimensional real spaces, and hence we use the standard Euclidean topology.

the model is given by the smallest magnitude of ϵ that yields ϵ -rationalizability. Moreover, the distribution $\hat{\psi}$ in Ψ which yields the minimal value ϵ , represents an intuitive estimator, which we now show to be strongly consistent.¹³

Theorem 2. *$\hat{\psi}$ is strongly consistent.*

The proof of the strong consistency of the estimators built on the basis of ϵ -rationalizability is related to known results of extrema estimators for a multinomial model (see, e.g., van der Vaart, 2000). The proof takes care of the fact that our model involves collections of multinomials, one per menu, the parameters of which are connected by the underlying distribution of types. Also, given that we only use the monotonicity and continuity of functions d and f , the argument applies to non-necessarily additive estimators.

6.2. Statistical Testing. We now discuss the method for statistically testing the model. Given the multinomial structure of the choices in each of the menus, we can intuitively construct the following statistic, which aggregates the standard Pearson statistic across menus in \mathcal{J} :

$$C(z) = \sum_{j \in \mathcal{J}} \sum_{x \in A_j} \frac{(z(x, j) - Z_j \psi(\mathcal{T}(x, j)))^2}{Z_j \psi(\mathcal{T}(x, j))}.$$

We can then show the following result.

Theorem 3. *$C(z)$ converges to a Chi-square with $\sum_{j \in \mathcal{J}} |A_j| - J$ degrees of freedom.*

The proof of the theorem immediately follows from the independence of the multinomial distributions across menus in a \mathcal{T} -RUM.

7. PRACTICAL IMPLEMENTATION OF \mathcal{T} -RUM

We now show how to implement our framework in practice, discussing each of the steps required for the analysis of a dataset using \mathcal{T} -RUMs and all the techniques developed in the previous sections. We use the experimental dataset analyzed in Apestegua and Ballester (2021) that involves choices over lotteries and, since the purposes are

¹³Several natural examples of estimators belong to this family. In standard maximum-likelihood or least squares or minimum Chi-square estimations, for instance, d is a map that operates on each subset of types $\mathcal{T}(x, j)$, considering the logarithmic ratio, or the square of the difference between the two distributions, or the normalized square of the difference. Then, in these three standard cases, f additively aggregates all of the above-mentioned distances across menus.

purely expository of the framework, we adopt the most convenient empirical implementation strategy, namely, we assume that all decisions are made following the same distribution over CRRA expected utility types.¹⁴

We start by describing the dataset. Eighty-seven UCL undergraduates were asked to choose lotteries from menus of sizes 2, 3, and 5, from the nine equiprobable monetary lotteries described in Table 1. Each participant 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 for the selection and order of presentation of the menus of 3 and 5 alternatives, and for the location of the lotteries on the screen and the monetary prizes within a lottery. There were two treatments, but for present purposes we merge all the data from the experiment.¹⁵

TABLE 1. Lotteries

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

Step 1 (Set of Types). The first step involves the selection of an ordered family of ordinal preferences $\{U_t\}_{t \in \mathcal{T}}$. For example, one may start with a well-known family of utilities that is sufficiently rich, and look for the parameter values that yield indifference between a pair of alternatives in the dataset. Then, \mathcal{T} could be specified by considering preferences right below and above the parameters that create indifference. Note that, in \mathcal{T} -RUMs, all that matters is the ordinal preference of alternatives, and two utility functions that are not separated by any of these thresholds are associated with the same ordinal preference.

Given that our dataset involves lottery choices, here we adopt the most standard practice and use expected utility with CRRA monetary utilities, where $u_\omega(x) = \frac{x^{1-\omega}}{1-\omega}$,

¹⁴To emphasize, the selection of expected utilities with CRRA monetary utilities, and the consideration of aggregate data with the assumption that all individuals are equally stochastic has only an illustrative purpose; we simply aim for a better understanding of how to implement our results in a given economic setting.

¹⁵Experimental payoffs were determined by randomly selecting one menu, and awarding subjects in accordance with their choice from that menu.

TABLE 2. Preferences and Estimation Results

t	ω	U_t	ψ	t	ω	U_t	ψ
1	-4.148	2-6-3-7-4-8-5-9-1	0.1166	16	0.374	(7,2)	0
2	-0.518	(1,9)	0	17	0.408	(8,2)	0
3	-0.313	(3,6)	0.1009	18	0.443	(9,2)	0.0915
4	-0.083	(4,7)	0	19	0.516	(4,3),(8,7)	0.0079
5	0.065	(5,8)	0	20	0.607	(6,2)	0
6	0.154	(4,6)	0	21	0.652	(5,3),(9,7)	0
7	0.209	(1,8)	0	22	0.808	(1,3)	0
8	0.229	(3,2),(7,6)	0.0561	23	0.844	(8,3)	0
9	0.258	(5,6)	0	24	1.000	(9,3)	0.1388
10	0.262	(1,6)	0.0274	25	1.124	(5,4),(9,8)	0
11	0.273	(5,7)	0	26	1.309	(1,4)	0
12	0.339	(4,2),(8,6)	0	27	2.000	(7,3)	0.0497
13	0.342	(1,7)	0	28	2.826	(9,4)	0.0489
14	0.358	(5,2),(9,6)	0	29	4.710	(1,5)	0
15	0.363	(1,2)	0	30	∞	(8,4)	0.3622

λ : 0.2516 log-likelihood: -10.7168

whenever $\omega \neq 1$, and $u_1(x) = \log x$, where ω represents the risk-aversion coefficient.¹⁶ Consider every risk aversion coefficient ω for which two lotteries in the experiment are indifferent. This corresponds to the 29 finite values reported in Columns 2 and 6 of Table 2. Now, to specify the family \mathcal{T} consider one utility in each of the 30 intervals determined by these values. Notice, moreover, that this implies that these 30 ordinal preferences account for all the possible CRRA ordinal preferences, given the set of lotteries. Columns 3 and 7 in Table 2 describe the ordinal preferences per type, beginning with the preference of the first type and then specifying the pair(s) of alternatives that flip from the preference order of the previous type.

