CHOICE-BASED FOUNDATIONS OF ORDERED LOGIT

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

Abstract. We provide revealed preference foundations to ordered logit for continuous decision problems.

Keywords: Ordered logit, Cumulative logit, Revealed preference.
JEL classification numbers: C00; D00.

1. Introduction

The ordered-logit model has attracted considerable attention across various fields including economics, political science, sociology or biology (see, e.g., the textbook treatments of Johnson and Albert (1999), O'Connell (2006), Agresti (2010) and Greene and Hensher (2010)). Within the field of economics, it has been used in areas as diverse as political economy, finance, labor, welfare, management, gender and networks (see, e.g., Besley and Persson (2011), Kaplan and Zingales (1997), Blau and Hagy (1998), Campante and Yanagizawa-Drott (2015), Cummings, (2004), Carlana (2019) and Bailey, Cao, Kuchler and Stroebel (2018)). Despite this widespread interest, the literature has not yet provided choice-based foundations for the ordered-logit model. The objective of this paper is to cover this gap.

Given that the ordered-logit model involves choices from one-dimensional decision problems, we consider a setting in which each problem is a line segment. This setting is flexible enough to study choices in a variety of economic environments such as political domains, consumption settings, state-contingent lotteries, inter-temporal

Date: November, 2022.

* Financial support by FEDER/Ministerio de Ciencia de Innovación (Agencia Estatal de Investigación) through Grant PGC2018-098949-B-I00 and through the Severo Ochoa Programme for Centres of Excellence in R&D (Barcelona School of Economics CEX2019-000915-S), and Balliol College is gratefully acknowledged.

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

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

payoffs, or payoff distributions. We consider arbitrary collections of decision problems and describe choice data in terms of the cumulative distribution function in each of the observed problems. We study the conditions for ordered-logit rationalizability, i.e., the existence of an ordered collection of utilities generating ordered choices, and a logistic distribution over these utilities, rationalizing the choice data.

The first result of the paper establishes that the conditions for ordered-logit rationalizability can be conveniently derived from those for deterministic rationalizability. Suppose that, for each probability value $p \in (0,1)$, we construct the hypothetical choice function c^p that identifies, for each decision problem, the first alternative attaining cumulative choice mass p. Then, Theorem 1 shows that data is generated by an ordered logit if and only if every c^p is rationalizable. An advantage of the structure of Theorem 1 is that it makes it portable to the analysis of any concrete economic environment within our setting. Foundations for the ordered-logit model in specific environments can be derived from existing, or perhaps relatively easy to obtain, deterministic rationalization results for these environments. Corollaries 1 to 3 illustrate this point for the cases of political choices with single-peaked utilities, consumption choices with strictly quasi-concave and monotone utilities, and lottery choices with expected utilities, respectively.

Importantly, the proof of Theorem 1 shows that the logistic assumption is indeed devoid of empirical content; whenever data is ordered-logit rationalizable, it can be equivalently explained, with appropriate relabelling of the same utilities, by means of other statistical distributions. Thus, Theorem 1 provides, indeed, foundations for ordered, non-parametric, stochastic choice models.

In empirical work, it is common practice to: (i) select a uni-dimensional class of utilities, and (ii) fix upfront a specific parametrization of these utilities. As an example, the analyst may try to understand choices over lotteries by using CRRA expected utilities, instead of any expected utility, and proceed to represent them throughout the textbook parametrization $\sum_i q_i \frac{x_i^{1-t}}{1-t}$, where parameter t captures relative risk aversion. Note that (i) simply restricts the class of utilities under consideration but leaves the nature of the problem unaltered. In the CRRA example, one may work out a parallel result to Corollary 3 using, instead of a deterministic characterization of expected utility, one of CRRA expected utility. Innocuous as it may seem, (ii) significantly affects the nature of the problem by precluding any relabelling of the utilities and paving the way for the logistic assumption to impose significant structure on choice data. In our example using CRRA expected utility, the chosen parametrization imposes that the logistic distribution must specifically operate over the risk aversion coefficient, which further shapes lottery choices. Given that assuming particular parametrizations have implications for the ordered-logit model, in Section 3 we study ordered-logit rationalizability for a given, yet generic, parametrization of utilities.

Theorem 2 in Section 3 shows that such an ordered-logit rationalization requires data to satisfy two properties, Corner Extremeness and Cumulative Logit Additivity. Corner Extremeness imposes that the corner alternatives in each decision problem can receive non-null choice probability if and only if they are maximizers for some of the utilities in the assumed parametrization. The property results from the fact that the ordered-logit model is a random utility model. Cumulative Logit Additivity uses the well-known cumulative logit notion, i.e., the log-ratio of masses below and above a given alternative, and states that equal sums of parameters must lead to equal sums of cumulative logits. The property arises from the fact that, in the ordered-logit model, the cumulative logit of any alternative coincides with the normalized (utility) parameter that selects it. As with Theorem 1, we illustrate how Theorem 2 may read in concrete economic environments through Corollaries 4 and 5, that cover the cases of political choices with Euclidean utilities, and lottery choices with CRRA expected utilities, respectively.

