TESTABLE IMPLICATIONS OF PARAMETRIC ASSUMPTIONS IN ORDERED RANDOM UTILITY MODELS

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

Abstract. We establish the testable implications of ordered random utility models under nonparametric, semi-nonparametric, and parametric assumptions. After characterizing these models in a continuous setting and in the absence of fundamental parametric restrictions, we show that assumptions about the type distribution alone are immaterial, while assumptions about the map linking types to utilities are relevant only insofar as they restrict the class of utilities at stake. Importantly, the joint presence of such parametric assumptions, as is common practice, further restricts the empirical content of the model. We then provide a characterization of commonly-used parametric ordered-logit models. We apply our results, both theoretically and empirically, to economically relevant settings.

Keywords: Ordered random utility model; empirical content; parametric restrictions; ordered logit; cumulative logit; cumulative log-odds.
JEL classification numbers: C00; D00.

1. Introduction

In many settings, decision problems are ordered and variation in choices can be intuitively explained as the result of variation in an underlying ordered latent variable. This simple structure is a fundamental instrument for empirical research, spanning diverse economic areas such as health, finance, labor, welfare, management, insurance, political economy, networks, and gender.[1] To formalize this notion, consider a latent variable on the real line, and let us refer to each of its values as a type. In what we refer to as Ordered Random Utility Models (ORUMs), choice data is the result of

Date: October, 2024.

*This paper supersedes our paper *Choice-based Foundations of Ordered Logit.* Financial support by FEDER/Ministerio de Ciencia e Innovación (Agencia Estatal de Investigación) through Grant PID2021-125538NB-I00 and through the Severo Ochoa Programme for Centers of Excellence in R&D (Barcelona School of Economics CEX2019-000915-S), and Balliol College is gratefully acknowledged. We thank Carlos Gonzalez-Perez for outstanding research assistance. Codes to replicate all the empirical results in this paper are available upon request.

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

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

[1]See, respectively, Barsky, Juster, Kimball, and Shapiro (1997), Kaplan and Zingales (1997), Blau and Hagy (1998), Campante and Yanagizawa-Drott (2015), Cummings (2004), Cohen and Einav (2007), Besley and Persson (2011), Bailey, Cao, Kuchler, and Stroebel (2018), and Carlana (2019).

the interaction of the following two components: (i) a type-utility map where higher types are associated with utility functions generating higher choices, and (ii) a type distribution describing the prevalence of each type.

The empirical use of ORUMs often involves a number of parametric assumptions on one or both of the components of the model. To illustrate, regarding the first component, applications to political or insurance choices are often built by referring to type t as the Euclidean utility centered at t or the expected utility with a relative risk aversion coefficient equal to t, respectively.[2] Regarding the second component, the most common assumptions may involve logistic (or Gaussian) variation in the latent variable, leading to the so-called ordered-logit (or ordered-probit) models. And in many cases, both components are jointly restricted, leading to a fully parametric model. It is then critical to carefully examine when and how parametric specifications bear down on the space of datasets that can be explained by ORUMs and, in each case, what type of properties characterize the corresponding model. This is the purpose of this paper.

We present our results within the context of cumulative choice data over an arbitrary collection of continuous decision problems. We start, in Section 3, with the nonparametric case and show that the model can be understood by means of a standard, deterministic notion of rationalizability. In a nutshell, suppose that for each probability value $p \in (0,1)$, we construct from data the quantile choice function c^p, i.e., c^p_j represents the alternative that first attains cumulative choice above p in decision problem A_j. Theorem 1 shows that data can be explained by an ORUM if and only if every quantile choice function is acyclical. The proof is rather direct but, fundamentally for our purposes, it is instrumental for understanding the exact content of semi-nonparametric ORUMs and flexible enough to be used in a variety of applied settings.

Section 4 is devoted to the semi-nonparametric analysis. Building upon the proof of Theorem 1, we show in Theorem 2 that constraining only the type distribution comes at no cost; as long as the type-utility map is not fixed, the analyst can freely reduce the dimensionality of the statistical distribution describing the latent variable. We then consider the empirical content of semi-nonparametric models that constrain the type-utility map but not the type distribution. Naturally, this restriction is relevant because we are now forced to use only the sub-class of utilities specified by the parametrization, and this limits our ability to explain patterns of choice across menus. However, we show in Theorem 3 that, modulo the sub-class of utilities in the parametrization, the assumption brings no further cost. Formally, if the type distribution remains unrestricted, fixing the type-utility map merely requires properly adjusting the deterministic rationalizability property used in Theorem 1 for it to apply to the corresponding sub-class of utilities in the parametrization.

[2] This type of parametric restriction is often referred to as semi-nonparametric (see, e.g., Barseghyan, Molinari, and Thirkettle (2021)). To simplify our exposition, we use the same term when the dimensionality of one, but not the other, component is constrained.

Section 5 brings the fully parametric analysis. In essence, fixing both the type-utility map and the type distribution fundamentally restricts the model, and does so in ways that cannot be readily described, even with the adjustment in Theorem 3. As a result, we are forced to analyze the properties of parametric ORUMs independently. Given the preponderance of ordered-logit models in the empirical literature, we study ORUMs in which, having fixed a type-utility map, the type distribution is assumed to be logistic.[3] We show in Theorem 4 that two simple properties, which we call corner extremeness (CE) and cumulative log-odds additivity (CLA), characterize these models. CE imposes that a corner alternative can receive a non-null choice probability if and only if this alternative is not dominated, i.e., considered sub-optimal by all utilities in the class. CLA uses the well-known notion of cumulative log-odds, i.e., the log-ratio of masses below and above a given alternative. The property states that equal sums of types must lead to equal sums of cumulative log-odds. Theorem 4 establishes the first characterization result of ordered-logit models, giving foundations to a popular tool in the empirical literature.

To illustrate our results from a more applied point of view, we study in detail two distinct applications to choices over arbitrary collections of linear budget sets: the first involving lotteries, and the second involving inter-personal allocations. Section 6 elaborates on the application to lotteries, with Corollary 1 providing nonparametric (as in Theorem 1) and semi-nonparametric results (as in Theorem 3) for ORUMs using expected utilities and CRRA expected utilities, respectively. Building upon Theorem 4, Corollary 2 characterizes the ordered-logit model that arises from logistic variation of the CRRA coefficient. Section 7 reproduces this analysis for the case of inter-personal allocations, where the semi-nonparametric and parametric cases adopt the CES functional form in modeling variation in altruism.

Sections 6 and 7 also provide comprehensive guidelines on how to bring the aforementioned theoretical results to data. We showcase our guidelines using simulated choice data and provide a step-by-step description that elaborates on: (i) the handling of data, (ii) the study of the nonparametric model, expanding on the computational convenience of the analysis, (iii) the study of the semi-nonparametric model, including statistical testing, and (iv) the study of the parametric model, its structural estimation, and its statistical testing.

2. Related literature

This paper contributes to the study of stochastic choice models in general, and to the literature on random utility models (RUMs) in particular. Classic works include Luce (1959), Block and Marschak (1960), and McFadden and Richter (1990). Here, we study RUMs with an ordered structure; see Small (1987) for an early study, Train (2009) for

[3]The majority of works cited in footnote 1 adopt the ordered-logit format, which gives a good sense of the popularity of the model.

an econometric treatment, and Greene and Hensher (2010) for the applicability of this type of model to a variety of economic questions. RUMs with an ordered structure have been theoretically discussed from different points of view in Apesteguia, Ballester, and Lu (2017), Barseghyan, Molinari, and Thirkettle (2021), Turansick (2022), Valkanova (2022), Apesteguia and Ballester (2023), Filiz-Ozbay and Masatlioglu (2023), Masatlioglu and Vu (2023), Petri (2023, 2024), and Yildiz (2024).

Part of this paper is devoted to establishing choice-based foundations for ORUMs with varying degrees of parametric assumptions.[4] Apesteguia, Ballester, and Lu (2017) provide the first axiomatization of a RUM with an ordered structure, which they call SCRUM. Their analysis is discrete and nonparametric, with unrestricted variation over single-crossing families of utilities. Their main result can be seen as a specific case of Theorem 1 because, for their analysis to hold, they need to observe data on every subset of a linearly ordered set X. Also within a discrete setting, Apesteguia and Ballester (2023) relax data requirements by working, as in the current paper, with arbitrary domains of ordered menus. Their main results, however, operate exclusively in the semi-nonparametric realm, in which the type-utility map is fixed. The present paper contributes to this literature in several ways. First, we establish axiomatic foundations for nonparametric, semi-nonparametric, and parametric versions of ORUMs, all with arbitrary data. Second, by adopting a common setting across all parametric versions, we are able to establish the differential empirical implications of parametric assumptions. Third, we provide, for the first time, an axiomatic treatment of the often-used parametric ordered logit, an analysis that requires the development of novel techniques drawn from the statistical literature. Finally, we adopt the empirically relevant but severely understudied continuous setting.[5]

Another relevant strand of literature pertains to the nonparametric identification of models that are non-additive in the error term. The connection is most evident in Matzkin (2003), where a model of the form $Y = m(X, \alpha)$, with observables X, Y and a non-additive unobserved error term α, is considered. The function m is assumed to be monotone in the latent variable α. Our framework follows this approach by considering a function m that selects the element Y optimizing the utility function of type α, given the menu characteristics X.[6] The main result in Matzkin (2003)

[4]The literature providing choice-based foundations for other stochastic models is extensive. Some recent contributions include Gul and Pesendorfer (2006), Manzini and Mariotti (2014), Caplin and Dean (2015), Fudenberg, Iijima, and Strzalecki (2015), Matejka and McKay (2015), Brady and Rehbeck (2016), Ahn, Echenique, and Saito (2018), Lu and Saito (2018), Cerreia-Vioglio, Dillenberger, Ortoleva, and Riella (2019), Frick, Iijima, and Strzalecki (2019), Natenzon (2019), Cattaneo, Ma, Masatlioglu, and Suleymanov (2020), Alós-Ferrer, Fehr, and Netzer (2021), Kovach and Tserenjigmid (2022), or He and Natenzon (2023).

[5]Our results can also be derived in the discrete choice setting. Appendix B does so and discusses in more detail the connection with the mentioned papers.

[6]One of the examples discussed in Matzkin (2003) is that of random demand over linear budget sets, which is closely related to the settings in our applications.

relevant to our purposes, Lemma 1, addresses the joint identification of the function m and the distribution of the error term α; essentially, monotonic transformations of the pair (m, α) yield the same data. Matzkin (2003) uses this result to explore several useful normalizations of the pair (m, α) across a range of econometric models, including cases where the distribution of α is selected and fixed. Despite the technical differences between Matzkin's paper and ours, the logic of her Lemma 1 is crucial in extending our Theorem 1 to our Theorem 2; we elaborate later in the paper on how Theorem 2 can be viewed as a combination of Theorem 1 and Matzkin's normalization analysis. Since the primary focus of the current paper is characterization rather than identification, we contribute to this literature in several ways. Most notably, we provide a precise description of the empirical content of nonparametric, semi-parametric, and parametric models, both in abstract terms and in applied contexts such as risk and altruism.

Another important strand of literature concerns the analysis of Random Utility Models (RUMs) in consumer settings. These works aim to simplify the empirical analysis of RUMs.[7] Although there are exceptions, such as Hoderlein and Stoye (2014, 2015), most of the literature focuses on general distributions of types, which are not necessarily ordered. A central message from this literature, pioneered by Kitamura and Stoye (2018), is that determining stochastic rationalizability in a demand setting can be accomplished by analyzing a finite set of points from the cumulative distribution functions that describe the data.[8] The technical details in Theorem 1, along with the applied results in Corollaries 1 and 3, and the political economy example discussed in Section 3.1, shed light on the possibility of extending this principle to other settings. In essence, the finite number of checks arises from the type of rationalizability used in consumer settings (e.g., via monotone and convex preferences) and the specific structure of linear budget sets. This principle does not extend to other menus and utilities. As we discuss in Section 3.1 using the case of political domains with single-peaked preferences, a finite number of checks may be insufficient when menus overlap. Furthermore, even with linear budget sets, if more structured utilities are considered (e.g., expected utilities in risk domains), the linkage of marginal utilities across all possible consumption bundles and budget sets requires a comprehensive analysis of quantiles. However, even in cases where all quantiles must be analyzed, ORUMs retain computational tractability due to the ordered structure of utilities and choices. We contribute to this literature by emphasizing this important aspect of the problem and providing new characterization results.

[7]Recent papers bridging the gap between choice-based foundations and the econometric implementation of stochastic models include Dardanoni, Manzini, Mariotti, and Tyson (2020), Aguiar and Kashaev (2021), Apesteguia and Ballester (2021), and Kovach and Tserenjigmid (2022).

[8]See Kashaev, Aguiar, Plávala, and Gauthier (2023) for the application of this technique to the study of dynamic RUMs.

3. Nonparametric ORUMs

We focus on a setting involving linear, continuous, decision problems for three key reasons. First, while ubiquitous in applications, the theoretical foundations of models adopting this setting remain under-explored in the stochastic choice literature. Second, the continuous structure facilitates theoretical treatment and makes the results easier to interpret. Third, as demonstrated in the applications in Sections 6 and 7, our results apply directly to various classic economic applications involving choices in linear budget sets. Nevertheless, our analysis extends readily to other settings involving non-linear menus, as briefly discussed in Section 8, or discrete menus, as we ellaborate in Appendix B.

Let $X \subseteq \mathbb{R}^K$ be a convex space of alternatives. There is a collection $\{A_j\}_{j=1}^J$ of decision problems, or menus, that are ordered line segments of X. That is, each menu A_j consists of two corner alternatives, $\underline{x}_j$ and $\overline{x}_j$, and their convex combinations, i.e.,

$$A_j = \{(1-a)\underline{x}_j + a\overline{x}_j : \underline{x}_j, \overline{x}_j \in X \text{ and } a \in [0,1]\}.$$

Thus, any alternative $x \in A_j$ is determined by its relative position in the line segment, $a_j(x) \in [0,1]$, which is the unique value such that

$$x = (1-a_j(x))\underline{x}_j + a_j(x)\overline{x}_j.$$

Using the relative position of alternatives in menus, we assume that choice data corresponds to a collection $F = \{F_j\}_{j=1}^J$ of cumulative distribution functions (CDFs) over the interval $[0,1]$. That is, given menu A_j and value $a \in [0,1]$, the value $F_j(a)$ describes the choice mass of alternatives x in menu A_j for which $a_j(x) \leq a$. We assume that each CDF is continuous on $[0,1)$ and strictly increasing.[9]

To discuss ORUM-rationalizability, let $\mathcal{U}$ denote the class of relevant utility functions on X. Assume for now that $\mathcal{U}$ is given by all utility functions that produce a unique maximizer in each menu. Given a utility function $U \in \mathcal{U}$, let $a_j(U)$ represent the a-value of the unique alternative in menu A_j that maximizes utility according to U.[10] An ORUM has two components:

(1) Ordered-choice: The model is based on an ordered set of latent types (or simply types), represented by $\mathbb{R}$, with each type associated with a utility function such that higher types select higher alternatives. Formally, there is a *type-utility map* $\gamma : \mathbb{R} \to \mathcal{U}$ such that, for every menu A_j, the map of maximizers $a_j(\gamma(t))$ is continuous and increasing in t. In alignment with our assumption on data, we

[9] Notice that this allows for mass at any of the corner points of the menu, as commonly observed in many applications. Continuity is a standard technical assumption, and strict monotonicity is analogous to the usual positivity assumption in discrete choice models.

[10] Expressions such as $a_j(x)$ and $a_j(U)$ will be used interchangeably, though there is no risk of confusion. The former refers to the position of an alternative x in menu A_j, while the latter refers to the position of the alternative that maximizes utility U in the same menu. As a result, $a_j(x) = a_j(U)$ conveniently captures the idea that x is the maximizer of utility U in menu A_j.

work with maps γ that generate maps of maximizers that are strictly increasing over the (menu-dependent) set of interior types, i.e., those leading to $a \in (0,1)$. Consequently, for every menu A_j and every $a \in (0,1)$, there exists a unique type, denoted by $t_j^\gamma(a)$, for which $a_j(\gamma(t_j^\gamma(a))) = a$.[11] We denote the class of all such maps as Γ.

(2) Stochasticity: The type making choices is subject to randomness. Formally, there is a *type distribution* in the form of a CDF $G : \mathbb{R} \to [0,1]$. We assume that G is continuous and strictly increasing, with $\lim_{t\to-\infty} G(t) = 0$ and $\lim_{t\to\infty} G(t) = 1$. We denote the class of such CDFs as $\mathcal{G}$.

In each menu, the distribution of choices is generated by randomizing over the set of types according to G, and then maximizing according to the associated utilities given by γ. Formally, we say that data F is ORUM-rationalizable with type-utility map γ and type distribution G whenever, for every menu A_j and every $a \in (0,1)$,

$$F_j(a) = G(t_j^\gamma(a)).$$

That is, for every menu and interior alternative, the observed cumulative choice mass at that alternative coincides with the mass of types that maximize below it. Given this, the continuity assumptions guarantee that the mass observed at the corners matches the mass of types maximizing at those corners, too. ORUMs clearly represent a restriction of the more general RUMs, with the mass distributed over a unidimensional, ordered collection of utilities.

We now provide a simple characterization of data that is ORUM-rationalizable. This property can be expressed in terms of a deterministic property of quantiles. For every probability value $p \in (0,1)$ and every menu A_j, the basic assumptions on data guarantee that there is a unique alternative $c_j^p \in A_j$ such that

$$c_j^p = \underset{x:F_j(a_j(x))\geq p}{\arg\min} \; a_j(x).$$

The choice function $c^p = \{c_j^p\}_{j=1}^J$ is called the p-quantile choice function, and, following standard deterministic notions, we say that the p-quantile choice function c^p is acyclical if for every collection $(A_{j_1}, c_{j_1}^p; \ldots; A_{j_K}, c_{j_K}^p)$ such that $c_{j_{k+1}}^p \in A_{j_k}$, $k \in \{1, \ldots, K-1\}$, it holds that $c_{j_1}^p \notin A_{j_K}$. In other words, the concatenation of revealed preferences at quantile p produces no cycle. We say that F satisfies Quantile Acyclicity whenever every c^p is acyclical.