Step 2 (Ordered domain). Next, it is necessary to check whether the domain of menus is ordered, given the selected family of utilities. Thus, one needs to identify in each menu $j \in \mathcal{J}$ and for every type $t \in \mathcal{T}$ the alternative $x \in A_j$ that is maximal.

¹⁶Since lotteries l_2 and l_6 involve 0 payoffs, we assume a small fixed positive background consumption.

Now, for every pair of maximal alternatives in a menu, one could check whether there is a unique $t^* \in \mathcal{T} \setminus \{T\}$ such that one alternative is preferred to the other if, and only if, $t \leq t^*$. Alternatively, one could follow Claim 1 in the characterization theorem and the discussion of interval domains in Section 4.1, and check whether, for every menu $j \in \mathcal{J}$ and alternative $x \in A_j$, $\mathcal{T}(x, j)$ is an interval.

It is routine to check that our domain of menus is ordered, given our set of types. Note that some of these menus do not fall within the classes covered in Lemma 1, thus reinforcing our claim regarding the applicability of the setting studied in this paper.

Step 3 (Identifiability). Having specified $\{U_t\}_{t \in \mathcal{T}}$ and shown that the domain is ordered, we now obtain the set of fully identifiable types \mathcal{T}^I . To do this, it is enough to determine those types that are the largest maximal type for an alternative in a menu. Given the ordered domain, it is sufficient to select every type whose maximal alternative in a menu is different from that of the next type.

In our case, we have selected all the ordinal CRRA preferences for the lotteries involved in the experiment, and, since our dataset includes all the binary menus, all 30 types are identifiable.

Step 4 (Characterizing property). In the main result of the paper we show that \mathcal{T} -Monotonicity is a necessary and sufficient property for \mathcal{T} -RUMs. The analyst may be interested in evaluating whether the property is satisfied within the sample. In Section 3.2 we provide two procedures for verifying this property in practice. If the property is fully satisfied, then the analyst can directly obtain a distribution over the type space that is consistent with the observed frequencies, as explained in the proof of Theorem 1, immediately after Claim 2.

In general, there are two possible relevant violations of \mathcal{T} -Monotonicity. The first, which involves the choice of dominated alternatives, constitutes a violation of the \mathcal{T} -Extremeness property, shown to be implied by \mathcal{T} -Monotonicity in the proof of Theorem 1. This type of violation cannot be explained by sampling issues, and hence, if observed, the appropriate approach involves the use of the trembling version of the model, studied in Section 5. In this case, one can work with the conditional choice probabilities \bar{p} defined in Section 5, and check whether \mathcal{T} -Monotonicity is satisfied for \bar{p} . Again, if \mathcal{T} -Monotonicity holds over \bar{p} , the analyst can obtain a distribution over types that is consistent with the conditional choice data.

\mathcal{T} -Monotonicity can also be violated in ways other than the violation of \mathcal{T} -Extremeness, perhaps due to finite sampling, as studied in Section 6. This can be evaluated by focusing on the conditional choice probabilities \bar{p} introduced in Section 5, and analyzing \mathcal{T} -Monotonicity using the exact two procedures described above.

In our dataset, it is rather direct that \mathcal{T} -Monotonicity holds neither over the observed choice probabilities nor over the conditional ones. In the case of the former, there are several menus for which some alternatives are not maximal for any type but are nevertheless chosen in the data, which is a violation of \mathcal{T} -Extremeness. In the binary menu $\{l_5, l_9\}$, for example, l_5 first-order stochastically dominates l_9 and l_9 is chosen with probability .17. In the case of the latter, consider the binary menus $\{l_4, l_7\}$ and $\{l_2, l_8\}$. The riskier lottery, l_7 (respectively, l_2), is maximal for types up to type 3 (respectively, type 16). \mathcal{T} -Monotonicity requires that the choice probability of l_7 in $\{l_4, l_7\}$ must be lower than that of l_2 in $\{l_2, l_8\}$. However, the observed choice probabilities are .38 and .22, respectively, which violates the property.¹⁷

Step 5 (Estimation). When \mathcal{T} -Monotonicity is not satisfied, an estimation exercise will find the closest distribution over types, as outlined in Section 6.1. This requires the selection of a particular estimator from the broad class of estimators shown in Section 6.1 to be strongly consistent. In addition, depending on whether \mathcal{T} -Extremeness holds, the estimation may involve the adoption of the trembling version of the model of Section 5. When this is the case, the analyst needs to decide which trembling structure to be adopted from the general class of specifications used in Section 5.

For our analysis, we adopt the standard maximum likelihood estimator, and, since our dataset violates \mathcal{T} -Extremeness, we implement the trembling version of the model. We choose to go with the simplest version and use a constant tremble parameter across all choice problems, and a uniform selection of non-maximal alternatives.¹⁸ Table 2 reports the estimated densities of the \mathcal{T} -RUMT (Columns 4 and 8). The results reveal a high degree of heterogeneity, including a very significant proportion of highly risk-averse types. The percentage of types with close to or higher than logarithmic curvature (type 24 and above) is 60%, with more than a third of all decisions belonging to the highest type, type 30, which exhibits risk-aversion levels as high as 4.7 and above. The

¹⁷Note that these choice probabilities also apply to the conditional ones, since there are no dominated options in these menus.

¹⁸The data and estimation programs are available for use on our websites.

results also show that a relevant proportion of the decisions reflect highly risk-seeking attitudes (22%), and even extreme risk-seeking, (12% of all decisions correspond to the lowest type, type 1, that is, risk-aversion coefficients below -4.8). The estimated probability of tremble is .25, thus confirming the behavioral relevance of non-maximal alternatives.