We close this introduction with a brief comment on the links of this paper with other strands of literature. First, the paper contributes to the study of choice-based foundations of various stochastic choice models. The classic works are those of Luce (1959) and Block and Marshak (1960).[1] Second, it is also worth mentioning its connection to recent papers seeking to bridge the gap between the choice-based foundations and the econometric implementation of stochastic models, such as Dardanoni, Manzini, Mariotti and Tyson (2020), Aguiar and Kashaev (2021), Barseghyan, Molinari, and Thirkettle (2021), and Apesteguia and Ballester (2021).

[1]Other recent contributions are 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), Apesteguia, Ballester and Lu (2017), Cerreia-Vioglio, Dillenberger, Riella and Ortoleva (2019), Frick, Iijima, and Strzalecki (2019), Natenzon (2019) or Cattaneo, Ma, Masatlioglu and Suleymanov (2020).

2. Ordered logit

Ordered logit involves choices from decision problems in which alternatives are ordered. We propose a setting to reflect this property while remaining general enough to cover a variety of economic applications, as we discuss below. Let $X \subseteq \mathbb{R}^K$ be a convex space of alternatives. Decision problems, also called menus of alternatives, are ordered line segments of X. Formally, there is a collection of decision problems $\{A_j\}_{j=1}^J$, each of them consisting of two distinct extreme alternatives and their convex combinations, i.e., $A_j = \{(1-a)\underline{x}_j + a\overline{x}_j : \underline{x}_j, \overline{x}_j \in X \text{ and } a \in [0,1]\}$. Any alternative $x \in A_j$ is, therefore, determined by its relative position in the line segment, i.e. by the unique value $a(x,j) \in [0,1]$ such that $x = (1-a(x,j))\underline{x}_j + a(x,j)\overline{x}_j$. The following are five simple examples that fall within this setting.

(1) *Political domain.* $X = \mathbb{R}$ represents the space of all possible policies, and each decision problem $A_j = [\underline{x}_j, \overline{x}_j]$ is a closed interval describing the feasible policies at a given situation.
(2) *Consumption setting.* $X = \mathbb{R}^2_+$ is the set of two-dimensional bundles, and each decision problem is the frontier of a linear budget set, i.e. $A_j = \{x \in X : x_1 + p_j x_2 = m_j\}$ with p_j denoting the (normalized) price of the second good and m_j representing (normalized) income, both parameters being strictly positive real values. Using the line segment notation, we can alternatively write A_j as the set of convex combinations of $\underline{x}_j = (m_j, 0)$ and $\overline{x}_j = (0, \frac{m_j}{p_j})$, ordering bundles by their amount of the second good.
(3) *Monetary allocations in dictator games.* $X = \mathbb{R}^2_+$ is the set of monetary allocations between a dictator and another person, with the amount received by the dictator represented in the first component of the vector. Decision problems can be described exactly as in the consumption setting.
(4) *State-contingent lotteries.* $X = \mathbb{R}^2_+ \times [0,1]$ is the set of two-state monetary lotteries. A decision problem is determined by setting a value $q \in [0,1]$ describing the probability of the first state, with monetary prizes x_1 and x_2 being subject to a budget constraint as in the previous examples.[2]
(5) *Inter-temporal allocations.* $X = \mathbb{R}^4_+$ is the set of two-period monetary streams. A decision problem is defined by fixing the time horizons t_1 and t_2 in which the

[2]Without loss of generality, assume that the second state is the one that pays less in expectation, i.e. $p_j > \frac{1-q}{q}$. This allows to interpret higher levels of x_2 as more risk aversion.

monetary payoffs x_1 and x_2 are awarded, where these payments are subject to a budget constraint as in the previous examples.

We start with a brief presentation of the classical, deterministic, notion of choice rationalization, that will later prove useful for our analysis. In the deterministic case, data is described by the observed chosen alternative in each decision problem, i.e., $c = \{c_j\}_{j=1}^J$, with $c_j \in A_j$. We say that c is rationalizable if there exists a utility function $U : X \to \mathbb{R}$ such that, in every menu, the observed choice is the alternative maximizing utility U. The utility function is usually assumed to have some basic structure in order to guarantee that maximizers are unique and, depending on the setting of interest, some further structure is also imposed. For example, in the political domain, utilities are frequently assumed to be single-peaked; in the consumption setting, utilities are commonly assumed to be strictly quasi-concave and strictly monotone in each of the goods; in the state-contingent lotteries setting, utilities are usually expected utilities. We denote by $\mathcal{U}$ the set of utility functions under consideration.

We now move to the stochastic setting and start with the description of choice data. In this case, data corresponds to the observed distribution of choices in each decision problem, i.e., $F = \{F_j\}_{j=1}^J$ where $F_j : A_j \to [0,1]$ is the cumulative distribution function (CDF) of choices in decision problem j. That is, the value $F_j(x) \in [0,1]$ represents the probability of choosing any alternative $y \in A_j$ such that $a(y,j) \leq a(x,j)$. As a CDF, F_j is non-decreasing with $F_j(\overline{x}_j) = 1$, and assumed to be continuous over $A_j \setminus \{\underline{x}_j, \overline{x}_j\}$.