Theorem 1. *F is ORUM-rationalizable if and only if F satisfies Quantile Acyclicity.*

ORUM-rationalizability is thus equivalent to each p-quantile choice function being acyclical. The intuition for the sufficiency part is as follows: for every $p \in (0,1)$, the acyclicity of the p-quantile choice function c^p allows us to identify a utility function U^p that rationalizes c^p. Choices must increase across different levels of p due to the quantile

[11] That is, t_j^γ is just the inverse of the map $a_j \circ \gamma$, which is bijective on the set of interior alternatives.

definition of c^p. Consequently, data can be explained by randomizing uniformly over these quantile utilities. To construct the type-utility map γ and the type distribution G, we project the interval $(0,1)$ onto the real line using a bijection, and consider the corresponding induced distribution over the reals. We do this via the standard logistic transformation. Other transformations could lead to different parameterizations, and we expand on this in Section 4.

3.1. **A political economy example.** We propose a simple political economy example in order to illustrate Theorem 1, the practical testing of the characterizing property, the comparison of ORUMs with the prominent Luce model, and a discussion on the handling of substitution patterns by the model.

Suppose that $X = \mathbb{R}$, and let $\mathcal{U}$ be the class of strictly quasi-concave utility functions on X. Following the analysis of Moulin (1984), menus will be closed intervals of the real line. ORUM-rationalizability requires the consideration of a type-utility map in which types are assigned strictly quasi-concave utilities, with the family being continuous and strictly increasing in the peaks. When a type distribution is assumed, this induces a distribution of peaks, which is all that matters for choices.

An important aspect of this application, as in many other economic settings, is that the family $\mathcal{U}$ is assumed to have some minimal structural property. For example, monotonicity in consumption settings, independence in risk settings, or strict quasi-concavity in the current example. Naturally, this requires strengthening the revealed preference notion that defines Quantile Acyclicity. In the present context, strict quasi-concavity implies that, when considering the revealed preference at quantile p, we learn more than the usual c^p_j *is preferred to every* $x \in A_j$. Specifically, if $c^p_j \neq \overline{x}_j$, we infer the ranking of all alternatives greater than or equal to c^p_j, with a similar reasoning for alternatives below c^p_j whenever $c^p_j \neq \underline{x}_j$. This revealed information should be incorporated into the acyclicity analysis. With this in mind, the rest follows: the rationalizability of c^p requires the absence of cycles in the concatenation of these strengthened revealed preferences, and ORUM-rationalizability is equivalent to the corresponding version of Quantile Acyclicity.

Now, suppose that choice data is generated by a uniform distribution of peaks on $(0,1)$, with no mass outside this interval. Consider two menus, $A_1 = [0,1]$ and $A_2 = [0, \frac{1}{2}]$. Trivially, the data generated by the ORUM is uniform on A_1, with no mass at the corners. When considering menu A_2, every type with a peak above $\frac{1}{2}$ selects the alternative $\overline{x}_2 = \frac{1}{2}$, implying that a mass of $\frac{1}{2}$ is uniformly distributed in the interior, and a mass of $\frac{1}{2}$ is observed at the upper corner. To see the role of our characterizing property in this setting, note that with two menus related by inclusion, Quantile Acyclicity simply implies asymmetry, i.e., the smaller menu cannot contradict the revelations of the larger one. In particular, when $p \in (0, \frac{1}{2})$, since $c^p_1 = p \in A_2 \subseteq A_1$, we must also have $c^p_2 = p$. As a result, for menu A_2, the mass of choices in $I_1 = [0, p]$ must be p. By considering p approaching $\frac{1}{2}$, the corner mass described above is derived.

Quantile Acyclicity requires every quantile to be rationalizable. Depending on the application, it may be sufficient to consider a finite number of checks. This is the case, for example, in linear budget sets involving consumption or altruism, as we discuss in Section 7. Yet, it can be seen that this feature results from the non-overlapping nature of linear budget sets and the family of utilities at stake. The present example shows that finite checks are not sufficient in some cases. Essentially, ORUM-rationalizability requires that the revealed distribution of peaks in $(0, \frac{1}{2})$ is exactly the same for A_1 and A_2, a feature that cannot be tested with a finite number of checks on F_1 and F_2. The analysis of expected utility in Section 6 constitutes another example where a finite number of quantiles is insufficient.

We now compare ORUMs to the best-known restriction of RUMs, the Luce model. A continuous version of the Luce model producing uniform choices in A_1 with no mass at the corners requires all alternatives to be indifferent. Hence, contrary to our example, one should also observe uniform choices in A_2 with no mass at the corners. From the point of view of the characterizing property of the Luce model, consider the continuous version of IIA, where the ratio of choice probabilities between two intervals does not depend on the menu to which they belong. IIA then requires that the ratio of choice probabilities for intervals $I_1 = [0, p]$ and $I_2 = [p, \frac{1}{2}]$ must be the same in both menus. That is, for menu A_2, the ratio of mass of choices in I_1 over I_2 should also be $\frac{p}{\frac{1}{2}-p}$, and thus the mass of choices belonging to I_1 should be equal to $\frac{p}{p+\frac{1}{2}-p} = 2p$, leading to the uniform F_2 with no mass at the corners.

Importantly, note that ORUM is built on the basis of mass over ordered types, which endogenously embeds the ordered substitution patterns contained in the space of alternatives and the family of utilities under consideration. In contrast to the Luce model, in ORUMs, the mass of the removed alternatives $[\frac{1}{2}, 1]$ is not uniformly distributed among the rest but follows the substitution pattern induced by the ordered structure of the real line and the strictly quasi-concave utilities of types. When these high alternatives are removed, high types concentrate on the highest available alternative in A_2. By ignoring the ordered structure underlying the political application, the Luce model struggles to accommodate these patterns.

4. Semi-nonparametric ORUMs

Empirical work usually involves a variety of parametric assumptions, either through the adoption of a specific type-utility map or by working with a particular class of type distributions. These assumptions may, for example, facilitate computational estimation or aid in interpreting results. In this section, we briefly discuss the potential implications of restricting each of these channels separately. We first consider the case in which the dimensionality of the type distribution is reduced.[12] The next result follows directly from Theorem 1 and shows that such restrictions alone have no further empirical implications.

Theorem 2. *Let $G \in \mathcal{G}$. F is ORUM-rationalizable with type distribution G if and only if F satisfies Quantile Acyclicity.*

In other words, parametric restrictions on the type distribution alone have no empirical content per se. The reason for this is that an appropriate relabelling of utilities allows for free modification of the type distribution structure. This can be viewed as a normalization result, akin to the classical analysis by Matzkin (2003). There, the focus is on identifying a class of econometric models of the form $Y = m(X, \alpha)$, where X and Y are observables, α is a non-additive unobserved error, and m in monotone in α. By considering how the observable choice Y depends on the observable menu of alternatives X through the unobserved latent variable α and the optimization process m, it becomes clear that ORUMs belong to this class of econometric models. For our purposes, the main result in Matzkin (2003) is Lemma 1, which demonstrates that the distribution of errors can be fixed without loss of generality. This is precisely the reasoning behind why Quantile Acyclicity is also sufficient for ORUM-rationalizability, even after fixing the type distribution. Thus, Theorem 2 can be viewed as a direct combination of our Theorem 1 and Lemma 1 in Matzkin (2003).[13]

Next, we consider the case in which the type-utility map is fixed. For example, in the study of decisions under risk, let $\mathcal{U}$ represent the general class of expected utilities. An analyst interested in describing risk aversion levels in a population might fix the type-utility map such that $\gamma(t)$ corresponds to expected utility with a constant relative risk aversion coefficient equal to t.[14] This modelling assumption is likely to restrict the analyst's flexibility, since the sub-class $\gamma(\mathbb{R})$ of utilities within the image of γ has a lower dimension than $\mathcal{U}$. For instance, CRRA utilities may fail to capture across-menu choice patterns that can only be explained by other expected utility functions. However, the logic presented in Theorem 1 can be directly applied to this semi-nonparametric case by appropriately adapting the required revealed preference property on each quantile of choices to capture the rationalizability by utilities in $\gamma(\mathbb{R})$. That is, instead of applying the acyclicity condition, that is relevant when all utilities are considered, we must strengthen the deterministic concept to apply precisely to the sub-class of utilities $\gamma(\mathbb{R})$. In our abstract setting, we refer to this as Quantile $\gamma(\mathbb{R})$-Rationalizability. In

[12] Notice that Theorem 2 below fixes a given type distribution. Since this is inconsequential, the theorem extends to cases where the type distribution is assumed to belong to a particular family of distributions, such as in ordered-logit or ordered-probit models.

[13] Note that the results in this paper are not primarily concerned with identification. Instead, our goal is to provide necessary and sufficient conditions for all cases, thereby demonstrating the exact empirical content of these models.

[14] The CRRA monetary utility function with coefficient t is of the form $\frac{z^{1-t}}{1-t}$ for $t \neq 1$ and $\log z$ for $t = 1$, with $z > 0$.

Section 6 we examine the CRRA expected utility case in detail, and in Section 7 we provide another example in the context of altruism with CES utilities.

Theorem 3. *Let $\gamma \in \Gamma$, with image $\gamma(\mathbb{R})$. F is ORUM-rationalizable with type-utility map γ if and only if F satisfies Quantile $\gamma(\mathbb{R})$-Rationalizability.*

Theorem 3 shows that, modulo the restriction imposed by the assumed sub-family of utilities, the specific parametrization chosen does not result in any loss of generality. Once again, we can adapt to any specific labelling of these utilities by finding the exact type distribution that fits this labelling.

Before concluding, we note that empirical work often involves restrictions on both components of the model; that is, the analyst typically fixes a specific type-utility map and imposes some structural property on the type distribution. For example, an analyst studying risk aversion might fix the CRRA expected utility map and assume logistic variation. The combination of these two restrictions generates stronger implications, which we illustrate using the political economy example discussed earlier. This issue is discussed in greater detail in Section 5.

4.1. **A political economy example.** We illustrate the semi-nonparametric results using, again, our political economy example involving the real line and the class of strictly quasi-concave utilities. For the purposes of this section, we simplify the rationalizability question by assuming that data is observed over a single interval $A_1 = [-k, k]$. Notice that, since there is only a single menu, Quantile Acyclicity for single-peaked preferences trivially holds, and thus any data is ORUM-rationalizable. In particular, suppose that $F_1(0) = \frac{1}{3}$, $F_1(\frac{1}{2}) = \frac{1}{2}$, and $\lim_{a \to 1} F_1(a) = \frac{4}{5}$.[15] ORUM-rationalizability can be achieved with any continuous and strictly increasing type-utility map and type distribution for which the induced CDF of peaks has value $\frac{1}{3}$ at $-k$, $\frac{1}{2}$ at 0, and $\frac{4}{5}$ at k.

Theorem 2 informs us that we can account for logistic variation of the latent variable by selecting an appropriate type-utility map. For example, define $\gamma(t)$ as the utility function $-(x - f(t))^2$, where $f(t) = \frac{t}{2}$ whenever $t \geq 0$, and $f(t) = t$ whenever $t < 0$. If we then select a logistic distribution with location and scale parameters set to 0 and 1, respectively, this model rationalizes F_1.

Theorem 3 shows that fixing the type-utility map γ may require an adaptation of the quantile rationalizability concept. Suppose, for example, that we fix the type-utility map where type t is assigned the Euclidean utility $E(t)$, defined by $-(x-t)^2$. Euclidean utilities form a sub-family of strictly quasi-concave utilities, and additional revelations are possible because, for example, observing an interior choice reveals the ranking of all alternatives in $\mathbb{R}$. In such cases, we may need to define a deterministic notion of Euclidean rationalizability. However, with choices over intervals, all that matters is the

[15]For the present discussion, it suffices to consider the values of F_1 at a few points. Therefore, we will show the rationalization only at these specific points.

distribution of peaks, and thus the Euclidean restriction becomes inconsequential; the same deterministic choice behaviors can still be rationalized. In particular, our example with a single menu remains rationalizable, and to explain the data using E, we simply need to select a type distribution with CDF values at $-k$, 0, and k, as discussed above. The combination of E and the constructed type distribution rationalizes the data.

Now, consider a scenario where both the specific type-utility map E and a restriction to logistic type distributions are imposed. Under these assumptions, the data F_1 is no longer ORUM-rationalizable. Notice that, given $F_1(\frac{1}{2}) = \frac{1}{2}$, the median value of the latent variable must be 0. Since the peaks of types $-k$ and k are symmetric with respect to 0, the logistic assumption would imply that $F_1(0) = 1 - \lim_{a\to 1} F_1(a)$, which results in a contradiction. Thus, parametric assumptions on both components of the model lead to consequences beyond those described in Theorem 3, which will be further analyzed in the next section.

5. Parametric ordered-logit model

As illustrated by our previous example, when parametric restrictions are imposed on both the type-utility map and the type distribution, the model becomes restricted in ways that cannot be easily captured by Theorem 1 or the adjustment proposed in Theorem 3. It is therefore evident that the foundations of parametric models often used in empirical work, such as ordered-logit models, require more sophisticated properties that restrict data across quantiles. In this section, we take an initial step toward describing the exact empirical content of such models. Specifically, we study parametric ordered-logit models, a popular tool in economics as well as other disciplines such as political science, sociology, and biology.[16]

To simplify the analysis, we assume the following richness condition: for any two menus A_j and $A_{j'}$, there exists a sequence of menus $A_j = A_{j_0}, \ldots, A_{j_k}, \ldots, A_{j_K} = A_{j'}$ such that, for each $k \in \{0, \ldots, K-1\}$, there exists an interval of types that produce interior maximizers in both A_{j_k} and $A_{j_{k+1}}$. To motivate this assumption, note that if at least one of the corner alternatives has no mass, the richness assumption is trivially satisfied.[17] However, if choices at both corners are relevant, the richness condition only requires that the intervals of utilities generating interior maximizers overlap weakly, after connecting a chain of menus. For example, in a consumption setting, any pair of menus should be connected by other menus (likely involving intermediate prices) to produce the desired overlap.

We now study the parametric ordered-logit model, in which a generic type-utility map $\gamma : \mathbb{R} \to \mathcal{U}$ is fixed, and the type distribution is restricted to the logistic family,

[16]For references to the diversity of economic applications of the ordered-logit model, see the discussion in footnote 1 of the Introduction.

[17]This feature is present in the two applications discussed in Sections 6 and 7.

denoted $\mathcal{G}^L$. Recall that a logistic type distribution takes the form

$$G^{(\tau,\sigma)}(t) = \frac{1}{1 + e^{-(t-\tau)/\sigma}},$$

where $\tau \in \mathbb{R}$ and $\sigma > 0$ are the location and scale parameters.

We now present the key properties of this ordered-logit model, which capture the requirements on corner and interior alternatives that arise when a logistic distribution is applied to the given type-utility map γ. The requirement on the corner alternatives follows from the fundamental properties of Random Utility Models (RUMs). In this continuous ordered setting, the existence of a single type leading to a corner implies the existence of an unbounded interval of types with the same property, resulting in a strictly positive mass for this corner.[18]

Corner Extremeness (CE). $F_j(0) > 0$ (respectively, $\lim_{a \to 1} F_j(a) < 1$) if and only if there exists $t \in \mathbb{R}$ such that $a_j(\gamma(t)) = 0$ (respectively, $a_j(\gamma(t)) = 1$).

To formulate our second property, consider the cumulative log-odds of any interior value $a \in (0, 1)$, defined as

$$\ell_j(a) = \log\left(\frac{F_j(a)}{1 - F_j(a)}\right).$$[19]

It is well known that when data is generated by a logistic distribution with parameters (τ, σ), the cumulative log-odds $\ell_j(a)$ correspond to the standardized type with a as the maximizer, i.e., $\ell_j(a) = \frac{t_j^\gamma(a)-\tau}{\sigma}$. Now, consider two menus and a pair of non-corner alternatives from each menu. If the sum of the types associated with the first pair of alternatives equals the sum of the types associated with the second pair, then the sum of the cumulative log-odds of both pairs should also coincide.[20]

Cumulative Log-odds Additivity (CLA). Let $a, b, a', b' \in (0, 1)$ and $A_j, A_{j'}$ be such that $t_j^\gamma(a) + t_j^\gamma(b) = t_{j'}^\gamma(a') + t_{j'}^\gamma(b')$. Then, $\ell_j(a) + \ell_j(b) = \ell_{j'}(a') + \ell_{j'}(b')$.

Note that when $a = b$ and $a' = b'$, CLA implies that the same type must have the same associated log-odds, and hence the same quantile, in every menu. This implies that quantiles are rationalized by utilities in $\gamma(\mathbb{R})$. Importantly, CLA not only imposes this requirement but also constrains the sum of log-odds for pairs of alternatives.

Theorem 4 shows that these two basic properties, CE and CLA, are not only necessary but also sufficient for data to be ORUM-rationalizable with type-utility map γ and a type distribution in $\mathcal{G}^L$.

Theorem 4. *Let γ be any type-utility map. The data F is ORUM-rationalizable with type-utility map γ and a type distribution in $\mathcal{G}^L$ if and only if F satisfies CE and CLA.*

[18] This is akin to the extremeness property described by Gul and Pesendorfer (2006).

[19] This is also called the logit function, which is the inverse of the logistic distribution.

[20] Since γ is fixed, the mapping from types to alternatives is known, and CLA can thus be verified.

The proof of the sufficiency part of Theorem 4 involves several steps. First, when considering the interior alternatives in a given menu, the data immediately induces a specific CDF over the corresponding interval of types. When relevant, we account for the censoring produced by corner choices, which are optimal for an unbounded interval of types. The masses observed at the corners must be appropriately distributed among all rationalizing types, guaranteed to exist by CE, in such a way that the constructed CDF over all types satisfies the additivity requirement imposed by CLA for interior alternatives. We address this using a recursive construction. Second, the ordered-logit functional form requires us to build upon Theorem 2.1.5 of Galambos and Kotz (1978). This classical statistical result provides a necessary and sufficient condition for a single CDF over the real line, that has been assumed to be symmetric about the origin, to be logistic. We extend this result to our revealed preference setting, where (i) distributions may have arbitrary means and are not proven to be symmetric, and (ii) we have a collection of menu-dependent distributions, not yet shown to coincide. Our CLA property proves sufficient to demonstrate that these menu-dependent distributions are all logistic and share the same location and scale parameters. Finally, it is important to note that the parameters (τ, σ) of the logistic type distribution that rationalize the data must be unique.