Step 6 (Statistical testing). The final step is the statistical testing of the model, as elaborated in Section 6.2. This involves comparing the observed choice frequencies with the predictions given by the estimated model, using the Pearson statistic. The analysis can be performed at the menu level with $|A_j| - 1$ degrees of freedom, or at the aggregate level with $\sum_{j \in \mathcal{J}} |A_j| - J$ degrees of freedom.

In the dataset, based on the Pearson statistic per each of the 246 menus, we are unable to reject the null hypothesis of equality between observed and predicted choices in 72% (85%) of the menus at the 5% (1%) significance level.¹⁹ Note that the degrees of freedom in the individual menus are 1, 2 and 4 for the menus with 2, 3 and 5 alternatives. The aggregated Pearson statistic across menus enables us to reject the null hypothesis of equality at conventional significance levels. In this case, the degrees of freedom are 708.

8. RELATED LITERATURE

There is growing theoretical, econometric and empirical interest in random utility models using ordered collections of types. The closest paper to this one is Apesteguia, Ballester and Lu (2017), which provides a theoretical analysis of the single-crossing random utility model (SCRUM).²⁰ The fundamental difference between the two papers is that the earlier one deals with the characterization of the model using a traditional stochastic choice theory approach, while the present one is of a more practical nature, being aimed at addressing the theoretical and econometric challenges posed by the empirical implementation of the model.²¹ Consequently, this paper differs from the

¹⁹Recall that the purpose of this exercise is to illustrate the applicability of the model, and for this purpose we chose the simplest possible implementation, involving a representative agent approach with CRRA expected utilities. Naturally, accounting for inter- and intra-personal variability and using other families of utility functions may improve these results.

²⁰See also Petri (2021) and Valkanova (2021).

²¹The stochastic choice theory literature is turning to the issue of the empirical implementation of the theoretical models developed in the field. For example, Dardanoni, Manzini, Mariotti and

previous one in four main ways. Firstly, it models the conventional approach used in applied work, where the analyst confronts the situation with a favored, fixed, family of ordered types. This lends content to the properties used in our characterization.²² Secondly, whereas SCRUMs impose the single-crossing condition on the grand set of alternatives, X , the assumption here is that the condition holds within each menu, and not necessarily on X .²³ Thirdly, whereas the characterization of SCRUMs uses a universal domain; i.e., it requires data on every single subset of the grand set of alternatives, here we work with arbitrary domains.²⁴ It is thanks to these last two differences that the present paper is closer to the empirical requirements for implementation. Both completeness on X and the universal choice domain are assumptions that facilitate the theoretical treatment of the model, but are very rarely met in practice. With the relaxation of these two assumptions, the model fits the majority of existing datasets. Finally, in line with the practical motivation for this study, the findings enable us to derive econometric results of direct interest for the implementation of the model. In particular, we introduce a class of estimators which are shown to be strongly consistent, and study the statistical testing of the model. All in all, the present paper takes steps towards endowing ordered random utility models with full-fledged empirical content.

In a recent theoretical contribution, Filiz-Ozbay and Masatlioglu (2020) study a random model using an ordered collection of choice functions rather than utilities and thus, importantly, provide the theoretical foundations for what can be considered a model of stochastic, boundedly rational, ordered choice. The main difference between their paper and ours is that we work on the practical implementation of random utility models.

Tyson (2020) study limited-attention stochastic choice models, where the choice domain involves a single menu of alternatives; while Cattaneo, Ma, Masatlioglu and Suleymanov (2020) establish non-parametric identification results for their random attention model.

²²In a similar fashion, Gul and Pesendorfer's study (2006) of random expected utility uses an extremeness property defined based on the class of expected utilities.

²³Indeed, in Section 4 we significantly relax this already mild assumption.

²⁴Note, therefore, that the properties characterizing SCRUMs may even be empty statements in the current domains, and thus useless. For instance, if the data comprise only two disjoint menus, the properties are emptily satisfied because they say nothing about disjoint menus. Our \mathcal{T} -Monotonicity property represents a novel approach to the characterization of ordered choice even in settings like this.

A handful of recent empirical papers focus on trying to exploit the single-crossing condition. Barseghyan, Molinari and Thirkettle (2021) use random utility models satisfying the single-crossing condition to provide a semi-parametric identification of attention models under risk taking. Chiappori, Salanié, Salanié and Gandhi (2019) also impose 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 provides theoretical foundations for the model and its estimation in general settings, beyond that of decisions under risk, while using arbitrary domains.

A series of applied papers have implemented parametric versions of the random utility model, over an ordered collection of types, to estimate a specific behavioral trait; most frequently, risk aversion.²⁵ Barsky, Juster, Kimball and Shapiro (1997) is one of the first examples of the use of this methodology, where the ordered structure of a decision problem involving lotteries is exploited to obtain population estimates of risk aversion and perform covariate analysis. Cohen and Einav (2007) use data on auto insurance contracts, showing that any given probability of accident leads to an ordered decision problem of deductibles and premiums, thereby facilitating the estimation of risk aversion. Andersson, Holm, Tyran and Wengström (2020) use decision problems involving two states with fixed probabilities to show that choice variability is determined by cognitive ability rather than risk aversion. Our paper contributes to this applied literature by providing foundations for a general version of the model.

The econometrics literature on the non-parametric identification of ordered discrete choice models is also of relevance here (see Cunha, Heckman and Navarro (2007), and references cited therein, and Greene and Hensher (2010) for a survey). The papers in this literature focus on identification relying on the existence of a relationship between the probability of choice of any alternative and the mass of types for which the alternative is optimal, which, given the structure, takes the form of an interval. However, it is crucial to stress that there are no axiomatic exercises of any kind in this literature. Hence, the novelty of our paper is to bring the ordered choice logic from the applied and econometrics literature to a revealed preference setting that imposes minimal requirements on the data structure, to provide a novel and easily testable

²⁵See 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.

property, \mathcal{T} -Monotonicity, and to show that it fully axiomatizes the ordered random type model.