We now present the notion of ordered-logit rationalization. Informally, we say that F is ordered-logit rationalizable if there exists an ordered collection of utilities, and a logistic distribution over these utilities, such that, for every decision problem, (i) ordered choices: higher utilities produce higher choices and (ii) logit rationalization: the CDF of observed choices at any given alternative is equal to the mass of utilities, according to the logistic, that maximize below the given alternative. Formally, in an ordered-logit rationalization, there is a set of types $\mathbb{R}$, a continuous mapping $\gamma : \mathbb{R} \to \mathcal{U}$ assigning a utility to each type, and a logistic distribution with location and scale parameters τ and σ, respectively, defined over the types. Denoting by x_j^t the unique maximizer of type t in decision problem j, part (i) requires that $t \leq t'$ implies $a(x_j^t, j) \leq a(x_j^{t'}, j)$, and part (ii) requires that $F_j(x)$ is the logistic mass, given parameters (τ, σ), of all types for which $a(x_j^t, j) \leq a(x, j)$.

An important advantage of the ordered-logit model is that the generated choice probabilities have a simple closed-form expression. Consider first the case of an alternative $x \in A_j \setminus \{\overline{x}_j\}$. Denote by t^x_j the supremum type with a maximizer in decision problem j that lies below x.[3] In the ordered-logit model, the cumulative choice probability at x in decision problem j corresponds to the logistic mass, given parameters (τ, σ), at type t^x_j, which is equal to $\frac{1}{1+e^{-(t^x_j-\tau)/\sigma}}$. Whenever the extreme alternative $\overline{x}_j$ has strictly positive choice mass, there must be a utility, indexed by $\overline{t}_j$, above which $\overline{x}_j$ is invariably the maximal alternative. Hence the probability of choosing alternative $\overline{x}_j$ is given by $1 - \frac{1}{1+e^{-(\overline{t}_j-\tau)/\sigma}}$.

The interpretation of the logistic parameters is straightforward. Parameter τ is a location parameter, and whenever x^τ_j is a non-extreme alternative, the ordered-logit cumulative mass at x^τ_j is .5.[4] Parameter σ is a scale parameter capturing the variability of the distribution around the location. When σ goes to zero, the model resembles one for deterministic rational choice, because the mass of any interval containing type τ, and hence its utility, approaches 1, and the mass of any interval not containing it approaches zero. When σ goes to ∞, the density of type τ becomes closer to that of any other type within any given interval.

2.1. **A characterization of ordered logit.** We now provide a necessary and sufficient condition for data to admit an ordered-logit rationalization. The property builds upon the deterministic notion of rationalization. Set any value $p \in (0, 1)$. For each decision problem A_j, denote by c^p_j the first alternative that attains cumulative choice mass p, i.e.,

$$c^p_j = \{x \in A_j : F_j(x) \geq p \text{ and } a(y, j) < a(x, j) \Rightarrow F_j(y) < p\}.$$

Denote by $c^p = \{c^p_j\}^J_{j=1}$ the collection of alternatives defined in such a way.

In Theorem 1, we show that F admits an ordered-logit rationalization if and only if c^p is rationalizable for every p. The intuition for this result is as follows. Thanks to the rationalizability of each c^p, we can find a utility function rationalizing these hypothetical choices and guarantee that a rationalizing map $\beta : (0, 1) \to \mathcal{U}$ exists. Moreover, thanks to the definition of c^p and the non-decreasing nature of each CDF,

[3] If there is no such type, set $t^x_j = -\infty$. In many applications, a bijection between types and choices exists and, whenever this happens, type t^x_j is merely the inverse of alternative x^t_j.

[4] If $x^\tau_j \in \{\underline{x}_j, \overline{x}_j\}$, the ordered-logit mass of alternative x^τ_j must be greater than .5.

this map must generate ordered choices in each menu. Stochastic rationalization of each choice distribution can be obtained by using the uniform distribution over $(0, 1)$, i.e., by randomizing uniformly over the utilities $\beta(p)$. In order to construct an ordered-logit rationalization, we merely need to use the same utilities after some relabeling, building the appropriate bijection between $(0, 1)$ and the set of types $\mathbb{R}$ to guarantee that the uniform distribution over $(0, 1)$ becomes a logistic distribution over $\mathbb{R}$.

Theorem 1. *F admits an ordered-logit rationalization if and only if c^p is rationalizable for every p.*

Proof of Theorem 1: We start by proving the necessity part. Suppose that F admits an order-logit rationalization with map γ and logistic parameters (τ, σ). Fix any $p \in (0, 1)$. Let $t(p) \in \mathbb{R}$ be the first type satisfying the equation $\frac{t(p)-\tau}{\sigma} = \log \frac{p}{1-p}$, i.e., the first type reaching cumulative mass p for the given logistic distribution. We claim that type $t(p)$ produces maximizers across menus that coincide with c^p, proving the rationalizability of c^p. Consider any decision problem j. Given the definition of c^p_j and the ordered choice structure, all types below $t(p)$ have maximizer below c^p_j. The definition of c^p_j and the continuity assumption guarantee that $x^{t(p)}_j = c^p_j$, and the claim follows.