5.1. **A political economy example.** The properties of CE and CLA may seem abstract, but this is primarily because the type-utility map γ has been presented abstractly. Let us return to the parameterized ordered-logit model using the Euclidean map E. In Section 4.1, we learned that this model restricts data, even with a single menu.[21] With respect to CE, notice that for any menu $A_j = [\underline{x}_j, \overline{x}_j]$, the corner alternatives are always the maximizers for some type (specifically, for types with peaks below $\underline{x}_j$ and above $\overline{x}_j$, respectively). Consequently, CE reads as (for any menu):

$$[\text{CE}_{PE}] \quad F_j(0) > 0 \quad \text{and} \quad \lim_{a \to 1} F_j(a) < 1.$$

Regarding CLA, notice that with Euclidean preferences, the utility maximized at x is the one with a peak at x, and hence $t_j^E(a_j(x)) = x$. Therefore, CLA reads as follows (for any menus and interior alternatives):

$$[\text{CLA}_{PE}] \quad x + y = x' + y' \Rightarrow \ell_j(a_j(x)) + \ell_j(a_j(y)) = \ell_{j'}(a_{j'}(x')) + \ell_{j'}(a_{j'}(y')).$$

Note that our analysis can be applied to the study of other type distributions.[22] When considering any other family of continuous and strictly increasing distributions, CE would still apply, while the CLA requirement should be modified to capture the structural implications of the corresponding statistical distribution. For instance, if we

[21]Indeed, we showed how data should be symmetric around the alternative occupying the median quantile, a property implied by the analysis here.

[22]An alternative approach is to consider mixtures of the logistic distribution. This can be seen as a parallel exercise to the mixed-logit analysis. In Appendix C, we show that any ORUM with a given type-utility map γ can be approximated by a sequence of mixed ordered logits using the same γ.

consider a Gaussian latent variable, the standardized type $\frac{t_j^\gamma(a)-\tau}{\sigma}$ would now take the form $\Phi_j^{-1}(a)$, where Φ_j^{-1} denotes the inverse of the CDF of the standard normal. Setting aside the lack of a closed-form solution for this expression, one could operationalize a version of the CLA property by simply replacing $\ell_j(a)$ with $\Phi_j^{-1}(a)$.

Moreover, the structure of the CLA property also helps in deciding which family of type distributions to work with. An analyst can start by fixing the type-utility map γ, understanding the implications of this assumption as outlined in Theorem 3. Then, if a parametric exercise is performed, an analysis of the changes in the sums of log-odds, Φ^{-1}, or the corresponding expression given by the distribution at stake, may guide the selection of the most appropriate type distribution.

6. Application: Risk

An advantage of our results is that they are portable to the analysis of specific economic settings of interest. As illustrated in Section 5.1 with the political economy example, when a specific setting is adopted, all properties can be expressed in terms of the appropriate fundamentals and take on a form that is familiar to the expert, allowing for direct empirical examination. We now showcase the applicability of our framework using two basic settings involving risk and altruism in standard linear budget sets. In each case, we first particularize Theorems 1 to 4, providing nonparametric, semi-nonparametric, and parametric characterizations. We then give detailed guidelines on how to empirically apply our results. We start with the setting of risk; Section 7 covers the study of altruism.

Let $X = [0,1]\times\mathbb{R}_+^2$ represent the set of all possible (two) state-contingent lotteries.[23] A menu is a collection of lotteries in which all have the same state probabilities, and their payouts belong to a linear budget set. Formally, wealth is equal to 1, and let $\pi_j^i > 0$ denote the price of allocating money to state i in menu A_j. Then, a menu is the collection of all affordable lotteries of the form $(q_j^1; x^1, x^2)$, where $q_j^1 \in [0,1]$ describes the menu-dependent probability of state 1 (hence, $q_j^2 = 1 - q_j^1$ describes the probability of state 2), and $(x^1, x^2) \in \mathbb{R}_+^2$ are the possible payout combinations. With the usual budget set notation, the affordable lotteries are $B_j = \{(q_j^1; x^1, x^2) \in X : \pi_j^1 x^1 + \pi_j^2 x^2 \leq 1\}$.[24] We assume, without loss of generality, that state 1 pays more in expectation, i.e., $\phi_j^1 \equiv \frac{\pi_j^1}{q_j^1} < \frac{\pi_j^2}{q_j^2} \equiv \phi_j^2$, and denote $\phi_j = \frac{\phi_j^2}{\phi_j^1}$. Given the assumptions on utility laid down below, potential maximizers belong to the line segment A_j between the corner

[23]We discuss the case of two states because ORUMs in general, and ordered-logit models in particular, are uni-dimensional in nature; they create one-to-one maps from a latent variable to the space of choices. We briefly address this point in Section 8. Moreover, as shown in Apesteguia, Ballester, and Gutierrez-Daza (2024), the lessons learned in the uni-dimensional analysis can be transferred via conditional analysis to the study of multi-dimensional settings.

[24]Given a menu, the proper decision variables involve x^1 and x^2, but we need to keep in the notation the menu constant q_j^1 to fully describe the lottery.

allocation $\underline{x}_j = (q_j^1; \frac{1}{\pi_j^1}, 0)$ and the equal-payout allocation $\overline{x}_j = (q_j^1; \frac{1}{\pi_j^1+\pi_j^2}, \frac{1}{\pi_j^1+\pi_j^2})$. Note that larger levels of risk aversion are intuitively reflected in larger investments in state 2. For this application, it is more convenient to represent each alternative $x \in A_j$ by its unique associated ratio $r_j(x) \in [0,1]$, given by $r_j(x) = \frac{x^2}{x^1}$, rather than by the value $a_j(x)$ used above.[25]

6.1. **Theoretical Results.** We begin with the nonparametric and semi-nonparametric analysis. Given the risk setting, we first focus on the class $\mathcal{U}$ of expected utilities associated with continuous, strictly increasing, and concave monetary utility functions. To illustrate the semi-nonparametric case, we fix the type-utility map $CRRA$, where type t corresponds to the expected utility with a relative risk aversion coefficient equal to t.[26]

The nonparametric study of ORUM-rationalizability requires expected utility rationalizability at the quantile level. We can borrow from the deterministic analysis of expected utility rationalizability by Kubler, Selden, and Wei (2014). Following their paper, we say that the quantile choice function c^p formed by the collection $\{c_j^p = (q_j^1; c_j^{p,1}, c_j^{p,2})\}_{j=1}^J$ satisfies the Strong Axiom of Revealed Expected Utility (SAREU) whenever, for every sequence of menus $A_{j_1}, \ldots, A_{j_K}$, it holds that $L(j_1, j_2) \cdot L(j_2, j_3) \cdot \ldots \cdot L(j_{K-1}, j_K) \cdot L(j_K, j_1) < 1$, where[27]

$$L(j,j') = \begin{cases} 0 & \text{whenever } c_{j'}^{p,s'} > c_j^{p,s} \text{ for all } s, s' \in \{1,2\}, \\ \max\limits_{s,s': c_j^{p,s} > c_{j'}^{p,s'}} \dfrac{\phi_j^s}{\phi_{j'}^{s'}} & \text{otherwise.} \end{cases}$$

The simplest violation of SAREU involves a single menu A_j such that $L(j,j) > 1$. This corresponds to $c_j^{p,2} > c_j^{p,1}$, i.e., a violation of first-order stochastic dominance (since, as discussed above, larger investments in the state paying more in expectation should be unequivocally observed). SAREU eliminates these and more complex weighted cycles, an issue that will be discussed in detail in the empirical part below. We say that F satisfies Quantile SAREU whenever every c^p satisfies SAREU.

Next, consider the semi-nonparametric model. From the standard first-order condition analysis, it is well-known that the relative risk aversion coefficient of the type maximizing at x in menu A_j is $-\frac{\log \phi_j}{\log r_j(x)}$. Thus, for a choice function to be rationalizable by a CRRA expected utility, this expression should be, for any given quantile, constant across menus. Formally, we say that the quantile choice function c^p satisfies

[25]Notice that the definition of quantiles is not affected by this ordinal transformation. This is merely more convenient for the exposition of the parametric case.

[26]Recall that, as shown in Theorem 2, the semi-nonparametric model imposing assumptions on the type distribution is equivalent to the nonparametric model, hence we omit its discussion here and in Section 7.

[27]As in Kubler, Selden, and Wei (2014), we assume that $c_j^{p,s} \neq c_{j'}^{p,s'}$ for all j, j' and s, s'.

the Strong Axiom of Revealed CRRA Expected Utility (SARCEU) whenever, for every pair of menus A_j and $A_{j'}$, it holds that

$$\frac{\log \phi_j}{\log r_j(c_j^p)} = \frac{\log \phi_{j'}}{\log r_{j'}(c_{j'}^p)}.$$[28]

We say that F satisfies Quantile SARCEU whenever every c^p satisfies SARCEU.

Corollary 1. *In the risk domain:*

(1) *F is ORUM-rationalizable if and only if F satisfies Quantile SAREU.*
(2) *F is ORUM-rationalizable with type-utility map CRRA if and only if F satisfies Quantile SARCEU.*

We now proceed to study the parametric case. We consider again the parametrization CRRA, and require the type distribution to belong to the class $\mathcal{G}^L$ of logistic distributions. We apply Theorem 4 to analyze the properties in this setting.

First, notice that the corner allocation $\underline{x}_j = (q_j^1; \frac{1}{\pi_j^1}, 0)$ corresponds to the menu-independent interval of types $(-\infty, 0]$, and must have strictly positive mass. Meanwhile, the corner allocation $\overline{x}_j = (q_j^1; \frac{1}{\pi_j^1+\pi_j^2}, \frac{1}{\pi_j^1+\pi_j^2})$ is sub-optimal for all utilities in the CRRA class and must have zero mass. Therefore, CE reads as follows, for every menu:

$$[\text{CE}_R] \quad F_j(0) > 0 \quad \text{and} \quad \lim_{r\to 1} F_j(r) = 1.$$

Since the type maximizing at $r \in (0,1)$ is given by $-\frac{\log \phi}{\log r}$, the CLA condition for any pair of menus and interior ratios is:

$$[\text{CLA}_R] \quad \log \phi_j \left(\frac{1}{\log r} + \frac{1}{\log s}\right) = \log \phi_{j'} \left(\frac{1}{\log r'} + \frac{1}{\log s'}\right) \Rightarrow \ell_j(r)+\ell_j(s) = \ell_{j'}(r')+\ell_{j'}(s').$$

Corollary 2 follows directly from these conditions:

Corollary 2. *In the risk domain, F is ORUM-rationalizable with type-utility map CRRA and a type distribution in $\mathcal{G}^L$ if and only if F satisfies CE_R and CLA_R.*

In relation to the comparison between the parametric and semi-nonparametric properties, note that the quantile version of SARCEU requires $\frac{\log \phi_j}{\log r} = \frac{\log \phi_{j'}}{\log r'} \Leftrightarrow \ell_j(r) = \ell_{j'}(r')$. This follows readily from CLA_R by considering the case in which $r = s$ and $r' = s'$. CLA_R strengthens SARCEU by requiring this principle to apply across pairs of quantiles.

6.2. **Empirical Analysis.** We now give a detailed guideline, comprising ten steps, on how to apply the results in the previous section to a dataset with linear budget sets. We illustrate the exercise with a simulated dataset as follows. First, we generate $J = 20$ menus with prices and probabilities randomly selected from the uniform distribution in the ranges $[1/100, 1/10]$ and $[1/5, 4/5]$, respectively. Second, finite choice data is

[28] When the choice is such that $c_j^{p,2} = 0$, the expression should be read as 0.

obtained by simulating, for each menu, N independent choices from the parametric ORUM with CRRA expected utilities and logistic distribution given by $\tau = .6$ and $\sigma = .3$.[29] We study the effect of sampling size by considering $N = \{20, 80, 160, 480\}$.

The first two steps are on the organization of the dataset in accordance with our framework.

Step 1. We start by organizing the menus such that state 1 is the one that pays more in expectation; that is, such that $\phi^1_j \equiv \frac{\pi^1_j}{q^1_j} < \frac{\pi^2_j}{q^2_j} \equiv \phi^2_j$. Table 11 in Appendix D reports the ordered menus in our dataset. ◇

Step 2. Since the analysis is based on the observed quantiles, we compute the empirical distribution functions; for every menu A_j we order the N choices by increasing consumption ratio r_j. Figure 1 reports the empirical CDF of a sample of ten randomly selected menus from our dataset of 20 menus. ◇

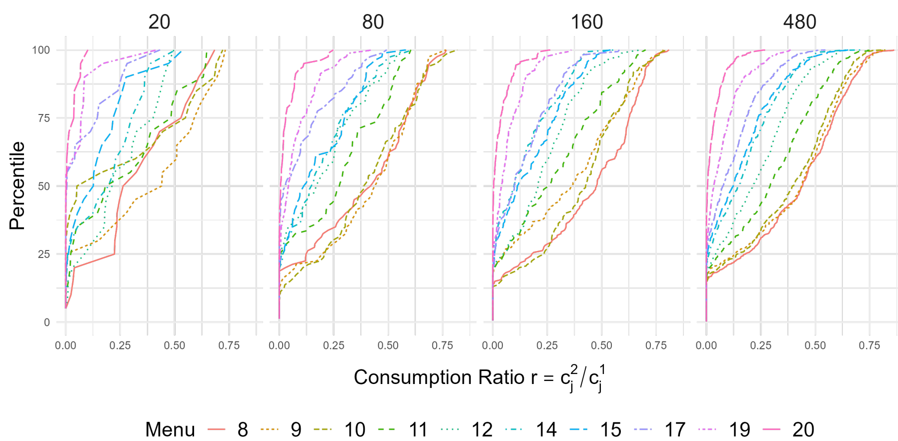


FIGURE 1. Empirical CDF of r_j across menus

The next two steps deal with the nonparametric analysis. This requires studying whether the quantile choice functions c^p, with $p \in \{\frac{1}{N}, \dots, \frac{N}{N}\}$, satisfy SAREU. To facilitate the computational exercise, we start by establishing an immediate connection between SAREU and the well-known problem of existence of negative cycles in a weighted directed graph.

Consider the complete directed graph where each node corresponds to one menu, and the weight of the directed edge from j to j' is

$$w(j, j') = \begin{cases} -\log L(j, j') & \text{whenever } L(j, j') > 0, \\ \infty & \text{otherwise.} \end{cases}$$

It is immediate that the logarithmic transformation and the change of sign guarantee that a violation of SAREU is equivalent to the existence of a negative cycle in the

[29] These parameters give choice patterns (e.g., corner choices) that are in line with empirical observations in budget set experiments (see Choi, Kariv, Müller, and Silverman (2014)).

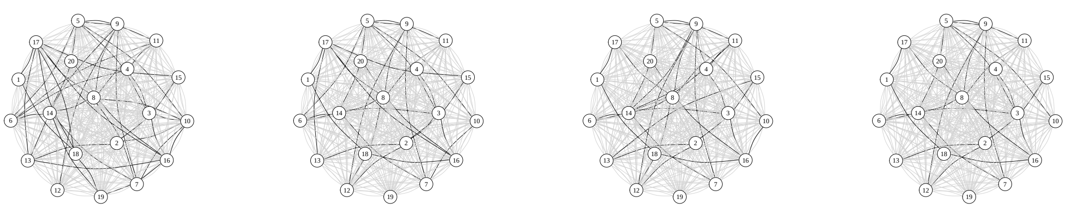
graph, i.e., a directed cycle where the sum of weights is negative. Now, in the context of a deterministic choice function c_j^p, notice that $w(j, j')$ is determined exclusively by the ranking of consumption choices $c_j^{p,1}, c_j^{p,2}, c_{j'}^{p,1}$ and $c_{j'}^{p,2}$, and that, given that $\phi_j^1 < \phi_j^2$, there are only six such possible rankings, as described in Table 1.

Case	$w(j, j')$	$w(j', j)$
$c_j^{p,1} > c_j^{p,2} > c_{j'}^{p,1} > c_{j'}^{p,2}$	$\log \phi_{j'}^1 - \log \phi_j^2 < 0$	∞
$c_j^{p,1} > c_{j'}^{p,1} > c_j^{p,2} > c_{j'}^{p,2}$	$\max\{\log \phi_{j'}^1 - \log \phi_j^1, \log \phi_{j'}^2 - \log \phi_j^2\}$	$\log \phi_j^2 - \log \phi_{j'}^1 > 0$
$c_j^{p,1} > c_{j'}^{p,1} > c_{j'}^{p,2} > c_j^{p,2}$	$\log \phi_{j'}^1 - \log \phi_j^1$	$\log \phi_j^2 - \log \phi_{j'}^2$
$c_{j'}^{p,1} > c_j^{p,1} > c_j^{p,2} > c_{j'}^{p,2}$	$\log \phi_{j'}^2 - \log \phi_j^2$	$\log \phi_j^1 - \log \phi_{j'}^1$
$c_{j'}^{p,1} > c_j^{p,1} > c_{j'}^{p,2} > c_j^{p,2}$	$\log \phi_{j'}^2 - \log \phi_j^{p,1} > 0$	$\max\{\log \phi_j^1 - \log \phi_{j'}^1, \log \phi_j^2 - \log \phi_{j'}^2\}$
$c_{j'}^{p,1} > c_{j'}^{p,2} > c_j^{p,1} > c_j^{p,2}$	∞	$\log \phi_j^1 - \log \phi_{j'}^2 < 0$

TABLE 1. Weights in the graph for menus A_j and $A_{j'}$

Hence, given J menus, the directed graph associated to the deterministic case, with a single choice per menu, must be one of at most $6^{\frac{J^2}{2}}$ types.[30]

Step 3. For every quantile, we use Table 1 to construct the corresponding weighted directed graph. Figure 2 illustrates this graph for the median quantile; it plots a snapshot of the graph without the weights, with black edges denoting negative weights. ◇



20 obs 80 obs 160 obs 480 obs

FIGURE 2. Directed graph using Table 1 for the median quantile and all menus

Step 4. In order to search for negative cycles, one can apply well-known algorithms such as the Bellman-Ford algorithm.[31] Table 2 presents the rationalizable and non-rationalizable quantiles based on the same set of 20 quantiles. As expected, one can immediately appreciate that the finer the sampling, the lower is the number of quantiles that are not rationalizable. ◇

The next two steps study the semi-nonparametric case. Quantile SARCEU imposes that for every pair of menus the expression $\frac{\log \phi_j}{\log r_j(c_j^p)}$ should be equal across all quantiles.