9. FINAL REMARKS

In this paper we have provided micro-foundations for ordered random utility models. The consideration of ordered settings is a widely-used approach in applied work. The aim of this paper has been to link actual empirical practice with the revealed preference approach. Moreover, in our attempt to bridge these two literatures, we have adopted a non-parametric approach and worked with arbitrary domains of ordered menus. Furthermore, we have provided a class of strongly consistent, micro-founded estimators, and developed statistical tools for testing the model.

We close by briefly commenting on some differences between \mathcal{T} -RUMs and additive RUMs (ARUMs), also very popular in applied work. First, note that, in a \mathcal{T} -RUM, there is a given ordered collection of utilities, over which the individual is assumed to have a preference distribution. In an ARUM, the analyst assumes that the individual has a particular utility function which is subject to additive, cardinal, shocks. This implies that, effectively, in an ARUM, the individual has a distribution with strictly positive mass over every single possible utility function over the set of alternatives, where the distribution is shaped by some assumption on the shocks (e.g., logistic, normal). Furthermore, \mathcal{T} -RUM requires only an ordinal understanding of the utility functions at stake, while ARUM requires a cardinal interpretation of the utility functions, since shocks enter additively and choice probabilities are determined by cardinal utility differences. Finally, \mathcal{T} -RUMs are monotone, in the sense that shifts in the distribution over types generate intuitive shifts in the choice distributions, thus facilitating the interpretation of the relevant behavioral parameters. However, as shown in Apestegua and Ballester (2018), the typical implementation of ARUM in risk settings, combining expected utility and i.i.d. additive errors, may suffer from severe non-monotonicity.

APPENDIX A. PROOFS

Proof of Lemma 1: In all three of the classes considered, let menu j contain two lotteries, l_1 and l_2 , that are maximal for at least one expected utility type, i.e., $l_1, l_2 \in \{l : \mathcal{T}(l, j) \neq \emptyset\}$. Assume, without loss of generality, that l_1 is maximal for t_1 , and l_2 for t_2 , with $t_1 < t_2$. Hence, the lotteries are not related by first order stochastic

dominance, and, in particular, we know that: in class (1), the good- (respectively, bad-) state payoff is strictly larger (resp, smaller) for l_1 ; in class (2), the probabilities of both the highest and the lowest payoffs are strictly larger for l_1 ; and, in class (3), the probability of winning (resp, the size of) the prize is strictly smaller (resp, larger) for l_1 . Hence, it becomes evident that l_1 and l_2 are related by second order stochastic dominance. Then, by Theorem 1 in Hammond (1974), $l_1 \triangleleft l_2$, and the domain is ordered. \blacksquare

Proof of Theorem 1: The necessity of the axiom is evident and is thus omitted. We now prove its sufficiency, by proceeding through a series of claims. We then assume that the domain is composed of ordered menus, and that p satisfies \mathcal{T} -Monotonicity. The first two claims do not use \mathcal{T} -Monotonicity; they follow exclusively from the ordered structure of the domain.

Claim 1. For every menu $j \in \mathcal{J}$ and alternative $x \in A_j$, $\mathcal{T}(x, j)$ is an interval.

Proof of Claim 1: Suppose, by way of contradiction, that the claim is false. Let (x, j) be a pair such that types $t_1 < t_2 < t_3$ exist, with $\{t_1, t_3\} \subseteq \mathcal{T}(x, j)$, but $t_2 \notin \mathcal{T}(x, j)$. Let $z \in A_j$ be the alternative for which $t_2 \in \mathcal{T}(z, j)$. From the joint consideration of types t_1 and t_2 , it must be that $z \not\triangleleft x$. From the joint consideration of types t_2 and t_3 , it follows that $x \not\triangleleft z$. Given that it must be that $x \neq z$, this contradicts the fact that menu j is ordered, and concludes the proof of the claim. \square

Claim 2. For every menu $j \in \mathcal{J}$ and alternative $x \in A_j$ such that $\mathcal{T}(x, j) \neq \emptyset$, $\bigcup_{y \in A_j, y \triangleleft x} \mathcal{T}(y, j) = \{1, 2, \dots, \max \mathcal{T}(x, j)\}$.

Proof of Claim 2: Suppose, by way of contradiction, that the claim is false. Then, there exists t^* such that either: (i) $t^* \leq \max \mathcal{T}(x, j)$ and $t^* \notin \bigcup_{y \in A_j, y \triangleleft x} \mathcal{T}(y, j)$, or (ii) $t^* > \max \mathcal{T}(x, j)$ with $t^* \in \bigcup_{y \in A_j, y \triangleleft x} \mathcal{T}(y, j)$ hold. In both cases, let $z \in A_j$ be the alternative for which $t^* \in \mathcal{T}(z, j)$. In case (i), we have $z \not\triangleleft x$ by assumption, implying that t^* and $\max \mathcal{T}(x, j)$ must be different types. Their joint consideration guarantees, furthermore, that $x \not\triangleleft z$, thus contradicting the fact that menu j is ordered. In case (ii), we know, by assumption, that $z \triangleleft x$. Since $t^* > \max \mathcal{T}(x, j)$, it must be that $x \neq z$ and $u_{t^*}(x) > u_{t^*}(z)$, which contradicts the assumption that $t^* \in \mathcal{T}(z, j)$, thus concluding the proof of the claim. \square

We now consider the sub-collection of types $\mathcal{T}^I \subseteq \mathcal{T}$ and the correspondence $F : \mathcal{T}^I \rightrightarrows [0, 1]$ defined by:

$$\mathcal{T}^I = \{t \in \mathcal{T} : \text{there exists } (x, j) \text{ such that } \max \mathcal{T}(x, j) = t\}, \text{ and}$$

$$k \in F(t) \text{ whenever there is } (x, j) \text{ such that } t = \max \mathcal{T}(x, j) \text{ and } k = \sum_{y \in A_j, y \preceq x} p(y, j).$$

Claim 3. F is a single-valued increasing map.