We now prove the sufficiency part. Since c^p is rationalizable for every $p \in (0, 1)$, and given the continuity of F, there exists a continuous mapping $\beta : (0, 1) \to \mathcal{U}$ such that the utility of type p rationalizes c^p. Consider then any menu A_j. We first prove that maximizers are non-decreasing in p which, given rationalizability, is equivalent to prove that choices c^p_j are non-decreasing in p. This follows immediately from the definition of such alternatives and the fact that F_j is non-decreasing.

We then consider the uniform distribution on $(0, 1)$ and claim that, for every decision problem j and every $x \in A_j$ such that $0 < F_j(x) < 1$, $F_j(x)$ coincides with the mass of utilities with a maximizer below x. Given the non-decreasing nature of the maximizing alternatives, we need to prove that the utility function $\beta(F_j(x))$ is the last utility with maximizer below x. First, consider $p > F_j(x)$. Since x has not reached cumulative probability p, it must be that $a(c^p_j, j) > a(x, j)$, and since utility $\beta(p)$ rationalizes c^p_j, the maximizer of $\beta(p)$ lies strictly above x. Second, consider the utility function $\beta(F_j(x))$. By construction, $c^{F_j(x)}_j$ lies below x and since $\beta(F_j(x))$ rationalizes $c^{F_j(x)}_j$, the maximizer of $\beta(F_j(x))$ lies below x.

Finally, we construct the order-logit rationalization as follows. For every type $t \in \mathbb{R}$, let $\gamma(t)$ be the utility $\beta(\frac{1}{1+e^{-t}})$. Given the continuous bijection performed over the same collection of utilities, the map γ is continuous and trivially produces non-decreasing maximal alternatives. We can then define a logistic distribution over the reals, with location parameter equal to zero and scale parameter equal to one. This produces exactly the same choice probabilities as the original uniform distribution over $(0, 1)$ and, hence, the ordered logit rationalizes the cumulative mass of any alternative x such that $0 < F_j(x) < 1$. Finally, if x is such that $F(x) = 0$ (respectively $F(x) = 1$), it is straightforward that no utility can produce a maximizer below x (respectively, all utilities produce a maximizer below x), and the mass of these utilities is zero (respectively, one). This concludes the proof. ■

Importantly, the proof of Theorem 1 has been constructed to transparently illustrate that the assumption of the logistic distribution has no particular empirical content per se. We have shown that, when every c^p is rationalizable, data could be rationalized by using other alternative probability distribution over the same set of utilities, by simply relabelling them appropriately. That is, without the imposition of further restrictions, the result is distribution-free, and thus, Theorem 1 gives foundations not only to the ordered-logit model but also to non-parametric stochastic ordered choice. In Section 3, we return to this issue and study the implications of the common, applied, case in which the analyst fixes upfront a specific parametrization of a class of utilities and wonder whether a logistic distribution over the assumed parametrization may explain the data. In this case, as we will see, the logistic assumption has bite.

Before this, we want to point out that an advantage of the structure of Theorem 1 is that it is portable to the analysis of specific economic settings of interest. All that is needed is to embed the known, or relatively easy to obtain, deterministic rationalization conditions of interest. We illustrate now with some examples.

2.2. **Characterizations of ordered logit in specific settings.** With the purpose of showing how Theorem 1 may read in specific economic settings, we present three immediate corollaries. The first one deals with decisions on a political domain and uses the usual single-peaked utilities.[5] We simply need to write $I_j^c = c(A_j)$ whenever

[5]A utility U is single-peaked if there exists an alternative $y \in \mathbb{R}$, the peak of the utility, such that for every pair satisfying $x' < x \leq y$ or $y \leq x < x'$, we have $U(x) > U(x')$. Remark 1 in Moulin (1984)

$c(A_j) \neq \{\underline{x}_j, \overline{x}_j\}$, $I_j^c = (-\infty, \underline{x}_j]$ whenever $c(A_j) = \underline{x}_j$ and $I_j^c = [\overline{x}_j, +\infty)$ whenever $c(A_j) = \overline{x}_j$. We say that c satisfies the intersection property whenever $I_j^c \cap I_{j'}^c \neq \emptyset$ holds for every pair of decision problems.

Corollary 1. *In the political domain, F admits an ordered-logit rationalization with single-peaked utilities if and only if c^p satisfies the intersection property for every p.*