[30]This may be particularly useful from a computational point of view in cases where different sub-populations or datasets are available. Notice that, it is feasible to determine which types of datasets are consistent with SAREU without the need of actual data and then verify whether actual data matches this structure.

[31]See, e.g., Bang-Jensen and Gutin (2008) for a textbook treatment.

Dataset	Number of violations	Non-rationalizable quantiles
20 obs	17	5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 95, 100
80 obs	11	10, 15, 20, 25, 30, 35, 40, 45, 50, 95, 100
160 obs	8	10, 15, 20, 25, 30, 35, 90, 95
480 obs	6	10, 15, 20, 25, 30, 35

TABLE 2. Non-rationalizable quantiles across datasets

Recall that this expression is the negative of the relative risk aversion coefficient induced by the corresponding consumption.

Step 5. For every menu we compute the empirical distribution of the induced risk coefficients. Figures 3 plots the CDFs for 10 randomly selected menus. ◇

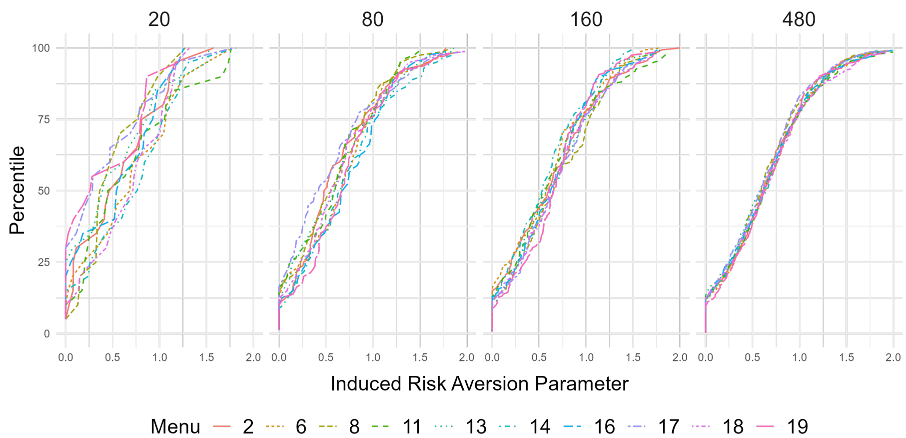


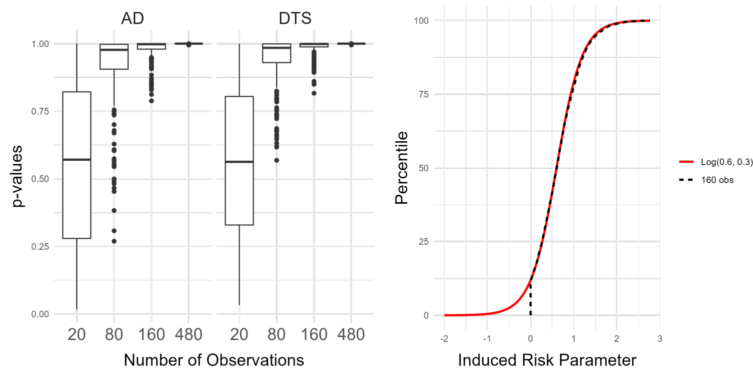
FIGURE 3. Empirical CDF of induced risk aversion parameters

Step 6. We test for the equality of distributions using the Anderson-Darling and the DTS tests. Figure 4a reports the distribution of p-values resulting from testing the equality of distributions across all pairs of menus, comparing these on the basis of the same 20 quantiles. In addition, we merge all observations and plot the empirical distribution of induced risk aversion coefficients. Figure 4b reports it, for the case of $N = 160$, together with the CDF of the underlying logistic. [32] ◇

The next four steps are devoted to the study the parametric model. We start with a discussion on properties CE_R and CLA_R.

Step 7. Property CE_R requires some mass at $r = 0$ and no mass at $r = 1$. Table 12 in Appendix D reports the observed mass at the corner of each one of the menus, which can all be seen as being strictly positive. Table 13 in Appendix D reports the largest value of r per menu. Most of these values are relatively large, with some exceptions that, intuitively, correspond to those menus with very high relative price of the second state. ◇

[32]Notice that a simple estimation of the logistic may be done by minimizing the distance to this empirical distribution function.



(A) (B)

Figure 4. Boxplot of distribution of p-values across each pair of menus (left); and aggregated empirical CDF of induced risk aversion coefficients for $N = 160$ (right)

Step 8. We now concentrate on the main property of the parametric model, CLA_R. Consider all triples formed by one menu A_j and two interior observations in this menu.[33] Take the induced risk aversion coefficients corresponding to these two observations, as calculated in Step 5, and compute its sum. Take the empirical quantiles of these two observations, as calculated in Step 2, transform them into log-odds and compute its sum. CLA_R requires the relationship between these two sums to be a linear function. Figure 5 plots this relationship for all pairs form by combining 5 randomly selected interior choices per menu.

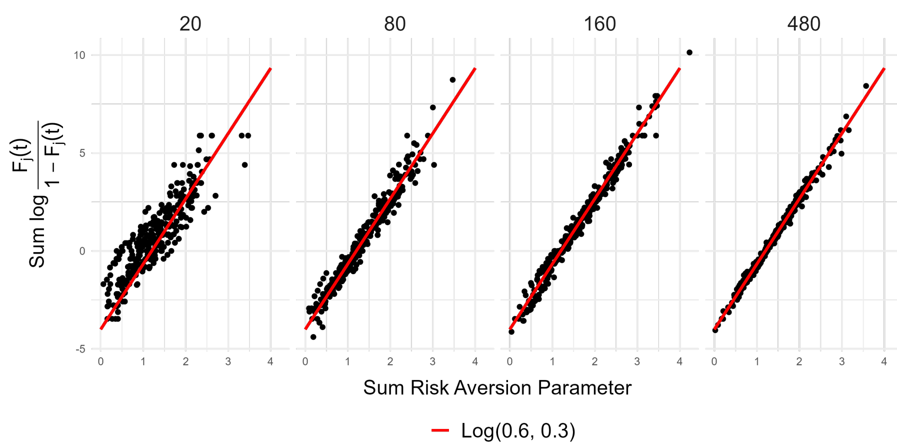


Figure 5. CLA property in risk application

Note, in addition, that these plots can be used in a simple estimation of the parameters. Linearly regressing the data plotted in Figure 5, the intersect of the sum of log odds with zero corresponds to twice the location, while the slope corresponds to the inverse of the scale of the logistic. Table 3 reports on this. ◇

[33]Given the finiteness, we also need to exclude the largest observation in each menu.

Parameter	20 obs	80 obs	160 obs	480 obs	True
τ	.503	.595	.611	.603	.6
σ	.382	.305	.300	.301	.3

Table 3. CLA parameter estimation via linear regression

The final two steps study the testing of the parametric model. In order to do so, we need to estimate the parameters of the model. We do so here by following standard methods.[34] Notice that we have a total of $M = J \times N$ observations, with $1 \leq m \leq M$ denoting a generic observation. We divide all the observations into two classes; $\{1, \dots, M_0\}$ are observations for which their associated consumption ratio is 0, and $\{M_0{+}1, \dots, N\}$ are observations such that the ratio is strictly positive. Notice that every corner observation must be the result of a type $t \leq 0$, i.e. of a negative relative risk aversion coefficient. Recall that for every interior observation $m \in \{M_0{+}1, \dots, M\}$ there is a unique induced type $t_m > 0$ that rationalizes this choice. Given the logistic parameters τ and σ, we denote the normalized values $\bar{0} = \frac{0-\tau}{\sigma}$ and $\bar{t}_m = \frac{t_m-\tau}{\sigma}$.

The log-likelihood can be written as:[35]

$$\sum_{1}^{M_0} \log G^{(\tau,\sigma)}(0) + \sum_{M_0+1}^{M} \log g^{(\tau,\sigma)}(t_m) =$$

$$-M_0 \log(1+e^{-\bar{0}}) - \sum_{m=M_0+1}^{M} \Big(\bar{t}_m + \log\sigma + 2\log(1+e^{-\bar{t}_m})\Big).$$

The first-order conditions with respect to τ and σ are:

$$\frac{1}{M}\Big[M_0 \frac{1+\frac{e^{-\bar{0}}}{1+e^{-\bar{0}}}}{2} + \sum_{m=M_0+1}^{M} \frac{e^{-\bar{t}_m}}{1+e^{-\bar{t}_m}}\Big] = \frac{1}{2},$$

$$\frac{1}{M}\Big[M_0(1-\bar{0}\frac{e^{-\bar{0}}}{1+e^{-\bar{0}}}) + \sum_{m=M_0+1}^{M} \bar{t}_m \frac{1-e^{-\bar{t}_m}}{1+e^{-\bar{t}_m}}\Big] = 1.$$

We now are in a position to present our statistical test. We can do so using a version, that accounts for type-I censoring (recall that all observations with $t \leq 0$ are recorded as 0), of the well-known Kolmogorov-Smirnov test for the logistic distribution using the estimated parameters:[36]

$$KS = \sqrt{M}\max\Big\{\Big|\frac{M_0}{M} - \frac{1}{1+e^{-\hat{\bar{0}}}}\Big|, \max_{M_0<m^*\leq M}\Big|\frac{m^*}{M} - \frac{1}{1+e^{-\hat{\bar{t}}_{m^*}}}\Big|\Big\} + \frac{0.19}{\sqrt{M}},$$

[34] Alternatively, we could use the simple method proposed in Step 8.

[35] $g^{(\tau,\sigma)}$ denotes the density of $G^{(\tau,\sigma)}$.

[36] See the standard procedures in Chapter 4 in D'Agostino and Stephens (1986). Alternatively, one could also use the Cramer-von Mises test.

where m^* is the counter referring to the interior observations after having been re-ordered to be increasing in their corresponding induced types, and $\hat{\hat{0}}$ and $\hat{\hat{t}}_{m^*}$ are the normalized values using the estimated parameters $(\hat{\tau}, \hat{\sigma})$.[37]

Step 9. The parametric estimation can be done by maximizing the above log-likelihood function or, using the first order conditions, or the method outlined in Step 8. Table 4 reports the results, conditional on the number of observations. ◇

Parameter	**20 obs**	**80 obs**	**160 obs**	**480 obs**	**True**
τ	.567	.605	.606	.606	.6
σ	.323	.315	.308	.305	.3

TABLE 4. Parameter estimation via Maximum Likelihood

Step 10. The application of the tests to our dataset is summarized in Table 5. ◇

N **obs**	Rounded share of non-censored data	Test Statistic KS	p-value KS	Test Statistic CVM	p-value CVM
20 obs	.9	.596	$> .5$	.068	$> .5$
80 obs	.9	.486	$> .5$	.041	$> .5$
160 obs	.9	.484	$> .5$	.031	$> .5$
480 obs	.9	.430	$> .5$	.036	$> .5$

TABLE 5. KS and CVM tests for logistic distribution with left random censoring

7. Application: Altruism

Let $X = \mathbb{R}^2_+$ represent the set of all possible monetary allocations in which the first component refers to the payment to oneself and the second refers to the payment to another person. Individuals confront linear budget sets of the form $B_j = \{x = (x^1, x^2) \in X \mid \pi^1_j x^1 + \pi^2_j x^2 \leq 1\}$, where $\pi^i_j > 0$ denotes the cost of allocating money to person i in budget set B_j. Below we assume monotonicity in preferences, leading choices to belong to the line segment, denoted A_j, between the corner allocations $\underline{x}_j = (\frac{1}{\pi^1_j}, 0)$ and $\overline{x}_j = (0, \frac{1}{\pi^2_j})$. Note that larger levels of altruism are reflected in larger transfers to individual 2. In this application, it is convenient to use the following parametrization of the allocations in a menu A_j: $v_j(x) \in [0,1]$ given by $v_j(x) = 1 - e^{-r_j(x)}$.[38] Also, denote $\pi_j = \frac{\pi^2_j}{\pi^1_j}$.

[37] Note that the ordering in terms of the induced types is equivalent to that of the normalized types.

[38] Where we set $v_j(x) = 1$ when $x^1 = 0$.

7.1. **Theoretical results.** We consider first the nonparametric analysis. In line with standard models of social preferences, we start by considering the class $\mathcal{U}$ of utilities that are continuous and either: (i) strictly monotone in the first component (and either constant or decreasing in the second component), or (ii) strictly increasing in both components and strictly convex. The characterization of this case will be based on the classical Weak Axiom of Revealed Preference (WARP). In our setting, the quantile choice function c^p satisfies WARP whenever for every pair of menus $A_j, A_{j'}$ such that $c^p_j \in B_{j'}$ and $c^p_{j'} \in B_j$, it is $c^p_j = c^p_{j'}$. As it is well-known, this is basically the asymmetry part of the Acyclicity property of Section 3, together with monotonicity. We say that F satisfies Quantile WARP whenever every c^p satisfies WARP.

In the semi-nonparametric case, we consider CES-type utility functions $(x^1)^\alpha + t(x^2)^\alpha$, where $\alpha \in (0,1)$ is a constant, and allow for variation in the altruism coefficient $t \in \mathbb{R}$. The study of this case follows by applying Theorem 3 in combination with the following deterministic analysis. From the standard first-order condition, the altruism coefficient is equal to $\pi_j r_j(x)^{1-\alpha}$, where recall that $r_j(x) = \frac{x^2}{x^1}$. For a given quantile, the altruism coefficient should be equal across pairs of menus. Hence, for any given quantile p, this allows to obtain α from two menus as $\frac{\log(\pi_j/\pi_{j'})}{\log(r_j(c^p_j)/r_{j'}(c^p_{j'}))} + 1$. We then say that the collection of choices c^p satisfies the Weak Axiom of Revealed CES Preference (WARCESP) whenever for every three menus $A_j, A_{j'}, A_{j''}$, either $c^{p,2}_j = 0$ always holds or

$$0 < \frac{\log(\pi_j/\pi_{j'})}{\log(r_j(c^p_j)/r_{j'}(c^p_{j'}))} + 1 = \frac{\log(\pi_j/\pi_{j''})}{\log(r_j(c^p_j)/r_{j''}(c^p_{j''}))} + 1 < 1.$$

We say that F satisfies Quantile WARCESP whenever every c^p satisfies WARCESP for the same constant.

Corollary 3. *In the altruism domain,*

(1) *F is ORUM-rationalizable if and only if F satisfies Quantile WARP.*
(2) *F is ORUM-rationalizable with type-utility map CES if and only if F satisfies Quantile WARCESP.*

Corollary 3 is similar in spirit to Corollary 1. The nonparametric case is obtained with a relatively classical and broad set of preferences, while the semi-nonparametric case is derived by using the CES family of utility functions. Importantly, we want to focus our attention on the altruism parameter t, and not on the curvature α, so the second part of Corollary 3 requires not only the assumption that every quantile is CES-rationalizable but also that all quantiles are CES-rationalized with the same curvature.

We now proceed to study the parametric case by using the CES map for a given value of α, and imposing the logistic distribution over the type space t (or equivalently, over the altruism parameter). The study of rationalizability is again a direct consequence of the application of Theorem 4. To understand how property CE reads in this setting, notice again that for any line segment A_j, the corner allocation $\underline{x}_j = (\frac{1}{\pi_j^1}, 0)$ is associated with a constant interval of CES utilities ($t \leq 0$) while the corner allocation $\overline{x}_j = (0, \frac{1}{\pi_j^2})$ is never selected. Then, for any menu:

$$[\text{CE}_A] \quad F_j(0) > 0 \text{ and } \lim_{v \to 1} F_j(v) = 1.$$

To see how property CLA reads, notice that the type that maximizes at the interior alternative x is the one with altruism coefficient equal to $\pi_j r_j(x)^{1-\alpha}$. Then, for any pair of menus and interior alternatives:

$$[\text{CLA}_A] \quad \pi_j \left(\left[\log \frac{1}{1-v} \right]^{1-\alpha} + \left[\log \frac{1}{1-w} \right]^{1-\alpha} \right) = \pi_{j'} \left(\left[\log \frac{1}{1-v'} \right]^{1-\alpha} + \left[\log \frac{1}{1-w'} \right]^{1-\alpha} \right) \Rightarrow$$
$$\ell_j(v) + \ell_j(w) = \ell_{j'}(v') + \ell_{j'}(w').$$

We omit the proof of the following immediate result.

Corollary 4. *In the altruism domain, F is ORUM-rationalizable with type-utility map CES and a type distribution in $\mathcal{G}^L$ if and only if F satisfies CE_A and CLA_A.*

7.2. **Empirical Analysis.** We now give a detailed guideline on how to apply the results in the previous section to a simulated dataset. This guideline follows closely the one given for the risk case. We randomly generate $J = 20$ menus with prices randomly selected from the uniform distribution in the range $[1/100, 1/10]$. We simulate $N = \{20, 80, 160, 480\}$ choices in each menu, using a parametric ORUM with CES utilities. In accordance with the risk analysis, we set a curvature of the CES functions of $\alpha = .4$. The parameters of the logistic distribution governing altruism are $\tau = .2$ and $\sigma = .2$.[39]

Step 1. We start by ordering the collection of menus in decreasing order of π_j. Table 14 in Appendix D reports the simulated menus. ◇

Step 2. We compute the empirical distribution functions. That is, for each menu, we order the responses in increasing order in terms of the consumption in state 2. Figure 6 reports the empirical CDF of a sample of ten randomly selected menus. ◇

We now illustrate the nonparametric case. Note that, as already discussed in the political economy example, this application is more amenable than the risk one, since the structure of WARP is simpler than that of SARCEU. In a consumer setting, Kitamura and Stoye (2018) show that one can discretize choices to work with finite partitions, called patches, and Hoderlein and Stoye (2015) exploit this technique for the case of two goods. This approach, within our setting, reads as follows. Suppose that a quantile c^p fails to satisfy WARP. Then, there must exist two menus $A_j, A_{j'}$, with $\pi_j > \pi_{j'}$, where this violation occurs. Denoting by $x_{j,j'}$ the allocation where these two budget lines intersect, the violation guarantees that $F_j(v_j(x_{j,j'})) < p \leq F_{j'}(v_{j'}(x_{j,j'}))$. Then,

[39] These parameters give choice patterns (e.g., corner choices) that are in line with empirical observations in budget set experiments (see Andreoni and Miller (2002)).