Proof of Claim 3: To see this, consider two types $t, t' \in \mathcal{T}^I$ such that $t \leq t'$. By definition of \mathcal{T}^I , there exist pairs (x, j) and (x', j') such that $t = \max \mathcal{T}(x, j)$ and $t' = \max \mathcal{T}(x', j')$. By Claim 2, and the fact that $t \leq t'$, we know that $\bigcup_{z \in A_j, z \preceq x} \mathcal{T}(z, j) = \{1, 2, \dots, t\} \subseteq \{1, 2, \dots, t'\} = \bigcup_{z \in A_{j'}, z \preceq x'} \mathcal{T}(z, j')$. The use of \mathcal{T} -Monotonicity guarantees that $\sum_{z \in A_j, z \preceq x} p(z, j) \leq \sum_{z \in A_{j'}, z \preceq x'} p(z, j')$, with equality when $t = t'$, which proves the claim. \square

Claim 4. \mathcal{T} -Monotonicity implies the following property, which we call \mathcal{T} -Extremeness: $p(x, j) > 0$ implies $\mathcal{T}(x, j) \neq \emptyset$.²⁶

Proof of Claim 4: Let $\mathcal{T}(x, j) = \emptyset$. Hence, $\bigcup_{y \in A_j} \mathcal{T}(y, j) = \mathcal{T} \subseteq \mathcal{T} = \bigcup_{y \in A_j \setminus \{x\}} \mathcal{T}(y, j)$.

Then, by \mathcal{T} -Monotonicity, we have $\sum_{y \in A_j} p(y, j) = 1 \leq \sum_{y \in A_j \setminus \{x\}} p(y, j)$. Given that p is a stochastic choice function, $\sum_{y \in A_j \setminus \{x\}} p(y, j) \leq 1$, and, given the last inequality, it must in fact be equal to 1. Consequently, $p(x, j) = 0$, proving the claim. \square

Claim 5. $T \in \mathcal{T}^I$, with $F(T) = 1$.

Proof of Claim 5: To see the first part, consider any menu j and let x be the alternative such that $T \in \mathcal{T}(x, j)$. It can only be the case that $\max \mathcal{T}(x, j) = T$, and hence, $T \in \mathcal{T}^I$. To see the second part, consider any menu j , and any alternative $y \in A_j$ such that $\mathcal{T}(y, j) \neq \emptyset$. Since menu j is ordered, the joint consideration of types $\max \mathcal{T}(y, j)$ and $\max \mathcal{T}(x, j) = T$ guarantees that $y \preceq x$. Hence, the use of \mathcal{T} -Extremeness and the definition of stochastic choice function guarantee that $1 = \sum_{y \in A_j} p(y, j) \geq \sum_{y \in A_j, y \preceq x} p(y, j) = F(t) \geq \sum_{y \in A_j: \mathcal{T}(y, j) \neq \emptyset} p(y, j) = 1$, which proves the claim. \square

²⁶Gul and Pesendorfer (2006) use a similar property in their study of random expected utility models.

Given Claims 3 and 5, we are able to construct a map $G : \mathcal{T} \rightarrow [0, 1]$ that extends F , i.e. $F(t) = G(t)$ for every $t \in \mathcal{T}^I$, and is weakly increasing, i.e. $t_1 < t_2$ implies that $G(t_1) \leq G(t_2)$. Trivially, $\psi(t) = G(t) - G(t - 1)$, with the notational assumption $G(0) = 0$, is a probability distribution over \mathcal{T} . We then consider the \mathcal{T} -RUM defined by ψ .

Claim 6. For every menu $j \in \mathcal{J}$, and every alternative $x \in A_j$, $p(x, j) = \psi(\mathcal{T}(x, j))$.

Proof of Claim 6: If $\mathcal{T}(x, j) = \emptyset$, we know that \mathcal{T} -Monotonicity implies \mathcal{T} -Extremeness, which, in turn, guarantees that $p(x, j) = 0$, which is precisely the probability assigned by the \mathcal{T} -RUM. Whenever $\mathcal{T}(x, j) \neq \emptyset$, Claim 1 guarantees that $\mathcal{T}(x, j)$ is an interval. If $1 \in \mathcal{T}(x, j)$, we know, by construction, that $p(x, j) = F(\max \mathcal{T}(x, j)) = G(\max \mathcal{T}(x, j)) = \psi(\mathcal{T}(x, j))$, as desired. If $1 \notin \mathcal{T}(x, j)$, let z be the highest alternative in A_j , according to \triangleleft , satisfying $\mathcal{T}(z, j) \neq \emptyset$ and $z \triangleleft x$. It must obviously be the case that $\mathcal{T}(x, j) = \{\max \mathcal{T}(z, j) + 1, \max \mathcal{T}(z, j) + 2, \dots, \max \mathcal{T}(x, j)\}$ and, by construction, $p(x, j) = F(\max \mathcal{T}(x, j)) - F(\max \mathcal{T}(z, j)) = G(\max \mathcal{T}(x, j)) - G(\max \mathcal{T}(z, j)) = \psi(\mathcal{T}(x, j))$. This proves the claim. \square

Having constructed a \mathcal{T} -RUM that rationalizes all choice probabilities, we have proved the sufficiency of the property and hence the theorem. \blacksquare

Proof of Corollary 1: Consider domain \mathcal{D}_{MM} and denote by \mathcal{S} the set of all strictly positive real numbers (slopes) s such that there exists $A_j \in \mathcal{D}_{MM}$ and a lottery $q^{j,k} \in \bar{A}_j$, with $k < K_j$, such that $s = \frac{q_a^{j,k} - q_a^{j,k+1}}{q_c^{j,k} - q_c^{j,k+1}}$. Write \mathcal{S} as an ordered set $\{s_1, s_2, \dots, s_{T-1}\}$ in a strictly increasing way. We can then consider any sequence of strictly positive numbers $\{s'_1, s'_2, \dots, s'_T\}$ with the property that $s'_1 < s_1 < s'_2 < s_2 < \dots < s'_t < s_t < \dots < s'_{T-1} < s_{T-1} < s'_T$. Recall that in a Marschak-Machina triangle, any expected utility is fully determined by the strictly positive slope of the linear indifference curves that it creates over the triangle, with indifference curves improving monotonically with the value of q_a . Hence, for each $t \in \{1, 2, \dots, T\}$, we can select an expected utility function EU_t with linear indifference curves of slope s'_t . We have then constructed a collection of types $\{EU_t\}_{t \in \mathcal{T} = \{1, 2, \dots, T\}}$. Given the construction, these utilities are ordered by increasing risk aversion and, from Lemma 1, we know that the domain is ordered in relation to this set of types.