Proof of Corollary 1: We start by proving that a choice function c defined over J decision problems has the intersection property if and only if it can be rationalized by a single-peaked utility function. For the necessity part, suppose that all choices are generated by a single-peaked utility function, with peak in $y \in \mathbb{R}$. Notice that for every decision problem j, I_j^c must include the peak alternative y: either because y belongs to A_j and is thus chosen, or because y does not belong to A_j and the chosen alternative is the closest extreme to y in A_j. Thus, the intersection property trivially holds. For the sufficiency part, consider the following two cases. First, there exists one menu A_j such that $c(A_j) \neq \{\underline{x}_j, \overline{x}_j\}$. Denote this chosen element by y and construct the Euclidean utility with peak at y, $-(x-y)^2$. Euclidean utilities are trivially single-peaked utilities. We show that this utility rationalizes all choices. It trivially rationalizes decision problem j. For any other j', it can be that $y \in A_{j'}$ or $y \notin A_{j'}$. In the former case, the intersection property requires that $c(A_{j'}) = y$, and rationalization holds. In the latter case, the intersection property requires that $c(A_{j'})$ is the extreme alternative closest to y, and rationalization holds. Second, suppose that extreme alternatives are selected in all decision problems. By the intersection property, it must be the case that whenever $c(A_j) = \underline{x}_j$ and $c(A_{j'}) = \overline{x}_{j'}$, it is $\underline{x}_j \geq \overline{x}_{j'}$. Hence, the largest right-extreme choice, denoted y_1, must be lower than the smallest left-extreme choice, denoted y_2. One can set the peak of the Euclidean utility to be any alternative in $[y_1, y_2]$ and this utility rationalizes all choices. The application of Theorem 1 concludes the proof. ■

presents a deterministic characterization of single-peaked rationalizability based upon the classical property of independence of irrelevant alternatives and a continuity requirement (see also Bossert and Peters (2009), who analyze further the problem). Given that their results require the choice function to be observed over all closed intervals, and not over a finite number of them as in our setting, we cannot borrow directly from their properties and need to provide our own property.

We now present a corollary for the consumption setting based upon the classical Weak Axiom of Revealed Preference (WARP).[6]

Corollary 2. *In the consumption setting, F admits an ordered-logit rationalization with utilities that are strictly quasi-concave and strictly monotone in both goods if and only if every c^p satisfies WARP.*

Proof of Corollary 2: The necessity of WARP for deterministic rationalizability is well-known. Similarly, sufficiency simply requires to use the classical argument developed by Rose (1958) for the two-dimensional consumption setting and conclude that a choice function satisfying WARP also satisfies the Strong Axiom of Revealed Preference. As it is well-known, this allows to construct a utility function that rationalizes all choices and, in particular, this can be selected to be strictly quasi-concave and strictly monotone in both goods (see, e.g., Matzkin and Richter (1991)). The application of Theorem 1 concludes the proof. ■

We conclude with the analysis of the state-contingent lotteries settings. Consider a choice function c defined over J menus, each of them determined by the corresponding parameters q_j, p_j, m_j, i.e. $A_j = \{(x_1, x_2; q_j) : x_1 + p_j x_2 = m_j\}$. Our result uses the deterministic rationalization of expected utilities using SAREU.[7]

Corollary 3. *In the state-contingent lotteries setting, F admits an ordered-logit rationalization where all expected utilities are either risk loving, risk neutral or risk averse if and only if every c^p is either equal to $(\underline{x}_1, \ldots, \underline{x}_J)$ or it satisfies SAREU.*

Proof of Corollary 3: It is immediate to see that the selection of $\underline{x}_j$ in every decision problem j is necessary for the case of risk loving and risk neutral expected utilities. Sufficiency follows from the fact that these corner solutions can be rationalized by any such expected utility. Finally, Kubler, Selden and Wei (2014) show that SAREU

[6]In the consumption setting, we define B_j as the budget set with frontier A_j. Then, c satisfies WARP if $c(A_j) \in B_{j'}$ and $c(A_{j'}) \in B_j$ implies $c(A_j) = c(A_{j'})$.

[7]In our context, the property of SAREU by Kubler, Selden and Wei (2014) reads as follows. For any decision problem j, denote the ratio of prices-to-probabilities in each state by $\rho_j^1 = \frac{1}{q_j}$ and $\rho_j^2 = \frac{p_j}{1-q_j}$. We say that c satisfies SAREU whenever it always selects non-extreme alternatives $c_j = (c_{j1}, c_{j2}) > (0,0)$ and, for every sequence of decision problems $j_1, \ldots, j_K$, it is $\prod_{k=1}^{K-1} L(j_k, j_{k+1}) < 1$, where $L(j_k, j_{k+1}) = \max_{\{s,s':c_{js} > c_{j's'}\}} \frac{\rho_j^s}{\rho_{j'}^{s'}}$.

is a necessary and sufficient condition for strictly risk averse expected utilities. The application of Theorem 1 concludes the proof. ■

3. Ordered logit for a specific parametrization

In empirical work, it is common practice to: (i) select a uni-dimensional class of utilities $\mathcal{U}^* \subseteq \mathcal{U}$, and (ii) fix upfront a specific parametrization γ over $\mathcal{U}^*$. Part (i) restricts the class of utilities of interest. For example, the analyst may consider the class of Euclidean utilities in the political domain, the classes of Cobb-Douglas or CES utilities in the consumption setting, or the classes of CRRA or CARA expected utilities in the state-contingent lotteries setting. These restrictions have no major consequences in terms of the lessons learnt in Theorem 1. If we want to know whether the data admits an ordered-logit rationalization using utilities from $\mathcal{U}^*$, it suffices to identify a deterministic characterization of $\mathcal{U}^*$-rationalizability and use it in combination with Theorem 1, as Corollaries 1 to 3 do. As an example, ordered-logit rationalizability with Euclidean utilities is straightforward. The intersection property used in Corollary 1 is also necessary and sufficient for deterministic rationalizability with Euclidean utilities, and hence we can reproduce Corollary 1 using Euclidean utilities instead.[8] Similarly, deterministic rationalizability with Cobb-Douglas utilities can be obtained by the well-known property that the fraction of wealth used in each good should be constant across menus. Thus, a version of Corollary 2 using only Cobb-Douglas utilities would follow.