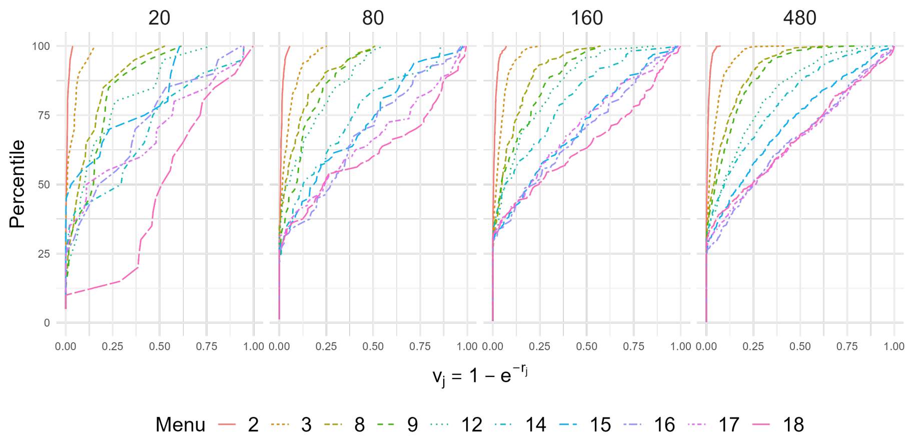


FIGURE 6. Empirical CDF of $v_j = 1 - e^{-r_j}$ across menus

the analysis of Quantile WARP simply requires to check $F_j(v_j(x_{j,j'})) \geq F_{j'}(v_{j'}(x_{j,j'}))$ for every pair of menus.

Step 3. Table 6 reports the number of violations of WARP. ◇

N obs	Number of violations	Menus with positive share of violations
20 obs	3	4, 13, 18
80 obs	2	5, 8
160 obs	0	-
480 obs	0	-

TABLE 6. Non-rationalizable menus in altruism application

For the semi-nonparametric case, we focus on the study of WARCESP for the altruism parameter. In order to do so, we first recover a value of α from the data as follows.

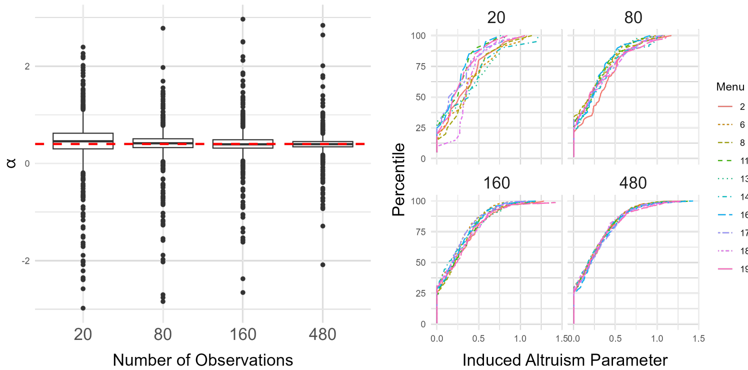
Step 4. We compute, for every pair of menus and interior quantiles, the value $\frac{\log(\pi_j/\pi_{j'})}{\log(r_j(c_j^p)/r_{j'}(c_{j'}^p))} + 1$, which corresponds to the value of α in the model. Then, we set α to be the median of these values. Figure 7a shows the distribution of α obtained when sampling and combining the interior quantiles $\{40, 60, 80, 100\}$ for all menus. Table 7 reports the median values. ◇

	20 obs	**80 obs**	**160 obs**	**480 obs**	**True**
α	.456	.415	.393	.397	.4

TABLE 7. Median α as described in Step 4

Having set the value of α, we can obtain the induced map between allocations and types, which, as discussed above, is given by $t = \pi_j r_j(x)^{1-\alpha} = \pi_j(\log \frac{1}{1-v_j(x)})^{1-\alpha}$.

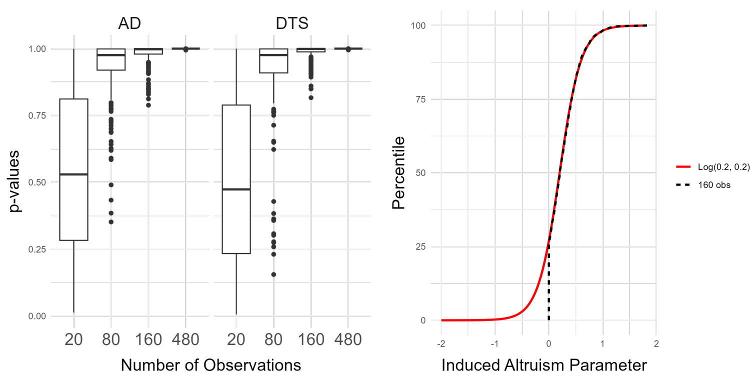
Step 5. We compute the empirical distributions of the induced altruism coefficients for the 20 menus, and Figure 7b plots 10 of them, randomly selected. ◇



(A) Step 4 (B) Step 5

FIGURE 7. Distribution of α across datasets (left); and Empirical CDF of induced altruism coefficients across menus (right)

Step 6. As required by Quantile WARCESP, we test for the equality of distributions. Figure 8a reports the distribution of p-values resulting from testing the equality of distributions across all pairs of menus in each dataset, comparing these on the basis of the same 20 quantiles, according to Anderson-Darling and DTS tests. In addition, we merge all observations and plot the empirical distribution of the induced altruism coefficients. Figure 8b reports it, for the case of $N = 160$, together with the CDF of the underlying logistic. ◇



(A) (B)

FIGURE 8. Boxplot of distribution of p-values across each pair of menus (left); and aggregated empirical CDF of induced altruism coefficients for $N = 160$ (right)

We now turn to the analysis of the parametric model.

Step 7. Note that CE_A requires some mass at $v = 0$ and no mass at $v = 1$. Table 15 in Appendix D reports the observed mass at the corner, and Table 16 in Appendix D the largest value of v per menu. ◇

Step 8. We now focus on property CLA_A by using the value of α obtained in Step 4. We consider all triples formed by one menu A_j and two interior observations in this menu. Take the induced altruism coefficients corresponding to these two observations, as calculated in Step 5, and compute its sum. Take the empirical quantiles of these two observations, as calculated in Step 2, transform them into log-odds and compute its sum. CLA_A requires the relationship between these two sums to be a linear function. Figure 9 plots this relationship for all pairs formed by combining 5 randomly selected interior choices per menu, and Table 8 reports the results of the estimation of the parameters via linear regression. ◇

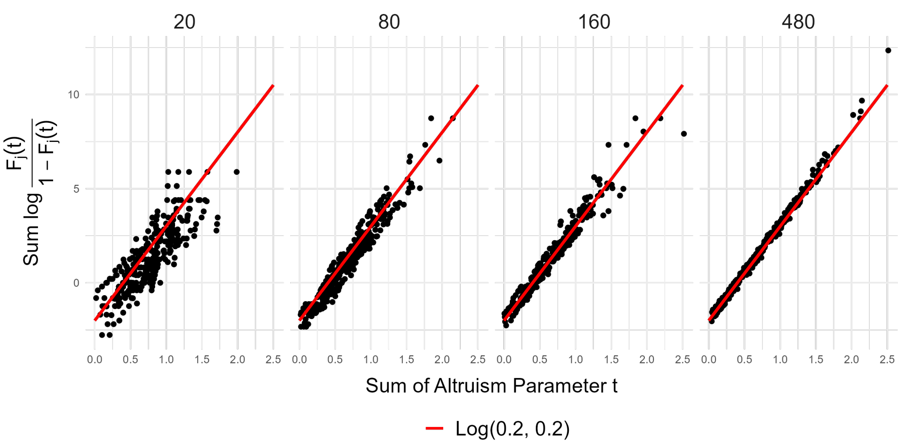


FIGURE 9. CLA property in altruism application

Parameter	20 obs	80 obs	160 obs	480 obs	True
τ	.215	.227	.198	.198	.2
σ	.259	.199	.202	.192	.2

TABLE 8. CLA parameter estimation via linear regression

The statistical analysis requires the same steps than in the risk domain.

Step 9. Table 9 reports the results of the maximum likelihood estimation, conditional on the number of observations. ◇

Parameter	20 obs	80 obs	160 obs	480 obs	True
τ	.272	.228	.205	.199	.2
σ	.211	.228	.198	.197	.2

TABLE 9. Parameter estimation via Maximum Likelihood

Step 10. Table 10 reports the results of the Kolmogorov-Smirnov and CVM tests. ◇

N **obs**	Rounded share of non-censored data	Test Statistic KS	p-value KS	Test Statistic CVM	p-value CVM
20 obs	.8	.595	$>.5$	.049	$>.5$
80 obs	.8	.461	$>.5$	.042	$>.5$
160 obs	.7	.472	$>.5$	.025	$>.5$
480 obs	.7	.382	$>.5$	.013	$>.5$

Table 10. KS and CVM tests for logistic distribution with left random censoring

8. Final Remarks

This paper investigates the empirical implications of commonly imposed parametric assumptions in ORUMs, specifically focusing on the case where menus are represented as linear segments, a structure that aligns well with several important economic applications. Importantly, notice that when a type-utility map is fixed, as in the semi-nonparametric and parametric cases, maximizers form a one-dimensional curve ordered by the type-utility map. As a result, the analysis can be based on these induced ordered curves, and hence, our results apply far beyond the setting and applications considered.[40] In Appendix B, we also extend our results to the case of finite problems, demonstrating the flexibility of the setting.

Appendix A. Proofs

Proof of Theorem 1: We start by proving the necessity part. Suppose that data F is ORUM-rationalizable with type-utility map γ and type distribution G, and consider any probability value $p \in (0,1)$. Given that G is continuous and strictly increasing, there is a unique type $G^{-1}(p)$ for which the cumulative mass is equal to p, i.e.,

$$G^{-1}(p) = \{t : G(t) = p\} = \min_{t:G(t)\geq p} t.$$

We claim that, for every menu A_j, the utility associated to type $G^{-1}(p)$, $\gamma(G^{-1}(p))$, has maximizer c_j^p. Suppose, by way of contradiction that this is not the case, i.e., suppose that there exists a menu A_j for which the maximizer of $\gamma(G^{-1}(p))$ is an alternative x such that $a_j(x) \neq a_j(c_j^p)$. If $a_j(x) < a_j(c_j^p)$, the ordered structure of choices generated by utilities guarantees that for every type $t \leq G^{-1}(p)$, $a_j(\gamma(t)) \leq a_j(x)$. Since (γ, G) rationalizes data, it must be $F_j(a_j(x)) \geq G(G^{-1}(p)) = p$, and x contradicts the definition of c_j^p. If $a_j(x) > a_j(c_j^p)$, it must be $c_j^p \neq \overline{x}_j$ and given the assumptions, we know that there is a type t^* such that $a_j(\gamma(t)) > a(c_j^p)$ if and only if $t > t^*$. It is $t^* < G^{-1}(p)$ and, given the assumptions on G and the rationalization of data, the mass below c_j^p would be strictly lower than p, a contradiction. We have then proved that

[40] In particular, no restriction on the linearity of menus or on the number of goods is of relevance in the semi-nonparametric and parametric cases. For the nonparametric results, it would be critical to examine the conditions under which such one-dimensional curves can be derived.

the maximizers of type $G^{-1}(p)$ are described by the choice function c^p. Given classical results, this choice function must be acyclical, and necessity follows.

We now prove the sufficiency part. Suppose that for a given probability value $p \in (0,1)$ the choice function c^p is acyclical. Given classical results, we can construct a utility function $U^p \in \mathcal{U}$ that is strictly positive over the finite set of alternatives $\{c_j^p\}_{j=1}^J$ and such that $U^p(x) = 0$ for the rest of alternatives in X, that rationalizes these p-quantile choices. That is, for every menu A_j, U^p gives c_j^p as a unique maximizer. Assign, to every type $t \in \mathbb{R}$, the utility function $\gamma(t)$ given by $U^{\frac{1}{1+e^{-t}}}$. We first claim that, for every menu, the maximizers induced by γ are increasing in t. To see this, consider any menu A_j and, given the construction, we merely need to prove that choices $c_j^{\frac{1}{1+e^{-t}}}$ are increasing in t, which follows immediately from the quantile definition of such alternatives and the obvious fact that F_j is non-decreasing. The strictly increasing nature in the interior and the continuity of choices follow similarly. Now, consider G to be the logistic distribution with location 0 and variance 1. We claim that, for every menu A_j and every $a \in (0,1)$, $F_j(a)$ coincides with the mass of types maximizing below a. Given the non-decreasing nature of the maximizing alternatives, we need to prove that the utility function $U^{F_j(a)}$ is the last utility with maximizer below a. First, consider $p > F_j(a)$. Since a has not reached cumulative probability p, it must be that $a(c_j^p) > a$, and since utility U^p rationalizes c_j^p, the maximizer of U^p lies strictly above a. Second, consider the utility function $U^{F_j(a)}$. By construction, $c_j^{F_j(a)}$ lies below a and since $U^{F_j(a)}$ rationalizes $c_j^{F_j(a)}$, the maximizer of $U^{F_j(a)}$ lies below a. This concludes the sufficiency part and the proof. ■

Proof of Theorem 2: Necessity follows directly from Theorem 1. For sufficiency, let $G \in \mathcal{G}$ and define U^p as in the proof of Theorem 1. Then, for every type $t \in \mathbb{R}$, define $\gamma(t)$ as the utility $U^{G(t)}$. That is, type t is assigned the utility that corresponds exactly to the quantile, according to G, of this type. The ordered-choice structure of type-utility map γ is immediate and ORUM-rationalizability follows from reproducing the proof of Theorem 1 with the pair (γ, G). ■

Proof of Theorem 3: Necessity follows directly from Theorem 1. For sufficiency, let $\gamma \in \Gamma$. Following Theorem 1, we can select $U^p \in \gamma(\mathbb{R})$. When constructing the bijection from the uniform distribution on $(0,1)$ to the reals, we need to respect map γ. This is always possible since no assumption is made on the type-distribution. ■

Proof of Theorem 4: Since the necessity of the axioms is straightforward, we will now prove sufficiency. Consider any menu A_j. We construct a sequence of open intervals of types, $\{I_j^0, I_j^1, \ldots, I_j^n, \ldots\}$, and a sequence of real functions defined over them, $\{G_j^0, G_j^1, \ldots, G_j^n, \ldots\}$, satisfying the following four properties:

(1) For every n, $I_j^n \subseteq I_j^{n+1}$.
(2) For every n, G_j^{n+1} extends G_j^n.

(3) For every n, G_j^n takes values in $(0,1)$, is continuous, and strictly increasing. Moreover, if I_j^n is bounded from above (respectively, from below), the function G_j^n must be strictly bounded from above by some value $k < 1$ (respectively, strictly bounded from below by some value $k > 0$).
(4) For every n and every four types t_1, t_2, t'_1, t'_2 in I_j^n, if $t_1 + t_2 = t'_1 + t'_2$ then $\log \frac{G_j^n(t_1)}{1-G_j^n(t_1)} + \log \frac{G_j^n(t_2)}{1-G_j^n(t_2)} = \log \frac{G_j^n(t'_1)}{1-G_j^n(t'_1)} + \log \frac{G_j^n(t'_2)}{1-G_j^n(t'_2)}$.

The first interval of types, I_j^0, corresponds to the set of types that have a maximizer in $A_j \setminus \{\underline{x}_j, \overline{x}_j\}$.[41] The first function, G_j^0, corresponds to the function that choice data F_j induces over these types, i.e., for every $t \in I_j^0$, $G_j^0(t) = F_j(a_j(\gamma(t)))$. The function G_j^0 is well-defined given the assumptions made on F_j. It is obviously strictly increasing and takes values in $(0,1)$. Moreover, if the interval I_j^0 is bounded from above (respectively, from below), there is an interval of types selecting $\overline{x}_j$ (respectively, $\underline{x}_j$) and property CE implies $\lim_{a \to 1} F_j(a) < 1$ (respectively, $\lim_{a \to 0} F_j(a) > 0$), and hence the boundedness conditions hold for G_j^0. That is, property (3) is satisfied. Property (4) for G_j^0 follows from CLA by considering $A_{j'} = A_j$.

The remaining intervals and functions are now defined recursively. Given collections $\{I_j^0, I_j^1, \ldots, I_j^n\}$ and $\{G_j^0, G_j^1, \ldots, G_j^n\}$ which satisfy all the properties, we define interval I_j^{n+1} and function G_j^{n+1} in such a way as to guarantee that collections $\{I_j^0, I_j^1, \ldots, I_j^{n+1}\}$ and $\{G_j^0, G_j^1, \ldots, G_j^{n+1}\}$ also satisfy the properties. The definition of the new interval of types, I_j^{n+1}, depends on the parity of n. If n is an even (respectively, an odd) integer, we define interval I_j^{n+1} as follows: (i) if I_j^n is not bounded from above (respectively, from below), define $I_j^{n+1} = I_j^n$ and (ii) if I_j^n is bounded from above (respectively, from below), define I_j^{n+1} as the union of the previous interval I_j^n, the least upper bound (respectively, the greatest lower bound) z_j^n of interval I_j^n, and the types t for which there exists $t' \in I_j^n$ with $t = 2z_j^n - t'$.[42]

We now consider the definition of function G_j^{n+1}. For every $t \in I_j^n$, define $G_j^{n+1}(t) = G_j^n(t)$. For the limit type z_j^n, define $G_j^{n+1}(z_j^n) = \lim_{s \to z_j^n} G_j^n(s)$, where the right-hand or left-hand bound must be considered, depending on the parity. Finally, for any other type t belonging to I_j^{n+1}, we know that there exists a unique value $t' \in I_j^n$ such that

[41] Notice that this set of types depends on the assumed type-utility map γ. Since γ is fixed and to simplify the exposition, we will avoid some references to γ in the arguments that follow.