We now prove that p satisfies \mathcal{T} -Extremeness. To see this, consider any menu $j \in \mathcal{D}_{MM}$ and assume that $p(q, j) > 0$. From Property 1, we know that there exists $q^{j,k} \in \bar{A}_j$ such that $q = q^{j,k}$. If $K_j = 1$, then q is the only lottery in \bar{A}_j and the construction guarantees that q is chosen by all types. If $k = 1 < K_j$, then it is immediate from the construction that q is chosen for types $1, 2, \dots, t$, where s_t is the slope determined by $q = q^{j,k}$ and $q^{j,k+1}$. If $k = K_j > 1$, then it is immediate from the construction that q is chosen for types $t + 1, \dots, T$, where s_t is the slope determined by $q^{j,k-1}$ and $q = q^{j,k}$. Finally, in any other case, q is chosen for types $t_1 + 1, t_1 + 2, \dots, t_2$, where s_{t_1} is the slope determined by $q^{j,k-1}$ and $q = q^{j,k}$ and t_2 is the slope determined by $q = q^{j,k}$ and $q^{j,k+1}$. In any case, we have proved that \mathcal{T} -Extremeness is satisfied.

We now prove that p satisfies \mathcal{T} -Monotonicity* and, hence, \mathcal{T} -Monotonicity. To see this, notice that for any lottery $q \in \bar{A}_j$, $\max \mathcal{T}(q, j)$ is, by construction, equal to the type t such that $s_t = \frac{q_a^{j,k} - q_a^{j,k+1}}{q_c^{j,k} - q_c^{j,k+1}}$ (or T if $q = q^{j,K_j}$). Similarly, given the ordered domain, lotteries lower than q with non-zero choice probability must belong to the collection $\{q^{j,1}, \dots, q^{j,k}\}$. Hence, Property 2 guarantees \mathcal{T} -Monotonicity* holds for the specific types chosen.

From Theorem 1 and the discussion in Section 3.2, we know that there exists a distribution over $\{EU_t\}_{t \in \mathcal{T} = \{1, 2, \dots, T\}}$ that explains all choices. We have proven the existence of a REUM explaining all choices and thus the sufficiency of the theorem. Necessity is immediate and the result follows. \blacksquare

Proof of Corollary 2: The proof is similar to that of Corollary 1, with the main difference being related to the treatment of corner solutions with Cobb-Douglas preferences. Consider the domain \mathcal{D} , and denote by \mathcal{S} the set of all strictly positive real numbers s for which there exists $A_j \in \mathcal{D}$ and an interior bundle $x^{j,l} \in \bar{I}_j$, $l < L_j$, such that $s = \frac{y_2^{j,l+1} - y_2^{j,l}}{y_1^{j,l} - y_1^{j,l+1}}$. Write \mathcal{S} as an ordered set $\{s_2, s_3, \dots, s_{T-2}\}$ in a strictly increasing way.²⁷ We can then consider any sequence of strictly positive numbers $\{s'_2, s'_3, \dots, s'_{T-1}\}$ with the property that $s'_2 < s_2 < s'_3 < s_3 < \dots < s'_t < s_t < \dots < s'_{T-2} < s_{T-2} < s'_{T-1}$. Notice that, when interior bundles are logarithmically transformed, any Cobb-Douglas with $\alpha \in (0, 1)$ is fully determined by the slope of the linear indifference curves that

²⁷Intuitively, \mathcal{S} includes the absolute value of the slope of every segment joining one interior maximal bundle with the consecutive interior maximal bundle, both expressed in logarithmic terms, within a menu.

it creates over the logarithmically-transformed bundles, with indifference curves improving monotonically with the amount of both goods. We can select a Cobb-Douglas, strictly monotone in both goods, CD_t with linear indifference curves of slope s'_t . We add the two Cobb-Douglas determined by weights $\alpha = 0$ and $\alpha = 1$, that we denote as CD_1 and CD_T , to the collection. We have then constructed a collection of types $\{CD_t\}_{t \in \mathcal{T}=\{1,2,\dots,T\}}$ and, given the construction, these utilities are ordered by increasing relevance of good 1. From the discussion in Section 2.3, we know that the domain is ordered in relation to this set of types.

From the above construction, and the analogous argument to the one followed in the proof of Corollary 1, it is evident that Property 3 guarantees that p satisfies \mathcal{T} -Extremeness.²⁸ To see that p satisfies \mathcal{T} -Monotonicity*, consider the set \bar{A}_j containing all possible maximal alternatives in a menu for our set of types. Consider menus $A_j, A'_j \in \mathcal{D}$ and alternatives $x = a^{j,k} \in \bar{A}_j, x' = a^{j',k'} \in \bar{A}_{j'}$, with $k < K_j$ and $k' < K_{j'}$ and suppose that $\max \mathcal{T}(x, j) \leq \max \mathcal{T}(x', j')$. The construction of the family of utilities guarantees that $\hat{r}(a^{j,k}, j) \leq \hat{r}(a^{j',k'}, j')$ and, by Property 4, this implies $\sum_{h \leq k} p(a^{j,h}, j) \leq \sum_{h' \leq k'} p(a^{j',h'}, j')$. Thus, \mathcal{T} -Monotonicity* holds and from Theorem 1 and the discussion in Section 3.2, we know that there exists a distribution over $\{CD_t\}_{t \in \mathcal{T}=\{1,2,\dots,T\}}$ that explains all choices. We have proven the existence of a RCDM explaining all choices and thus the sufficiency of the theorem. Necessity is immediate and the result follows. ■