Part (ii) may appear innocuous at first sight. However, fixing a specific parametrization of the class of utilities at stake has consequences. It precludes the type of exercise performed in the proof of Theorem 1, where we relabel the utilities and transform the probability distribution over them. As a result, assuming any specific distribution, such as the logistic, constitutes a relevant restriction. We start by illustrating the point with an example.[9]

[8]Euclidean utilities are defined in the proof of Corollary 1. Given that the sufficiency part of Corollary 1 rationalizes data by using exclusively Euclidean utilities, the intersection property must also characterize ordered-logit rationalizability with Euclidean utilities.

[9]Here is a more abstract reasoning. Whenever $J = 1$, data is trivially ordered-logit rationalizable. However, as we fix the parametrization γ, ordered-logit rationalizability is no longer vacuous and requires data to inherit some of the properties of the logistic distribution, such as symmetry.

Example 1. Suppose that data admits an ordered-logit rationalization over the class $\mathcal{U}^*$ of Euclidean utilities, with $\gamma(t) = -(x - f(t))^2$, where f is the function $f(t) = \frac{t}{2}$ whenever $t \geq 0$ and $f(t) = t$ whenever $t < 0$, and a logistic distribution that has location and scale parameters equal to 0 and 1, respectively. Consider the menu $A_1 = [-\log 2, +\log 2]$. From the closed-form expression of the ordered-logit model, we know that $F_1(-\log 2) = \frac{1}{3}$, $F_1(0) = 0.5$ and $F_1(+\log 2) = \frac{4}{5}$.

Suppose now that we try to explain the above data with an ordered-logit model after having fixed the usual parametrization of $\mathcal{U}^*$ given by $\gamma'(t) = -(x - t)^2$, where we interpret the parameter t as the peak of the utility function. In order to explain the observed data on A_1, we are forced to set the location parameter of the logistic equal to 0. But in that case, given the symmetry of the logistic distribution and of γ', any scale parameter will produce a choice mass for alternative $-\log 2$ equal to the choice mass for alternative $+\log 2$. This is a contradiction with the observed data and hence no ordered-logit rationalization using γ' exists. ■

Consider then a given parametrization of utilities $\gamma : \mathbb{R} \to \mathcal{U}$ that produces ordered choices in every decision problem. We study the exact conditions for data to be rationalized by a logistic distribution over γ. To simplify the exposition, we consider the usual case in which F_j is, in addition to our basic assumptions of Section 2, strictly increasing over the set of non-extreme alternatives $A_j \setminus \{\underline{x}_j, \overline{x}_j\}$ that, from now on, we denote as $(\underline{x}_j, \overline{x}_j)$. We also consider the following richness assumption: for every two decision problems j and j', there exists a sequence of decision problems $j^0 = j, j^1, \ldots, j^k, \ldots, j^K = j'$ such that, for every $k \in \{0, \ldots, K-1\}$, there exists an interval of utilities producing non-extreme maximizers in both A_{j^k} and $A_{j^{k+1}}$. The following example illustrates that this richness condition is commonly met in practice.

Example 2. In some cases, alternatives in $(\underline{x}_j, \overline{x}_j)$ may be in bijection with the set of utilities for every decision problem j. For example, in the standard consumption setting, this is the case for Cobb-Douglas utilities with strictly positive weights. When this happens, the richness assumption is trivially met. Alternatively, $(\underline{x}_j, \overline{x}_j)$ may be in bijection with an unbounded interval of utilities for every decision problem j, as in the case of strictly convex quasi-linear utility functions in the consumption setting, or CRRA expected utilities in decisions under risk.[10] But then, since two right-unbounded

[10] In the latter case, recall that risk averse individuals maximize in non-extreme lotteries approaching but never reaching the 45-degree line.

intervals of utilities must have an unbounded intersection, the assumption is again trivially met. In some occasions, though, $(\underline{x}_j, \overline{x}_j)$ may be the result of a bounded interval of utilities that varies across decision problems. However, we argue that even in this scenario, the richness assumption is not hard to meet. For instance, considering environments involving budget sets, suppose that two bounded intervals of utilities fail to intersect due to the fact that the budget sets have extreme, opposite, prices. Our richness assumption simply requires the presence of some other decision problems with intermediate prices that produce the desired linkage. □

3.1. **A characterization of ordered logit for a fixed parametrization.** As discussed above, fixing a given parametrization γ separates the ordered-logit model from other stochastic ordered-choice models, via the logistic assumption of the distribution of utilities. The general strategy of Theorem 1 is thus insufficient, and we need to provide properties that capture the logistic patterns. Our first property imposes that extreme alternatives have strictly positive mass if and only if there exists some type for which this alternative is optimal. Notice that the ordered-logit model is a random utility model and hence, an alternative can have strictly positive mass only if there is some type for which it is optimal. Moreover, if one such type exists, the ordered nature of choices provokes that extreme alternatives are indeed optimal for an interval of types, and strictly positive mass must result.