[42] Intuitively we are extending the original right-bounded (respectively, left-bounded) interval I_j^n beyond its boundary and adding the boundary point. This step is not needed when there are no corner choices because then the initial interval I_j^0 equals the set of all types, $\mathbb{R}$. When choices are observed in only one of the corner alternatives, or, equivalently, I_j^0 is bounded on one side, the logic requires a unique duplication, which already forms the entire real line. If choices are observed in both corner alternatives, or, equivalently, the initial interval is bounded on both sides, we need to duplicate the initial bounded interval an infinite number of times, as the proof indicates.

$t = 2z_j^n - t'$, so we can define $G_j^{n+1}(t)$ as the unique real value satisfying the equation:

$$\log \frac{G_j^{n+1}(t)}{1 - G_j^{n+1}(t)} = 2 \log \frac{G_j^{n+1}(z_j^n)}{1 - G_j^{n+1}(z_j^n)} - \log \frac{G_j^n(t')}{1 - G_j^n(t')}.$$

It is then evident that the function G_j^{n+1} is well defined on I_j^{n+1} and it is straightforward to see that $I_j^n \subseteq I_j^{n+1}$ and, hence, property (1) holds. Similarly, note that the construction guarantees that the function G_j^{n+1} extends G_j^n, and therefore property (2) is satisfied.

We now discuss property (3). Notice that, by the continuity of G_j^n and the fact that all values belong to $(0, 1)$, it is guaranteed that the limit value at z_j^n is well defined when needed. The continuity of the function G_j^{n+1} is then a direct consequence of this limit definition at z_j^n. To appreciate the strictly increasing nature of the new function, consider two types $t_1 < t_2$. If both types belong to I_j^n, we know that $G_j^{n+1}(t_1) < G_j^{n+1}(t_2)$ must hold because G_j^{n+1} extends the strictly increasing function G_j^n. If $t_1 \in I_j^n$ but t_2 does not, it must be the case that n is even and there exists $t'_2 \in I_j^n$ such that $t_2 = 2z_j^n - t'_2$. Since $\log \frac{G_j^n(z_j^n)}{1-G_j^n(z_j^n)} > \log \frac{G_j^n(t'_2)}{1-G_j^n(t'_2)}$, it is $\log \frac{G_j^{n+1}(t_1)}{1-G_j^{n+1}(t_1)} = \log \frac{G_j^n(t_1)}{1-G_j^n(t_1)} < \log \frac{G_j^n(z_j^n)}{1-G_j^n(z_j^n)} < 2 \log \frac{G_j^n(z_j^n)}{1-G_j^n(z_j^n)} - \log \frac{G_j^n(t'_2)}{1-G_j^n(t'_{,,})} = \log \frac{G_j^{n+1}(t_2)}{1-G_j^{n+1}(t_2)}$, as desired. If t_1 is not in I_j^n but t_2 is, an analogous argument applies in which n is odd and z_j^n is the lower bound of I_j^n. If neither is in I_j^n, they must both be above or below z_j^n, depending on the parity. There must exist $t'_1, t'_2 \in I_j^n$ such that $t_1 = 2z_j^n - t'_1$ and $t_2 = 2z_j^n - t'_2$. It clearly must be that $t'_1 > t'_2$ and we know that $G_j^n(t'_1) > G_j^n(t'_2)$. The definition of $G_j^{n+1}(t_1)$ and $G_j^{n+1}(t_2)$ guarantees that the former is strictly smaller than the latter. Hence, we have shown that G^{n+1} is strictly increasing and, to complete property (3), we need to show that this function takes values in $(0, 1)$ and is bounded as required. We show the case of n being even, the other case being analogous. If I_j^n is not bounded from above, the new function replicates the original one and the property holds. If I_j^n is bounded from above, we know that the value $G^{n+1}(z_j^n)$ must be strictly lower than 1 by virtue of the boundedness condition. For every $t \in I_j^{n+1}$ with $t > z_j^n$, the construction guarantees that G_j^{n+1} takes values in $(0, 1)$. To show boundedness, notice that nothing changes at the lower end of the interval and, since G_j^{n+1} extends G_j^n, the property is satisfied. For the upper end of the interval, suppose that I_j^{n+1} is bounded from above, in which case it must be that I_j^n is bounded from below (say, with largest lower bound k). It then follows that $\log \frac{G_j^{n+1}(t)}{1-G_j^{n+1}(t)} < 2 \log \frac{G_j^{n+1}(z_j^n)}{1-G_j^{n+1}(z_j^n)} - \log \frac{G_j^n(k)}{1-G_j^n(k)}$, and hence $G_j^{n+1}(t)$ must be strictly lower than 1. This completes the proof that G_j^{n+1} satisfies property (3).

To see that property (4) holds, consider any four types t_1, t_2, t'_1, t'_2 in I_j^{n+1} such that $t_1 + t_2 = t'_1 + t'_2$ and assume, without loss of generality, that $t_1 < t'_1 \leq t'_2 < t_2$.[43] Again, we show the case of n even, the other case being analogous. We start by noticing

[43]Notice that if the types were equal across the two pairs, the property would be trivially satisfied.

that property (4) holds over the closure of I_j^n, denoted by $\overline{I}_j^n$, thanks to the recursive assumption on G_j^n, the fact that G_j^{n+1} extends G_j^n, and the limit construction at z_j^n. Hence, we only need to consider cases where not all four types belong to $\overline{I}_j^n$:

- Case 1: None of the four types belongs to $\overline{I}_j^n$. There must exist $s_1, s_2, s_1', s_2' \in I_j^n$ such that $t_1 = 2z_j^n - s_1$, $t_1' = 2z_j^n - s_1'$, $t_2 = 2z_j^n - s_2$ and $t_2' = 2z_j^n - s_2'$. Clearly, it must be that $s_1 + s_2 = s_1' + s_2'$ and hence, we know that $\log \frac{G_j^n(s_1)}{1-G_j^n(s_1)} + \log \frac{G_j^n(s_2)}{1-G_j^n(s_2)} = \log \frac{G_j^n(s_1')}{1-G_j^n(s_1')} + \log \frac{G_j^n(s_2')}{1-G_j^n(s_2')}$, which is equivalent to $\log \frac{G_j^n(s_1)}{1-G_j^n(s_1)} + \log \frac{G_j^n(s_2)}{1-G_j^n(s_2)} + 4G_j^{n+1}(z_j^n) = \log \frac{G_j^n(s_1')}{1-G_j^n(s_1')} + \log \frac{G_j^n(s_2')}{1-G_j^n(s_2')} + 4G_j^{n+1}(z_j^n)$, which implies $\log \frac{G_j^n(t_1)}{1-G_j^n(t_1)} + \log \frac{G_j^n(t_2)}{1-G_j^n(t_2)} = \log \frac{G_j^n(t_1')}{1-G_j^n(t_1')} + \log \frac{G_j^n(t_2')}{1-G_j^n(t_2')}$, as desired.
- Case 2: $t_1 \in \overline{I}_j^n$. There must exist $s_2, s_1', s_2' \in I_j^n$ such that $t_1' = 2z_j^n - s_1'$, $t_2 = 2z_j^n - s_2$ and $t_2' = 2z_j^n - s_2'$. It must be that $t_1 + s_1' + s_2' = s_2 + 2z_j^n$. Define $\hat{t} = s_2 + z_j^n - t_1$, which belongs to I_j^n. Given that $t_1 + \hat{t} = s_2 + z_j^n$, property (4) holds over these four types. Now, notice that it must also be that $s_1' + s_2' = \hat{t} + z_j^n$ and hence property (4) holds over these four types. We can combine the two expressions to verify that property (4) holds over t_1, t_2, t_1' and t_2', as desired.
- Case 3: $t_1, t_1' \in \overline{I}_j^n$. There must exist $s_2, s_2' \in I_j^n$ such that $t_2 = 2z_j^n - s_2$, and $t_2' = 2z_j^n - s_2'$. It must be that $t_1 + s_2' = t_1' + s_2$ and hence, we know that $\log \frac{G_j^n(t_1)}{1-G_j^n(t_1)} + \log \frac{G_j^n(s_2')}{1-G_j^n(s_2')} = \log \frac{G_j^n(t_1')}{1-G_j^n(t_1')} + \log \frac{G_j^n(s_2)}{1-G_j^n(s_2)}$, which implies $\log \frac{G_j^n(t_1)}{1-G_j^n(t_1)} + \log \frac{G_j^n(s_2')}{1-G_j^n(s_2')} + 2G_j^{n+1}(z_j^n) = \log \frac{G_j^n(t_1')}{1-G_j^n(t_1')} + \log \frac{G_j^n(s_2)}{1-G_j^n(s_2)} + 2G_j^{n+1}(z_j^n)$, which implies $\log \frac{G_j^n(t_1)}{1-G_j^n(t_1)} + \log \frac{G_j^n(t_2)}{1-G_j^n(t_2)} = \log \frac{G_j^n(t_1')}{1-G_j^n(t_1')} + \log \frac{G_j^n(t_2')}{1-G_j^n(t_2')}$, as desired.
- Case 4: $t_1, t_1', t_2' \in \overline{I}_j^n$. There must exist $s_2 \in I_j^n$ such that $t_2 = 2z_j^n - s_2$. It must be that $t_1 + 2z_j^n = t_1' + t_2' + s_2$. Define $\hat{t} = t_1 + z_j^n - t_1'$, which belongs to I_j^n. Given that $t_1' + \hat{t} = t_1 + z_j^n$, property (4) holds over these four types. Now, notice that it must also be that $\hat{t} + z_j^n = t_2' + s_2$ and hence property (4) holds over these four types. We can combine the two expressions to verify that property (4) holds over t_1, t_2, t_1' and t_2', as desired.

This completes the proof that the collections $\{I_j^0, I_j^1, \dots, I_j^{n+1}\}$ and $\{G_j^0, G_j^1, \dots, G_j^{n+1}\}$ satisfy all the properties. The limit interval of the sequence $\{I_j^0, I_j^1, \dots, I_j^n, \dots\}$ is the entire set of reals. The limit function of the sequence $\{G_j^0, G_j^1, \dots, G_j^n, \dots\}$, which we denote by G_j, must be a continuous, strictly increasing CDF over the reals. Moreover, it extends G_j^0 and must also satisfy property (4) above.

Consider the median type of distribution G_j, i.e., the type τ_j such that $G_j(\tau_j) = .5$. Define the function H_j over the reals as follows:

$$H_j(w) = G_j(\tau_j + w).$$

We claim that H_j is a continuous, strictly increasing CDF over the reals, and is symmetric with respect to the origin. We need to show symmetry. For this, consider $t_1 = \tau_j - w$, $t_2 = \tau_j + w$ and $t_1' = t_2' = \tau_j$. Then, since $t_1 + t_2 = t_1' + t_2'$, we know that $\log \frac{G_j(t_1)}{1-G_j(t_1)} + \log \frac{G_j(t_2)}{1-G_j(t_2)} = \log \frac{G_j(t_1')}{1-G_j(t_1')} + \log \frac{G_j(t_2')}{1-G_j(t_2')} = 0 + 0 = 0$. Hence, it must be that $\log \frac{G_j(t_1)}{1-G_j(t_1)} = \log \frac{1-G_j(t_2)}{G_j(t_2)}$ and $G_j(t_1) = 1 - G_j(t_2)$ follows. As a result, $H_j(-w) = G_j(t_1) = 1 - G_j(t_2) = 1 - H_j(w)$, and the symmetry of H_j has been proved.

Consider now the following function defined over the positive reals:

$$O_j(w) = \frac{1 - H_j(w)}{H_j(w)}.$$

Since H_j is a continuous, strictly increasing CDF over the reals with $H_j(0) = .5$, $1 - O_j(w)$ must be a continuous, strictly increasing CDF over the positive reals with no strictly positive mass at zero. Moreover, given that G_j satisfies property (4) above, the definition of H_j and O_j guarantees that $O_j(w)O_j(z) = O_j(w + z)$ must hold for every pair of positive real values w and z. One can then reproduce the standard argument dating back to Cauchy (1821), which is described in Galambos and Kotz (1978; Theorem 1.3.1), to guarantee that O_j must be an exponential distribution, with no strictly positive mass at the origin.[44] That is, there exists $\sigma_j \in \mathbb{R}_{++}$ such that

$$1 - O_j(w) = 1 - \frac{1 - H_j(w)}{H_j(w)} = 1 - e^{-w/\sigma_j},$$

and hence, for every $w \geq 0$, it is true that $H_j(w) = \frac{1}{1+e^{-w/\sigma_j}}$. Moreover, given the symmetry of H_j with respect to the origin, for every $w < 0$, it must also be true that $H_j(w) = 1 - H_j(-w) = 1 - \frac{1}{1+e^{w/\sigma_j}} = \frac{1}{1+e^{-w/\sigma_j}}$. That is, H_j is a logistic distribution with location parameter equal to zero and scale parameter σ_j, and G_j is ordered logistic with location parameter τ_j and scale parameter σ_j. Since G_j extends G_j^0, all choices in menu A_j are explained by this distribution.

Consider now two menus A_j and $A_{j'}$. By our richness assumption, there exists a sequence of menus $A_{j_0} = A_j, A_{j_1}, \ldots, A_{j_k}, \ldots, A_{j_K} = A_{j'}$ such that, for every $k \in \{0, \ldots, K-1\}$, $I^0_{j^k} \cap I^0_{j^{k+1}} \neq \emptyset$. Consider $t \in I^0_{j^k} \cap I^0_{j^{k+1}}$ and take $t_1 = t_2 = t_1' = t_2' = t$. Using the ordered-logit structure of G_{j^k} and $G_{j^{k+1}}$, it follows that they must both have a common location parameter τ and a common scale parameter σ. The recursive application of this argument shows that G_j and $G_{j'}$ must have the same common parameters τ and σ, which concludes the proof. ■

Proof of Corollary 1: The first part follows immediately from Theorem 1 in Kubler, Selden and Wei (2014) and from our Theorem 1. The second part follows from our Theorem 3 after noticing that $\frac{\log \phi}{\log r}$ represents the negative of the curvature of the

[44]The property is satisfied by exponential distributions with and without strictly positive mass at zero. Since we know that O has no strictly positive mass at zero, it must be one of the latter.

CRRA function maximizing at the consumption ratio r, with strictly negative values representing risk-aversion and null values corresponding to any other risk attitude. ■

Proof of Corollary 3: The first part follows immediately from classical deterministic results on consumption theory. Notice that utilities that are strictly monotone in the first component always select the corner $x^2 = 0$, and generate choices that are consistent with WARP. Also, utilities that are strictly increasing in both components and strictly convex must also lead to choices consistent with WARP. Moreover, satisfaction of WARP allows to rationalize choices with such utility functions (see, e.g. Rose (1958) and Matzkin and Richter (1991)). Then, the application of our Theorem 1 concludes the proof. For the second part, notice that when $t \leq 0$, all choices are such that $x^2 = 0$. When $t > 0$, the first-order condition analysis for a given budget set allows to obtain $t = \pi_j r_j(x)^{1-\alpha}$. By equalizing this value for two different budget sets $A_j, A_{j'}$, one can then see that $\alpha = \frac{\log(\pi_j/\pi_{j'})}{\log(r_j(x)/r_{j'}(x'))} + 1$, which must then be constant across pairs of menus. If deterministic data has this constancy with interior solutions, one can define α to be such constant, and then define t from one menu, using α. For corner solutions, one can set $t \leq 0$ and choose freely α. The utility function defined in such ways rationalizes the observations. To apply our Theorem 1, notice that we want the utility function to use a constant α across quantiles, which is guaranteed by our assumption. ■

Appendix B. Discrete choice

We present here some simple discrete versions of our results, serving several purposes. First, they illustrate that our analysis is not confined to continuous settings. Second, they may be useful for practitioners working with discrete choice environments. Third, they help provide a more comprehensive understanding of the theoretical results developed in this paper in relation to existing literature.

Let $(X, <)$ be a finite, partially ordered set of alternatives. Denote by $\mathcal{X}$ the set of all subsets of X containing at least two alternatives, and by $\mathcal{D} \subseteq \mathcal{X}$ an arbitrary domain consisting of fully ordered menus, i.e., the restriction of $<$ to any $A \in \mathcal{D}$ is complete. An ordered menu A_j can be written as $A_j = \{1_j, 2_j, \ldots, i_j, \ldots, I_j + 1\}$, where $1_j < 2_j < \cdots < i_j < \cdots < I_j + 1$.

Choice data is described by a stochastic choice function ρ defined on $\mathcal{D}$. Specifically, $\rho(x, A) \in [0, 1]$, with $x \in A$ and $A \in \mathcal{D}$, represents the probability of choosing alternative x from menu A, with the constraint $\sum_{x \in A} \rho(x, A) = 1$. We denote cumulative choice data by $\bar{\rho}$, defined as $\bar{\rho}(i_j, A_j) = \sum_{k \leq i_j} \rho(k, A_j)$.

The definition of ORUM-rationalizability in this discrete setting follows directly from the one in the main text, with F replaced by $\bar{\rho}$. Given the finite nature of this setting, the image of γ can be interpreted as a finite ordered collection of linear orders, with G being the CDF of a finite probability distribution g.

B.1. **Nonparametric case.** To present a version of Theorem 1 in this discrete setting, we define the p-quantile choice function c^p. Formally, $c^p(A)$ is the first alternative in A that reaches cumulative probability mass p, i.e., $\bar{\rho}(c^p(A)-1, A) < p \leq \bar{\rho}(c^p(A), A)$.[45] A quantile choice function c^p is said to be rationalizable if there exists a linear order P on X such that, for every $A \in \mathcal{D}$, $c^p(A) = m(A, P)$, i.e., $c^p(A)$ is the maximal element in A according to P. It is well known that, in arbitrary and finite domains, deterministic rationalizability is equivalent to the acyclicity notion used in the main text.