Proof of Corollary 3: Claim 1 in the proof of Theorem 1 is now an assumption and can be used to define menu-dependent relations \triangleleft_j given by: $x \triangleleft_j y$ if there exist $t \in \mathcal{T}(x, j)$ and $t' \in \mathcal{T}(y, j)$ such that $t < t'$. Notice that the corresponding \trianglelefteq_j is complete on $\{x : \mathcal{T}(x, j) \neq \emptyset\}$ and can replace \trianglelefteq in the definition of F . The rest of the proof is analogous and thus, omitted. ■

Proof of Corollary 4: The proof of Corollary 3 can be reproduced by working exclusively with interval menus, leaving us only to prove that the constructed ψ explains the choices in non-interval menus. Consider one such menu j and alternative $x \in A_j$. The only non-trivial case is $\mathcal{T}(x, j) \neq \emptyset$, where the replica menu

²⁸The only relevant difference is that corner options can only be selected by CD_1 and CD_T .

j' can be considered. By part (ii) of the definition of replica menus, we can identify a subset of alternatives $B \subseteq A_{j'}$ such that $\mathcal{T}(x, j) = \bigcup_{y \in B} \mathcal{T}(y, j')$. By \mathcal{T} -Monotonicity $p(x, j) = \sum_{y \in B} p(y, j')$. By part (i) of the definition of replica, we know that j' is an interval menu, and thus, all of its choices are rationalized by ψ . Hence, $\sum_{y \in B} p(y, j') = \psi(\bigcup_{y \in B} \mathcal{T}(y, j')) = \psi(\mathcal{T}(x, j))$, which concludes the proof. ■

Proof of Proposition 1: Consider menu j as inducing the sequence of type intervals $\mathcal{T}(x_1, j) = \{1, 2, \dots, t_1\}$, $\mathcal{T}(x_2, j) = \{t_1 + 1, t_1 + 2, \dots, t_2\}$, \dots , $\mathcal{T}(x_\kappa, j) = \{t_{\kappa-1} + 1, t_{\kappa-1} + 2, \dots, T\}$, where the sequence of chosen lotteries $x_1, x_2, \dots, x_\kappa$ may include repetitions, thereby violating the interval assumption. Given the Marschak-Machina triangle, denote by $s(t)$ the (strictly positive) slope of the indifference curve of type t . Notice that $s(t)$ must be strictly increasing in t , because the types are ordered by risk aversion. Now, construct an interval menu $A_{j'} = \{y_1, \dots, y_\kappa\}$ in the Marschak-Machina triangle as follows. Let the vector of probabilities of lottery y_1 be $q_1 = (\frac{1}{2}, 0, \frac{1}{2})$ and, for $k \in \{2, \dots, \kappa\}$, let the vector of probabilities of lottery y_k be $q_k = q_{k-1} + (-\alpha_{k-1}, \alpha_{k-1} + \beta_{k-1}, -\beta_{k-1})$, with $s(t_{k-1}) < \frac{\alpha_{k-1}}{\beta_{k-1}} < s(t_{k-1} + 1)$ 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 within the triangle. It is then evident that lottery y_k is maximal for, and only for, types $\{t_{k-1} + 1, \dots, t_k\}$, making j' an interval menu. Hence, it follows that for every lottery x in j it is the case that $\mathcal{T}(x, j) = \bigcup_{k: x_k = x} \mathcal{T}(x_k, j) = \bigcup_{k: x_k = x} \mathcal{T}(y_k, j')$, and therefore j' is a replica of j for alternative x , as desired. ■

Proof of Corollary 5: Necessity is immediate. To see sufficiency, notice that \bar{p} must be a \mathcal{T} -RUM. The techniques in the proof of Theorem 1 can be used to construct the corresponding distribution ψ , and the definition of $\lambda_j = \sum_{x \in D_j} p(x, j)$ completes the \mathcal{T} -RUMT. The claim then follows immediately. ■

Proof of Corollary 6: We start by proving the first part. For a single menu $j \in \mathcal{J}$, notice that the set of distributions that rationalizes data in j is nonempty. As a result, the existence of a minimum value of ϵ such that z is ϵ -rationalizable follows directly from the compactness of $\Psi \times \Psi^J$ and the continuity of d and f . Moreover, if there exists a distribution ψ rationalizing all choices across menus, 0-rationalizability holds due to the fact that $f(d(\psi, \psi), \dots, d(\psi, \psi)) = f(0, \dots, 0) = 0$. Finally, if such a distribution

does not exist, the data cannot be 0-rationalizable because of the strict monotonicity properties of both d and f . This concludes the proof. ■

Proof of Theorem 2: Given any distribution over types ψ and data frequencies \tilde{z} , consider the value $g(\psi, \tilde{z}) = \min f(d(\psi, \psi_1), \dots, d(\psi, \psi_J))$, subject to ψ_j rationalizing the choice probabilities \tilde{z}_j . Note that, using the same logic as in Corollary 6, g is well-defined. Consider a sequence of data functions $\{z^n\}_{n=1}^\infty$ with $\lim_{n \rightarrow \infty} Z_j^n = \infty$ for every $j \in \mathcal{J}$. Given the definition of the estimator and the properties of f and d , it follows that the estimator for z^n is $\hat{\psi}^n = \arg \min_{\psi \in \Psi} g(\psi, \tilde{z}^n)$.