Corner extremeness (CE). $F_j(\underline{x}_j) > 0$ (respectively, $\lim_{x \to \overline{x}_j} F_j(x) < 1$) if and only if there exists $t \in \mathbb{R}$ such that $x_j^t = \underline{x}_j$ (respectively, $x_j^t = \overline{x}_j$).

The cumulative logit of alternative x at decision problem j, $\ell(x, A_j) = \log \frac{F_j(x)}{1-F_j(x)}$, is a well-known concept in the analysis of the ordered-logit model. The reason is that when data is generated by a logistic with parameters (τ, σ), it turns to be the case that $\ell(x, A_j) = (t_j^x - \tau)/\sigma$, i.e., the cumulative logit corresponds to the normalized type, under the given parametrization γ, for which alternative x is optimal in decision problem j. Thus, equal sums of types must correspond to equal sums of cumulative logits. Our second property establishes this additive property over non-extreme alternatives.

Cumulative Logit Additivity (CLA). Let $x_j^{t_1}, x_j^{t_2} \in (\underline{x}_j, \overline{x}_j)$ and $x_{j'}^{t'_1}, x_{j'}^{t'_2} \in (\underline{x}_{j'}, \overline{x}_{j'})$ be such that $t_1 + t_2 = t'_1 + t'_2$. Then, $\ell(x_j^{t_1}, A_j) + \ell(x_j^{t_2}, A_j) = \ell(x_{j'}^{t'_1}, A_{j'}) + \ell(x_{j'}^{t'_2}, A_{j'})$.

Theorem 2 shows that, given parametrization γ, CE and CLA are not only necessary but also sufficient for an ordered-logit rationalization. The proof of Theorem 2 comprises a number of steps. First, notice that the mapping between non-extreme alternatives and (a subset of) types is a bijection and hence, the data immediately induces a collection of (yet possibly different across menus) distributions over the subsets of types leading to non-extreme choices. However, we need to account for the censoring generated by extreme choices, that may be optimal for many types, and possibly different in different menus. Whenever this happens, the respective masses observed at the extremes must be appropriately distributed among all the rationalizing types, in such a way as to ensure that the constructed distribution also satisfies the CLA requirement that we know already holds for non-extreme alternatives. We address this requirement by using a recursive construction. Second, the ordered-logit functional form requires us to build upon Galambos and Kotz's (1978) Theorem 2.1.5. This classical, statistical, result provides a necessary and sufficient condition over triplets of real numbers for a single CDF over the reals, which is assumed to be symmetric with respect to the origin, to be logistic. We naturally need to extend this result to our revealed preference setting, where: (i) distributions may have any mean and have not yet been proven to be symmetric and, (ii) we have not one, but a collection of menu-dependent distributions. Our CLA property using quadruplets proves sufficient to show that our menu-dependent distributions are all logistic and, in fact, share the same location and scale parameters.[11]

Theorem 2. *Given γ, F admits an ordered-logit rationalization if and only if F satisfies CE and CLA.*

Proof of Theorem 2: Since the necessity of the axioms is straightforward, we will now prove their sufficiency, as follows. Consider any decision problem j. We construct a sequence of open intervals of types, $\{I_j^0, I_j^1, \dots, I_j^n, \dots\}$, and a sequence of real functions defined over them, $\{G_j^0, G_j^1, \dots, G_j^n, \dots\}$, 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 $k > 0$).

[11]Note also that the parameters (τ, σ) of the logistic distribution that rationalizes the data must be unique, which follows from the assumption that the set of interior bundles has strictly positive mass.

(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 $(\underline{x}_j, \overline{x}_j)$.[12] The first function, G_j^0, corresponds to the function that choice data F_j induces over these types, i.e., for every $t \in I_j^0$, $G_j^0(t) = F_j(x_j^t)$. 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 hence, $\lim_{x\to\overline{x}_j} F_j(x) < 1$ (respectively, $\lim_{x\to\underline{x}_j} F_j(x) > 0$), and the boundedness conditions hold for G_j^0. That is, property (3) is satisfied. We now show that G_j^0 must satisfy property (4). To see this, notice that we can apply CLA with decision problem $j' = j$, using the collection of non-extreme alternatives $x_j^{t_1}, x_j^{t_2}, x_j^{t_1'}, x_j^{t_2'}$, and the equality of cumulative logits over choice data corresponds exactly to property (4) over G_j^0.

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 $t = 2z_j^n - t'$.[13]

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

[12]Notice that this set of types depends on the assumed parametrization γ. Since γ is fixed, we avoid referring to γ in the arguments that follow.

[13]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 extreme 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 extreme 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 extreme 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.