Proposition 1. *ρ is ORUM-rationalizable if and only if ρ satisfies Quantile Acyclicity.*

Proof of Proposition 1: Suppose first that ρ is ORUM-rationalizable by means of (γ, g). Given the finiteness of the setting, (γ, g) induces a finite, ordered collection of linear orders $\{P_t\}_{t=1}^T$, with increasing maximizers in every menu, and a finite probability distribution over $\{1, \ldots, T\}$, denoted by g for simplicity. Consider any $p \in (0, 1)$, and let $t(p)$ be such that $\sum_{t<t(p)} g(t) < p \leq \sum_{t \leq t(p)} g(t)$. Given the increasing nature of the maximizers of $\{P_t\}_{t=1}^T$, it follows that, for every $A \in \mathcal{D}$, the quantile choices $\{c^p(A)\}_{A \in \mathcal{D}}$ coincide with the maximizers of $P_{t(p)}$. Hence, these choices must be rationalizable and, by classical results, must satisfy Acyclicity.

Conversely, suppose that c^p satisfies Acyclicity for every $p \in (0, 1)$. These choices can be rationalized by a linear preference, denoted P_p. Since the setting is finite, there are only finitely many distinct linear preferences in the family $\{P_p\}_{p \in (0,1)}$, denoted $\{P_t\}_{t=1}^T$. By discarding preferences that generate identical maximizers in the domain $\mathcal{D}$, we can assume, without loss of generality, that distinct preferences in this collection yield different choice functions. Furthermore, given the quantile definition, each of these preferences must correspond to an interval of p-values. Thus, we can assume that the family $\{P_t\}_{t=1}^T$ is ordered, i.e., P_t corresponds to lower p-values than $P_{t'}$ whenever $t < t'$. We can then assign to each preference P_t the Lebesgue measure of the corresponding interval of p-values, denoted $g(t)$. This pair $(\{P_t\}_{t=1}^T, g)$ rationalizes the data, so ρ is ORUM-rationalizable. ■

Proposition 1, a discrete version of Theorem 1 in the main text, clarifies the difference with the SCRUM model of Apesteguia, Ballester, and Lu (2017). Essentially, ORUM-rationalization is a more general model that allows for X to be partially ordered and the consideration of an arbitrary domain of menus.

Proposition 2 below shows that when X is fully ordered and choices are observed from every subset of X, ORUM-rationalization reduces to SCRUM-rationalization. To do so, we use the discrete version of the Weak Axiom of Revealed Preference (D-WARP) at the quantile level, where the direct revelations of the quantile choice function c^p must contain no cycles, and property α, where $c^p(A) \in B \subseteq A$ implies $c^p(B) = c^p(A)$. We say that ρ satisfies Quantile D-WARP (respectively, Quantile α) if every c^p satisfies WARP

[45]When $c^p(A) = 1_j$, we set $\bar{\rho}(c^p(A)-1, A) = 0$. Similarly, $i_j - 1$ is denoted as $(i-1)_j$. In the proof below, when $t(p) = 1$, we write $\sum_{t<t(p)} g(t) = 0$.

(respectively, property α). Consider also the two properties characterizing SCRUMs in Apesteguia, Ballester, and Lu (2017): Regularity and Centrality.[46]

Proposition 2. *Let $\mathcal{D} = \mathcal{X}$. ρ is ORUM-rationalizable (or SCRUM-rationalizable) if and only if ρ satisfies Quantile D-WARP, if and only if ρ satisfies Quantile α, if and only if ρ satisfies Regularity and Centrality.*

Proof of Proposition 2: Under $\mathcal{D} = \mathcal{X}$, classical results ensure that c^p satisfies Acyclicity if and only if it satisfies D-WARP, if and only if it satisfies property α, allowing us to apply Proposition 1. To prove the final statement, note that since $\mathcal{D} = \mathcal{X}$, all binary sets are part of the domain, and hence X must be completely ordered. It follows that ORUM-rationalization is equivalent to SCRUM-rationalization (a probability distribution over a single-crossing collection of preferences). The characterization result in Apesteguia, Ballester, and Lu (2017) completes the proof. ■

Thanks to Proposition 2, one can also see better how the properties of nonparametric ORUMs relate to the existing properties of SCRUMs. To illustrate, we show how the main property of SCRUMs, Centrality, is implied by Quantile Acyclicity in this restricted setting. Consider three alternatives, $x < y < z$, and suppose that the probability of choosing y from menu $\{x, y, z\}$ is strictly positive, i.e., $\rho(y, \{x, y, z\}) > 0$. Suppose, by way of contradiction, that $\rho(z, \{x, y, z\}) \neq \rho(z, \{y, z\})$. If $\rho(z, \{x, y, z\}) > \rho(z, \{y, z\})$, it must be $\rho(x, \{x, y, z\}) + \rho(y, \{x, y, z\}) < \rho(y, \{y, z\})$ and the quantile definition together with the ordered nature of choices guarantee the existence of a value p such that $c^p(\{x, y, z\}) = z$, but $c^p(\{y, z\}) = y$. If $\rho(z, \{x, y, z\}) < \rho(z, \{y, z\})$, and given that $\rho(y, \{x, y, z\}) > 0$, there must exist some p such that $c^p(\{x, y, z\}) = y$, but $c^p(\{y, z\}) = z$. In both cases, there is a contradiction with the fact that c^p satisfies the classical property α and, consequently, a violation of the acyclicity of c^p.

B.2. **Semi-nonparametric case.** We now present a discrete choice version of the semi-nonparametric results in the main text, which allows us to relate the present results to Apesteguia and Ballester (2023). To do so, fix a parametrization γ that generates ordered choices. That is, in any given menu $A_j \in \mathcal{D}$, types $(-\infty, t_j^1)$ uniquely select 1_j, types (t_j^1, t_j^2) uniquely select 2_j, and so on and so forth, with types $(t_j^{I_j}, \infty)$ uniquely selecting $I_j + 1$.[47] Let the utility functions in γ be characterized by a certain deterministic property, that we call $\gamma(\mathbb{R})$-Rationalizability and define Quantile $\gamma(\mathbb{R})$-Rationalizability accordingly.

The Monotonicity property of Apesteguia and Ballester (2023) reads as follows: for every $B \subseteq A \in \mathcal{D}$ and $B' \subseteq A' \in \mathcal{D}$, if the set of types maximizing in B within

[46]Centrality: $x < y < z$ or $z < y < x$ and $\rho(y, \{x, y, z\}) > 0$ imply $\rho(z, \{x, y, z\}) = \rho(z, \{y, z\})$.

[47]Since in the continuous setting types have no mass, the choice at threshold types is irrelevant and ties can be avoided. Moreover, we simplify notation and discard from a menu alternatives that are not maximal for any type.

A is included in the set of types maximizing in B' within A', then $\sum_{x\in B}\rho(x,A) \leq \sum_{x\in B'}\rho(x,A')$.

Proposition 3. *Let γ be any type-utility map. ρ is ORUM-rationalizable with type-utility map γ if and only if every ρ satisfies Quantile $\gamma(\mathbb{R})$-Rationalizability, if and only if ρ satisfies Monotonicity.*

Proof of Proposition 3 The result follows immediately from Proposition 1 and the characterization result of Apesteguia and Ballester (2023). ■

Proposition 3 shows that the model of Apesteguia and Ballester (2023) is equivalent to a discrete choice version of the semi-nonparametric model described in Theorem 3. The connection between the quantile rationalizability approach of the present paper, and the property of Monotonicity can be presented as follows. Suppose that ρ satisfies quantile rationalizability and assume, by way of contradiction, that Monotonicity fails. When Monotonicity fails, it can be shown that a cumulative violation of Monotonicity must exist. Namely, there exist (i_j, A_j) and $(i'_{j'}, A_{j'})$ such that the types maximizing below i_j in A_j contain the types maximizing below $i'_{j'}$ in $A_{j'}$, i.e., $t^i_j \geq t^{i'}_{j'}$, but we observe $\bar{\rho}(i_j, A_j) < \bar{\rho}(i'_{j'}, A_{j'})$. That is, every preference with property $\gamma(\mathbb{R})$-Rationalizability maximizing below $i'_{j'}$ in $A_{j'}$ must also maximize below i_j in A_j. Hence, given the inequality in cumulative choice probabilities the quantile choice function $c^{\rho(i'_{j'}, A_{j'})}$ violates property $\gamma(\mathbb{R})$-Rationalizability, a contradiction.

B.3. **Parametric case.** We now consider the parametric version of ordered logit in a discrete setting. To do so, fix a map γ and assume the logistic distribution. In the ordered-logit model the cumulative choice probability of alternatives $\{1_j, 2_j, \ldots, i_j\}$ in decision problem A_j is determined by the threshold type t^i_j as

$$G^{(\tau,\sigma)}(t^i_j) = \frac{1}{1+e^{-(t^i_j-\tau)/\sigma}}.$$

The ordered-logit rationalization of choice data requires, therefore, that $G^{(\tau,\sigma)}(t^i_j) = \bar{\rho}(i_j, A_j)$.

We assume Positivity that, given the ordered nature of the setting, is equivalent to $0 < \bar{\rho}(1_j, A_j) < \cdots < \bar{\rho}(i_j, A_j) < \cdots < \bar{\rho}(I_j, A_j) < 1 = \bar{\rho}(I_j+1, A_j)$ for every $A_j \in \mathcal{D}$. In addition, we make the following technical assumption: there exist threshold types $t^{i_1}_{j_1}, t^{i_2}_{j_2}, t^{i_3}_{j_3}, t^{i_4}_{j_4}$, and $t^{i^*}_{j^*}$ such that $\bar{\rho}(i_1, A_{j_1}) < \bar{\rho}(i^*, A_{j^*}) = .5 < \bar{\rho}(i_2, A_{j_2})$, and the value $\frac{t^{i_3}_{j_3}-t^{i^*}_{j^*}}{t^{i_4}_{j_4}-t^{i^*}_{j^*}}$ is an irrational number.

We introduce a version of CLA that characterizes ordered-logistic choice in discrete settings. Consider any two equally-sized collections of threshold types. The property states that the sum of the first collection is larger than that of the second if and only if the corresponding sum of cumulative log-odds in the former is larger than that in the latter.[48]

Discrete Cumulative Log-Odds Additivity (D-CLA). For every positive integer M, and for every two collections of threshold types $\{t^{i_1}_{j_1},\dots,t^{i_m}_{j_m},\dots,t^{i_M}_{j_M}\}$ and $\{t^{i'_1}_{j'_1},\dots,t^{i'_m}_{j'_m},\dots,t^{i'_M}_{j'_M}\}$, $\sum_{m=1}^{M} t^{i_m}_{j_m} \geq \sum_{m=1}^{M} t^{i'_m}_{j'_m}$ if and only if $\sum_{m=1}^{M} \ell_{j_m}(i_m) \geq \sum_{m=1}^{M} \ell_{j'_m}(i'_m)$.

Proposition 4. *Let γ be any type-utility map. ρ is ORUM-rationalizable with type-utility map γ and a type distribution in $\mathcal{G}^L$ if and only if ρ satisfies D-CLA.*

Proof of Proposition 4: In a discrete setting, it is immediate to see that any choice data that is rationalized by an ordered-logit model must satisfy D-CLA. We then need to prove the sufficiency part of the result. For this, we start by constructing a function over the reals. By assumption, there is a threshold type $t^{i^*}_{j^*}$ such that $\bar{\rho}(i^*, A_{j^*}) = .5$. Consider then the subsets of real numbers

$$\mathcal{T} = \{x : x = t^i_j - t^{i^*}_{j^*} \text{ for some } A_j \in \mathcal{D} \text{ and } i_j < I_j + 1\},$$

$$\mathcal{T}^{IC} = \{x : x \text{ is an integer combination of elements in } \mathcal{T}\}.$$

It is immediate to see that $\mathcal{T}^{IC}$ is a subgroup of the reals, and given our technical assumption on the existence of threshold types producing a ratio that is irrational, well-known results guarantee that $\mathcal{T}^{IC}$ must be dense in the reals.[49] We can then find, for every $x \in \mathbb{R}$, a sequence of elements $(x_1, x_2, \dots, x_k, \dots)$ in $\mathcal{T}^{IC}$ such that $x_k \to x$. Each of the elements x_k in this sequence is an integer combination of elements in $\mathcal{T}$ and hence, we can find collections of threshold types $\{t^{i_1}_{j_1},\dots,t^{i_v}_{j_v},\dots,t^{i_V}_{j_V}\}$ and $\{s^{i_1}_{j_1},\dots,s^{i_w}_{j_w},\dots,s^{i_W}_{j_W}\}$ such that $x_k = \sum_{v=1}^{V} n_v(t^{i_v}_{j_v} - t^{i^*}_{j^*}) - \sum_{w=1}^{W} n_w(s^{i_w}_{j_w} - t^{i^*}_{j^*})$, where all n_v and n_w are strictly positive integers. Consider the real value $H_k(x)$ that solves the equality $\log \frac{H_k(x)}{1-H_k(x)} = \sum_{v=1}^{V} n_v \ell_{j_v}(i_v) - \sum_{w=1}^{W} n_w \ell_{j_w}(i_w)$. Denoting by $H(x)$ the limit of the sequence of values formed by $H_k(x)$, we have constructed a function H over the reals.

First of all, notice that for any given threshold type t, there may be several decision problems which have this value t as a threshold type. D-CLA guarantees that the cumulative choice probability is the same in both decision problems, and hence the function H is well defined.[50] We now prove that H is increasing. Let $x < x'$. We know that there exist sequences of elements $(x_1, x_2, \dots, x_k, \dots)$ and $(x'_1, x'_2, \dots, x'_k, \dots)$ in $\mathcal{T}^{IC}$ such that $x_k \to x$ and $x'_k \to x'$. Since $x < x'$, there exists K such that $x_k < x'_k$ for every $k \geq K$. Let $k \geq K$, and consider the integer representations of x_k given by $\{n_v, t^{i_v}_{j_v}\}_{v=1}^{V}$ and $\{n_w, s^{i_w}_{j_w}\}_{w=1}^{W}$ and of x'_k given by $\{n'_v, t^{i'_v}_{j'_v}\}_{v=1}^{V'}$ and $\{n'_w, s^{i'_w}_{j'_w}\}_{w=1}^{W'}$.

[48] The cumulative log-odds in the discrete setting follow from the definition used in the continuous setting, with F replaced by $\bar{\rho}$. Notice that the continuous setting helps in the presentation of the property by focusing on pairs of types.

[49] See, e.g., Theorem 1.6 in Salzmann, Grundhöfer, Hähl and Löwen (2007).

[50] Indeed, the same idea applies to the extension of H to any real number, by using limits of integer combinations of threshold types. This argument is omitted below.

Consider the two positive integer values $\sum_{v=1}^{V} n_v + \sum_{w=1}^{W'} n'_w$ and $\sum_{v=1}^{V'} n'_v + \sum_{w=1}^{W} n_w$, one of which must be larger than the other. Consider w.l.o.g, that the former is the larger and, on this basis, construct the following two collections of threshold types. In collection one, we perform n_v repetitions, from $v = 1$ to V, of the threshold type $t^{i_v}_{j_v}$, and n'_w repetitions, from $w = 1$ to W', of the threshold type $s^{i'_w}_{j'_w}$. In the second collection, we perform n'_v repetitions, from $v = 1$ to V', of the threshold type $t^{i'_v}_{j'_v}$; n_w repetitions, from $w = 1$ to W, of the threshold types $s^{i_w}_{j_w}$; and, finally, $\sum_{v=1}^{V} n_v + \sum_{w=1}^{W'} n'_w - \sum_{v=1}^{V'} n'_v - \sum_{w=1}^{W} n_w$ repetitions of the threshold type $t^{i^*}_{j^*}$. By construction, these two collections have the same number of components, all of which are threshold types. Moreover, given that $x_k < x'_k$, the sum of types is strictly smaller in the former collection than in the latter, and the application of D-CLA guarantees that the sum of cumulative log-odds is strictly smaller in the former collection than in the latter. The limit definition of H guarantees that $H(x) \leq H(x')$. Similarly, it is immediate to see that H is continuous and, given the definition of $t^{i^*}_{j^*}$, it is also obvious that $H(0) = .5$. Moreover, by assumption, there are threshold types $t^{i_1}_{j_1}$ and $t^{i_2}_{j_2}$ such that $\bar{\rho}(i_1, A_{j_1}) < \bar{\rho}(i^*, A_{j^*}) = .5 < \bar{\rho}(i_2, A_{j_2})$ and hence by taking into account the sequences of real numbers given by $\{t^{i_1}_{j_1} - t^{i^*}_{j^*}, 2(t^{i_1}_{j_1} - t^{i^*}_{j^*}), \ldots, k(t^{i_1}_{j_1} - t^{i^*}_{j^*}), \ldots\}$ and $\{t^{i_2}_{j_2} - t^{i^*}_{j^*}, 2(t^{i_2}_{j_2} - t^{i^*}_{j^*}), \ldots, k(t^{i_2}_{j_2} - t^{i^*}_{j^*}), \ldots\}$, it is obvious that H approaches 0 (respectively 1) when considering real values approaching $-\infty$ (respectively, ∞). It is also immediate to see that H satisfies property 4 as described in the proof of Theorem 4.