Now suppose that the sequence of data functions is generated by a \mathcal{T} -RUM with probability distribution $\psi^* \in \Psi$. Consider menu j . For every alternative $x \in A_j$ such that $\psi^*(\mathcal{T}(x, j)) = 0$, either because $\mathcal{T}(x, j) = \emptyset$ or because no mass is associated to the types for which x is maximal, we know that $\tilde{z}^n(x, j) = 0$ always holds. For every alternative $x \in A_j$ such that $\psi^*(\mathcal{T}(x, j)) > 0$, standard arguments guarantee that the multinomial i.i.d. choices in menu j generate frequencies $\tilde{z}^n(x, j)$ that converge, almost surely, to $\psi^*(\mathcal{T}(x, j))$. Thus, the finiteness of each menu and of \mathcal{J} guarantees that \tilde{z}^n converges, almost surely, to the choice probabilities generated by ψ^* . It then follows immediately that $\hat{\psi}^n$ converges, almost surely, to ψ^* . This concludes the proof. ■

Proof of Theorem 3: As discussed in the proof of Theorem 2, the i.i.d. nature of the \mathcal{T} -RUM guarantees that we can understand choices in any given menu j as a multinomial distribution where entry x has probability $\psi(\mathcal{T}(x, j))$. It is well known that, as the number of observations grows, the statistic $\sum_{x \in A_j} \frac{(z(x, j) - Z_j \psi(\mathcal{T}(x, j)))^2}{Z_j \psi(\mathcal{T}(x, j))}$ converges to a chi-square distribution with $|A_j| - 1$ degrees of freedom. The i.i.d. nature of the model also guarantees that, even though their parameters are linked by ψ and the domain structure is ordered, the multinomial distributions across menus are independent. Hence, it follows immediately that $C(z)$ converges to a chi-square with $\sum_{j \in \mathcal{J}} (|A_j| - 1) = \sum_{j \in \mathcal{J}} |A_j| - J$ degrees of freedom, as desired. ■

REFERENCES

- [1] Andersson, O., H.J. Holm, J.R. Tyran and E. Wengström (2020). “Robust Inference in Risk Elicitation Tasks,” *Journal of Risk and Uncertainty*, 61:195–209.
- [2] Apesteguia, J. and M.A. Ballester (2018). “Monotone Stochastic Choice Models: The Case of Risk and Time Preferences,” *Journal of Political Economy*, 126(1):74–106.

- [3] Apesteguia, J. and M.A. Ballester (2021). “Separating Predicted Randomness from Residual Behavior,” *Journal of the European Economic Association*, 19(2):1041–1076.
- [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 (2021). “Discrete Choice under Risk with Limited Consideration,” *American Economic Review*, 111(6):1972–2006.
- [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] Camerer, C., (1989). “An Experimental Test of Several Generalized Utility Theories”, *Journal of Risk and Uncertainty*, 2: 61-104
- [9] Cattaneo, M., X. Ma, Y. Masatlioglu and E. Suleymanov (2020). “A Random Attention Model,” *Journal of Political Economy*, 128(7):2796–2836.
- [10] Chiappori, P-A, B. Salanié, F. Salanié and A. Gandhi (2019). “From Aggregate Betting Data to Individual Risk Preferences,” *Econometrica*, 87(1):1–36.
- [11] 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.
- [12] Cohen, A. and L. Einav (2007). “Estimating Risk Preferences from Deductible Choice,” *American Economic Review*, 97(3):745–788.
- [13] Coller, M. and M.B. Williams (1999). “Eliciting Individual Discount Rates,” *Experimental Economics*, 2(2):107–27.
- [14] Cunha, F., J.J. Heckman and S. Navarro (2007). “The Identification and Economic Content of Ordered Choice Models with Stochastic Thresholds,” *International Economic Review*, 48(4):1273–1309.
- [15] Dardanoni, V., P. Manzini, M. Mariotti and C.J. Tyson (2020). “Inferring Cognitive Heterogeneity from Aggregate Choices,” *Econometrica*, 88(3):1269–1296.
- [16] Filiz-Ozbay, E. and Y. Masatlioglu (2020). “Progressive Random Choice,” mimeo.
- [17] Greene, W.H. and D.A. Hensher (2010). *Modeling Ordered Choices*, Cambridge University Press.
- [18] Gul, F. and W. Pesendorfer (2006). “Random Expected Utility,” *Econometrica*, 74(1):121–146.
- [19] Hammond, J.S. (1974). “Simplifying the Choice between Uncertain Prospects Where Preference is Nonlinear,” *Management Science* 20(7):1047–1072.
- [20] Holt, C.A. and S.K. Laury (2002). “Risk Aversion and Incentive Effects,” *American Economic Review*, 92(5):1644–1655.
- [21] Jagelka, T. (2020). “Are Economists’ Preferences Psychologists’ Personality Traits? A Structural Approach,” mimeo.
- [22] Milgrom, P., and C. Shannon (1994). “Monotone Comparative Statics,” *Econometrica*, 62(1):157–180.

- [23] Mirrlees, J.A. (1971). “An Exploration in the Theory of Optimum Income Taxation,” *The Review of Economic Studies*, 38(2):175–208.
- [24] Petri, H. (2021). “Binary Single-Crossing Random Utility Models,” mimeo.
- [25] Spence, A.M. (1974). *Market Signaling: Information Transfer in Hiring and Related Screening Processes*, Harvard University Press.
- [26] Tanaka, T., Camerer, C. and Q. Nguyen (2010). “Risk and Time Preferences: Linking Experimental and Household Survey Data from Vietnam,” *American Economic Review*, 100:1, 557–571.
- [27] Topkis, D.M. (1978). “Minimizing a Submodular Function on a Lattice,” *Operations Research*, 26: 305–321.
- [28] Valkanova, K. (2021). “Inferring Single-Peakedness from Random Choice,” mimeo.
- [29] van der Vaart, A.W. (2000), *Asymptotic Statistics*, Cambridge University Press.
- [30] Warner, J.T. and S. Pleeter (2001). “The Personal Discount Rate: Evidence from Military Downsizing Programs,” *American Economic Review*, 91(1):33–53.