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 $t = 2z_j^n - t'$, 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 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 has changed at the lower end of the interval and hence the property is satisfied, as G_j^{n+1} extends G_j^n. 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 $t_1 < t_1' \leq t_2' < t_2$.[14] Again, we show the case of n even, the other case being analogous. We start by noticing 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.

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

> 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 $\{I_j^0, I_j^1, \ldots, I_j^{n+1}\}$ and $\{G_j^0, G_j^1, \ldots, G_j^{n+1}\}$ satisfy all the properties. The limit interval of the sequence $\{I_j^0, I_j^1, \ldots, I_j^n, \ldots\}$ is the entire set of reals. The limit function of the sequence $\{G_j^0, G_j^1, \ldots, G_j^n, \ldots\}$, 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.[15] That is, there exists $\sigma_j \in \mathbb{R}_{++}$ such that

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

[15] 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.

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 j are explained by this ordered-logistic distribution.

Consider now two decision problems j and j'. By our richness assumption, there exists a sequence of decision problems $j^0 = j, j^1, \dots, j^k, \dots, j^K = j'$ such that, for every $k \in \{0, \dots, K-1\}$, $I^0_{j^k} \cap I^0_{j^{k+1}} \neq \emptyset$. 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. ■

3.2. **Specific settings.** As we did with Theorem 1, we now present some immediate corollaries that illustrate how Theorem 2 may read in specific economic settings. We omit the proofs, as it is apparent that what we write as properties (i) and (ii) correspond, respectively, to the implications of CE and CLA under the setting at hand. The first of the corollaries deals with decisions on a political domain and uses the usual parametrization of Euclidean utilities in which $\gamma(t)$ is the utility function $-(x-t)^2$, where type t has peak at alternative t.

Corollary 4. *In the political domain, there exists a logistic distribution over the peaks of Euclidean utilities rationalizing F if and only if F satisfies: (i) $F_j(\underline{x}_j) > 0$ and $\lim_{x \to \overline{x}_j} F_j(x) < 1$ and (ii) if $x, y \in (\underline{x}_j, \overline{x}_j)$ and $x', y' \in (\underline{x}_{j'}, \overline{x}_{j'})$ are such that $x + y = x' + y'$, then $\ell(x, A_j) + \ell(y, A_j) = \ell(x', A_{j'}) + \ell(y', A_{j'})$.*

The second deals with decision problems involving lotteries and uses the textbook parametrization of CRRA expected utilities in which $\gamma(t)$ corresponds to the utility function $q\frac{x_1^{1-t}}{1-t} + (1-q)\frac{x_2^{1-t}}{1-t}$, i.e. t captures the relative Arrow-Pratt coefficient of risk aversion.[16] Denote by $\kappa(x, A_j) = \frac{\log \frac{x_1}{x_2}}{\log p_j - \log \frac{1-q_j}{q_j}}$, an expression that normalizes the log-consumption ratio by the log-ratio of probabilities and the log-ratio of prices in the menu. We say that c satisfies the constant log-ratio property whenever either

[16]This holds for $t \in \mathbb{R} \setminus \{1\}$ while $\gamma(1)$ is the utility function $q \log x_1 + (1-q) \log x_2$.

$c_j = (\frac{m_j}{p_j}, 0)$ holds for every decision problem j, or $\kappa(c_j, A_j) = \kappa(c_{j'}, A_{j'}) > 0$ for every pair of decision problems. Then,

Corollary 5. *In the state-contingent lotteries domain, there exists a logistic distribution over the relative risk aversion coefficients of CRRA expected utilities rationalizing F if and only if F satisfies: (i) $F_j(\underline{x}_j) > 0$ and $\lim_{x \to \overline{x}_j} F_j(x) = 1$ and (ii) if $x, y \in (\underline{x}_j, \overline{x}_j)$ and $x', y' \in (\underline{x}_{j'}, \overline{x}_{j'})$ are such that $\kappa(x, A_j) + \kappa(y, A_j) = \kappa(x', A_{j'}) + \kappa(y', A_{j'})$, then $\ell(x, A_j) + \ell(y, A_j) = \ell(x', A_{j'}) + \ell(y', A_{j'})$.*

References

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

[2] Agresti, A. (2010), *Analysis of Ordinal Categorical Data*, 2nd edition. New York: Wiley.

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

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

[5] 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.

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

[7] Besley T. and T. Persson (2011), "The Logic of Political Violence," *The Quarterly Journal of Economics*, 126(3):1411–1445.

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

[9] 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.

[10] Bossert, W. and H. Peters (2009), "Single-Peaked Choice," *Economic Theory*, 41(2):213–230.

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

[12] 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.

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

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

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

[16] 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.

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

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

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

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

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

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

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

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

[25] Johnson, V.E. and J.H. Albert (1999), *Ordinal Data Modeling*, New York: Springer.

[26] 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.

[27] Kubler, F., L. Selden, and X. Wei (2014), "Asset Demand Based Tests of Expected Utility Maximization," *American Economic Review*, 104 (11):3459–80.

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

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

[30] 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.

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

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

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

[34] O'Connell, A.A. (2006), *Logistic Regression Models for Ordinal Response Variables*, Sage.

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

[36] 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.