Consider now the following function defined over the positive reals:

$$O(x) = \frac{1 - H(x)}{H(x)}.$$

From the properties of H, it is immediate that $1 - O(x)$ must be a continuous CDF over the positive reals and that $O(x)O(z) = O(x + z)$ must hold for every pair of positive real values x and z. The same arguments used in the proof of Theorem 4 can be used to show the logistic nature of H, as desired. ■

Proposition 4 is the discrete version of Theorem 4. Note that the strategies followed in the corresponding proofs are rather different. In the case of continuous choice, choice data from a decision problem provide information on an interval of types that must be extended to the real line in a way that satisfies the additive requirements of CLA. Then, we use the intersection of the intervals of types across menus and the CLA property to guarantee that all these extensions follow the same logistic distribution. However, each discrete choice problem provides information only over a finite number of thresholds. Given the sparsity of these thresholds, we need to consider all of them together, and expand the information to the real line. We do this by using integer combinations of these thresholds, which, from classic results, are known to form a dense subset of the reals. In order to implement this strategy, we need a stronger additivity property, D-CLA, which operates not only over pairs of types but over two equally-sized collections of types. Once the extension to the reals is done, we can complete the proof using arguments from the proof of Theorem 4.[51]

Appendix C. Mixed Ordered Logit

It is well-known that RUMs can be approximated by mixed-logit models (see, McFadden and Train (2000)). We show here that the same kind of approximation can be established in the ORUM framework. Given Theorem 2 we only need to discuss the case in which the type-utility map γ is fixed. Denote by $\mathcal{G}^{ML}$ the set of all possible mixed-logistic distributions, i.e., the convex hull of $\mathcal{G}^{L}$.

Theorem 5. *Let γ be any type-utility map. If F is ORUM-rationalizable with type-utility map γ, there is a sequence $\{G^1, \dots, G^n, \dots\}$ with $G^n \in \mathcal{G}^{ML}$, such that $F^n \to F$, where F^n is ORUM-rationalizable with type-utility map γ and type distribution G^n.*

Proof of Theorem 5: Consider any F that is ORUM-rationalizable with type-utility map γ. We first show that for any utility function U in the image of γ, i.e., $\gamma(t) = U$ for some t, there exists a sequence of ordered logits producing choice distributions that converge to the deterministic choice function generated by U. Consider a sequence of ordered logits, all of them using map γ, in which the $n-$th element of the sequence has mean $\tau_n = t$ and standard deviation $\frac{1}{n}$. For every menu A_j, consider the element x_j such that $a_j(x_j) = a_j(U)$. It is immediate that for every $\epsilon > 0$, we can find a neighborhood of $a_j(x_j)$ such that the cumulative choice probability in that neighborhood approaches 1 whenever $n \to \infty$. Given finiteness of the domain of menus, this concludes the proof of the claim.

We now prove the main statement. Let G be the type distribution rationalizing F. For every U in the image of γ, consider the sequence of real-valued intervals $\{I_n\}$, where $I_n = [-n, n]$. Consider a sequence of finite subsets of real numbers K_n, where K_n is formed by those elements resulting from a partitioning of I_n into n^2 equal-size intervals. This sequence has the property of both expanding its support and becoming finer when n grows. For every value of n, consider the following mixed ordered logit: for every $t \in K_n$, use the ordered logit such that $\tau = t$ and $\sigma = \frac{1}{n}$. When n grows large, we know that this ordered logit approximates the choice frequencies of utility $\gamma(t)$. Denote by (t', t'') the interval formed by the elements in K_n right before and right after t. Given continuity, this ordered logit approximates well the choice frequencies generated by the truncated distribution over types (t', t''). To each sub-interval, assign the cumulative mass assigned that corresponds to γ and G. Given finiteness of the domain of menus, the mixed ordered logit approximates well the choice frequencies predicted by the truncation of the ORUM to interval I_n. The result follows. ■

[51] As in the continuous case, the discrete setting also has a unique pair of parameters (τ, σ) rationalizing the data.

Appendix D. Additional Tables

Menu	π_j^1	π_j^2	q_j^1	ϕ_j^1	ϕ_j^2	Menu	π_j^1	π_j^2	q_j^1	ϕ_j^1	ϕ_j^2
1	.024	.011	.696	.034	.035	11	.026	.066	.461	.057	.122
2	.025	.045	.393	.064	.074	12	.039	.089	.542	.073	.194
3	.043	.052	.500	.085	.103	13	.059	.056	.774	.076	.247
4	.068	.051	.644	.106	.144	14	.010	.071	.328	.031	.106
5	.054	.099	.425	.126	.172	15	.045	.088	.638	.070	.244
6	.075	.051	.669	.112	.154	16	.078	.097	.756	.104	.398
7	.091	.046	.752	.122	.186	17	.027	.076	.611	.044	.196
8	.090	.074	.664	.135	.219	18	.030	.075	.767	.040	.324
9	.045	.043	.639	.071	.120	19	.022	.096	.658	.034	.280
10	.073	.039	.764	.095	.166	20	.013	.081	.779	.017	.368

Table 11. Risk: Menus with states ordered by ϕ_j

Menu	ϕ_j	20 obs	80 obs	160 obs	480 obs
1	1.019	5	8.75	9.375	13.958
2	1.168	5	10.00	15.000	12.292
3	1.216	15	13.75	17.500	11.875
4	1.360	10	16.25	12.500	10.625
5	1.364	15	21.25	8.750	10.833
6	1.374	10	13.75	15.625	11.667
7	1.532	10	10.00	10.000	12.500
8	1.622	5	15.00	11.250	12.500
9	1.699	20	11.25	14.375	13.750
10	1.743	25	6.25	9.375	11.875
11	2.151	10	15.00	13.125	13.125
12	2.672	0	16.25	9.375	12.292
13	3.241	25	11.25	10.000	13.125
14	3.445	15	8.75	12.500	11.458
15	3.457	5	12.50	10.625	11.042
16	3.832	20	10.00	13.125	12.500
17	4.451	30	16.25	11.250	12.500
18	8.178	10	12.50	11.250	12.500
19	8.343	30	10.00	8.750	9.792
20	21.841	15	8.75	15.000	11.042

Table 12. Risk: Share of observations with $r = 0$

Menu	ϕ_j	20 obs	80 obs	160 obs	480 obs
1	1.019	.989	.990	.990	.992
2	1.168	.906	.917	.925	.946
3	1.216	.905	.895	.924	.925
4	1.360	.777	.857	.883	.888
5	1.364	.835	.880	.894	.880
6	1.374	.836	.879	.837	.891
7	1.532	.784	.842	.843	.832
8	1.622	.684	.767	.810	.863
9	1.699	.733	.765	.783	.810
10	1.743	.721	.818	.802	.822
11	2.151	.648	.604	.702	.767
12	2.672	.496	.577	.684	.652
13	3.241	.463	.531	.615	.632
14	3.445	.505	.610	.567	.713
15	3.457	.537	.582	.541	.651
16	3.832	.343	.535	.561	.592
17	4.451	.433	.494	.581	.547
18	8.178	.203	.379	.406	.488
19	8.343	.418	.431	.361	.387
20	21.841	.102	.248	.264	.270

Table 13. Risk: Largest r across menus

Menu	π_j^1	π_j^2	π_j	Menu	π_j^1	π_j^2	π_j
1	7.064	.010	.071	11	.958	.045	.043
2	6.208	.013	.081	12	.824	.052	.043
3	2.512	.026	.066	13	.821	.090	.074
4	1.977	.046	.091	14	.679	.075	.051
5	1.963	.045	.088	15	.541	.099	.054
6	1.856	.039	.073	16	.444	.024	.011
7	1.805	.025	.045	17	.444	.089	.039
8	1.331	.051	.068	18	.402	.075	.030
9	1.234	.078	.097	19	.353	.076	.027
10	1.056	.056	.059	20	.231	.096	.022

Table 14. Altruism: Menus

Menu	π_j	20 obs	80 obs	160 obs	480 obs
1	7.064	25	26.25	23.750	24.375
2	6.208	20	21.25	25.000	26.042
3	2.512	35	21.25	24.375	25.417
4	1.977	25	28.75	27.500	23.750
5	1.963	30	20.00	22.500	28.542
6	1.856	15	28.75	28.125	30.000
7	1.805	25	21.25	24.375	27.917
8	1.331	15	33.75	23.750	29.167
9	1.234	10	21.25	28.125	26.875
10	1.056	35	23.75	22.500	26.458
11	.958	25	27.50	23.750	26.250
12	.824	15	28.75	25.000	24.583
13	.821	30	26.25	26.875	28.125
14	.679	25	22.50	30.625	29.792
15	.541	35	25.00	28.125	23.958
16	.444	25	22.50	28.750	25.000
17	.444	25	31.25	31.250	26.042
18	.402	10	25.00	30.000	28.125
19	.353	20	23.75	30.625	26.458
20	.231	25	21.25	26.875	26.875

TABLE 15. Altruism: Share of observations with $v_j = 0$

Menu	ϕ_j	20 obs	80 obs	160 obs	480 obs
1	7.064	.045	.045	.045	.058
2	6.208	.037	.056	.072	.077
3	2.512	.154	.254	.255	.421
4	1.977	.294	.394	.415	.499
5	1.963	.168	.359	.310	.566
6	1.856	.310	.294	.330	.585
7	1.805	.213	.464	.613	.629
8	1.331	.528	.510	.574	.649
9	1.234	.616	.499	.570	.794
10	1.056	.617	.648	.860	.804
11	.958	.486	.701	.737	.888
12	.824	.772	.555	.862	.889
13	.821	.640	.971	.796	.894
14	.679	.949	.859	.993	.965
15	.541	.607	.978	.985	.995
16	.444	.935	.969	.994	.999
17	.444	.947	.990	.989	1.000
18	.402	.998	.995	1.000	1.000
19	.353	.994	1.000	1.000	1.000
20	.231	.996	1.000	1.000	1.000

TABLE 16. Altruism: Largest v_j across menus

REFERENCES

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

[2] Ahn, D. S., F. Echenique, and K. Saito (2018), “On path independent stochastic choice,” *Theoretical Economics*, 13(1), 61-85.

[3] Alós-Ferrer, C., E. Fehr, and N. Netzer (2021), “Time will tell: Recovering preferences when choices are noisy,” *Journal of Political Economy*, 129(6), 1828-1877.

[4] Andreoni, J. and J. Miller (2002), “Giving according to GARP: An experimental test of the consistency of preferences for altruism,” *Econometrica*, 70(2), 737-753.

[5] Apesteguia, J. and M.A. Ballester (2021), “Separating Predicted Randomness from Residual Behavior,” *Journal of the European Economic Association*, 19(2), 1041-1076.

[6] Apesteguia J, M.A. Ballester, and J. Lu (2017), “Single-Crossing Random Utility Models,” *Econometrica*, 85(2), 661-674.

[7] Apesteguia, J., and M.A. Ballester (2023), “Random utility models with ordered types and domains,” *Journal of Economic Theory*, 105674.

[8] Apesteguia J, M.A. Ballester, and A. Gutierrez-Daza (2024), “Random Discounted Expected Utility,” *Review of Economics and Statistics*, forthcoming.

[9] Bailey M., R. Cao, T. Kuchler, and J. Stroebel (2018), “The economic effects of social networks: Evidence from the housing market,” *Journal of Political Economy*, 126(6), 2224-2276.

[10] Bang-Jensen, J., and G. Z. Gutin (2008), *Digraphs: theory, algorithms and applications*. Springer Science and Business Media.

[11] Barseghyan, L., F. Molinari, and M. Thirkettle (2021), “Discrete Choice under Risk with Limited Consideration,” *American Economic Review*, 111(6), 1972-2006.

[12] Barsky, R. B., Juster, F. T., Kimball, M. S., and M.D. Shapiro (1997), “Preference parameters and behavioral heterogeneity: An experimental approach in the health and retirement study,” *The Quarterly Journal of Economics*, 112(2), 537-579.

[13] Besley T. and T. Persson (2011), “The Logic of Political Violence,” *The Quarterly Journal of Economics*, 126(3), 1411-1445.

[14] Blau, D. M. and A. P. Hagy (1998), "The demand for quality in child care," *Journal of Political Economy*, 106(1), 104-146.

[15] Block H.D. and J. Marschak (1960), "Random Orderings and Stochastic Theories of Response," In *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, edited by Olkin I. et al., Stanford University Press, pp. 97-132.

[16] Brady R.L. and J. Rehbeck (2016), "Menu-Dependent Stochastic Feasibility," *Econometrica*, 84(3), 1203-1223.

[17] Campante F. and D. Yanagizawa-Drott (2015), "Does Religion Affect Economic Growth and Happiness? Evidence from Ramadan," *The Quarterly Journal of Economics*, 130(2), 615-658.

[18] Caplin A. and M. Dean (2015), "Revealed Preference, Rational Inattention, and Costly Information Acquisition," *American Economic Review*, 105(7), 2183-2203.

[19] Carlana M. (2019), "Implicit Stereotypes: Evidence from Teachers' Gender Bias," *The Quarterly Journal of Economics*, 134(3), 1163-1224.

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

[21] Cauchy, A.L. (1821), *Cours d'Analyse de l'École Royale Polytechnique, I.re Partie. Analyse Algébrique*, L'Imprimerie Royale, Debure frères, Libraires du Roi et de la Bibliothèque du Roi.

[22] Cerreia-Vioglio, S., D. Dillenberger, P. Ortoleva, and G. Riella (2019), "Deliberately Stochastic," *American Economic Review*, 109 (7), 2425-45.

[23] Choi, S., S. Kariv, W. Müller, and D. Silverman (2014), Who is (more) rational?, *American Economic Review*, 104(6), 1518-1550.

[24] Cohen, A. and L. Einav (2007), "Estimating risk preferences from deductible choice," *American Economic Review*, 97(3), 745-788.

[25] Cummings, J.N. (2004), "Work Groups, Structural Diversity, and Knowledge Sharing in a Global Organization," *Management Science*, 50(3), 352-364.

[26] D'Agostino, R. B. and M. A. Stephens (1986), *Goodness-of-Fit-Techniques*. Marcel Dekker, Inc.

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

[28] Filiz-Ozbay, E. and Y. Masatlioglu (2023), "Progressive random choice," *Journal of Political Economy*, 131(3), 716-750.

[29] Frick, M., R. Iijima, and T. Strzalecki (2015), "Dynamic Random Utility," *Econometrica*, 87(6), 1941-2002.

[30] Fudenberg D., R. Iijima, and T. Strzalecki (2015), "Stochastic Choice and Revealed Perturbed Utility," *Econometrica*, 83(6), 2371-2409.

[31] Galambos, J and S. Kotz (1978), *Characterizations of Probability Distributions*, Lecture Notes in Mathematics vol. 675.

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

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

[34] He, J. and P. Natenzon (2023), "Moderate Utility," *American Economic Review: Insights*, forthcoming.

[35] Hoderlein, S. and J. Stoye (2014), "Revealed preferences in a heterogeneous population," *Review of Economics and Statistics*, 96(2), 197-213.

[36] Hoderlein, S. and J. Stoye (2015), "Testing stochastic rationality and predicting stochastic demand: the case of two goods," *Economic Theory Bulletin*, 3(2), 313-328.

[37] Kaplan, S.N. and L. Zingales (1997), "Do Investment-Cash Flow Sensitivities Provide Useful Measures of Financing Constraints?," *The Quarterly Journal of Economics*, 112(1), 169-215.

[38] Kashaev, N., V.H. Aguiar, M. Plávala, and C. Gauthier (2023), "Dynamic and stochastic rational behavior," Mimeo.

[39] Kitamura, Y. and J. Stoye, (2018), "Nonparametric Analysis of Random Utility Models," *Econometrica*, 86(6), 1883-1909.

[40] Kovach, M. and G. Tserenjigmid (2022), "Behavioral Foundations of Nested Stochastic Choice and Nested Logit," *Journal of Political Economy*, 130(9), 2411-2461.

[41] Kubler, F., L. Selden, and X. Wei (2014), "Asset Demand Based Tests of Expected Utility Maximization," *American Economic Review*, 104 (11), 3459-3480.

[42] Lu, J. and K. Saito (2018), "Random Intertemporal Choice," *Journal of Economic Theory*, 177, 780-815.

[43] Luce, R.D. (1959), *Individual Choice Behavior: A Theoretical Analysis*, New York, NY: Wiley.

[44] McFadden, D. and M.K. Richter (1990), "Stochastic Rationality and Revealed Stochastic Preference," In *Preferences, Uncertainty, and Optimality, Essays in Honor of Leo Hurwicz*, Westview Press: Boulder, CO, 161-186.

[45] McFadden, D. and K. Train (2000), "Mixed MNL models for discrete response," *Journal of applied Econometrics*, 15(5), 447-470

[46] Manzini P. and M. Mariotti (2014), "Stochastic Choice and Consideration Sets," *Econometrica*, 82(3), 1153-1176.

[47] Masatlioglu Y. and T.P. Vu (2023), "Growing Attention," Mimeo.

[48] Matzkin, R.L (2003) "Nonparametric estimation of nonadditive random functions," *Econometrica*, 71(5), 1339-1375.

[49] Matejka F. and A. McKay (2015), "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *American Economic Review*, 105(1), 272-98.

[50] Matzkin, R.L and M.K. Richter (1991), "Testing Strictly Concave Rationality," *Journal of Economic Theory*, 53(2), 287-303.

[51] Moulin, H. (1984), "Generalized Condorcet-Winners for Single Peaked and Single-Plateau Preferences," *Social Choice and Welfare*, 1(2), 127-147.

[52] Natenzon P. (2019), "Random Choice and Learning," *Journal of Political Economy*, 127(1), 419-457.

[53] Petri, H. (2023), "Binary single-crossing random utility models," *Games and Economic Behavior*, 138, 311-320.

[54] Petri, H. (2024), "Characterizing ordered models of behavioral heterogeneity," Mimeo.

[55] Rose, H. (1958), "Consistency of Preference: the Two-Commodity Case," *Review of Economic Studies*, 25(2), 124-125.

[56] Salzmann, H., T. Grundhöfer, H. Hähl, and R. Löwen (2007), *The Classical Fields: Structural Features of the Real and Rational Numbers*, Cambridge: Cambridge University Press.

[57] Small, K. A. (1987), "A discrete choice model for ordered alternatives," *Econometrica*, 409-424.

[58] Train, K. E. (2009), *Discrete choice methods with simulation*, Cambridge university press.

[59] Turansick, C. (2022), "Identification in the random utility model," *Journal of Economic Theory*, 203, 105489.

[60] Valkanova, K. (2022), "Inferring Single-Peakedness from Random Choice," Mimeo.

[61] Yildiz, K. (2024), "Self-progressive choice models," Mimeo.