Barcelona School of Economics

WORKING PAPERS

Choice-Based Foundations of Ordered Logit

BSE Working Paper 1323| February 2022

Jose Apesteguia, Miguel A. Ballester

CHOICE-BASED FOUNDATIONS OF ORDERED LOGIT

JOSE APESTEGUIA† AND MIGUEL A. BALLESTER‡

ABSTRACT. We provide revealed preference foundations to ordered logit, for discrete and continuous decision problems. In both cases, the axiomatizations are based on a simple property that reflects the additivity of cumulative logits.

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

1. INTRODUCTION

Ordered logit involves (i) a mapping between preferences and alternatives in which lower preferences are associated with lower choices, and (ii) a logistic distribution over preferences. As a result, choices in a given decision problem are stochastic. Denoting by (τ, σ) the location and scale parameters of the logistic distribution, and by t_j^a the last (threshold) preference where alternative a is maximal in decision problem j, the ordered logit probability of selecting alternatives below or equal to a in decision problem j is

$$F_j^{ol}(a) = \frac{1}{1 + e^{-(t_j^a - \tau)/\sigma}}.$$

Ordered logit is one of the most commonly used models in empirical analysis across various fields such as Economics, Sociology or Biology.[1] In spite of the massive empirical use of ordered logit, it lacks a proper revealed-preference analysis. This paper is the first to provide choice-based foundations for the ordered logit model.

Date: February, 2022.

* Financial support by the Spanish Ministry of Economics and Competitiveness through Grant PGC2018-098949-B-I00 and through the Severo Ochoa Program for Centers of Excellence in R&D (CEX2019-000915-S) and Balliol College is gratefully acknowledged.

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

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

[1]See, e.g., the textbook treatments of Johnson and Albert (1999), O'Connell (2006), Agresti (2010), Greene and Hensher (2010).

We assume a given continuous family of preferences and observe choice behavior over an arbitrary collection of decision problems. We pose the question as to what choice data should be like in order to be rationalized by a logistic distribution over the family of preferences. We study the two most fundamental choice settings in the application of the model, one involving discrete choice and the other continuous choice. In the discrete case, choices are a censored version of preferences, in the sense that each alternative is chosen by an interval of preferences. In the continuous case, we have to distinguish between interior choices and the corners. In the interior, there is a bijection between choices and preferences. However, choices may be censored at the corners, where they may be maximal for their corresponding intervals of preferences.

Our characterizations build upon the statistical literature. Galambos and Kotz (1978), using classical results on the exponential distribution by Cauchy (1821), give a condition under which a probability distribution defined over the real numbers, and assumed to be symmetric with respect to the origin, can be a logistic distribution. Our choice-based approach needs to depart from the statistical approach in a number of ways. First, for any given decision problem, we observe the distribution of choices, not the underlying distribution over preferences. Due to censoring, choices provide only partial information over the preference distribution, and the challenge here is how to expand the partial information to the totality of preferences. A second issue to be addressed is that the distribution of choices may not generate a symmetric distribution at the origin. A third is that, since there are multiple decision problems, we observe multiple choice distributions, and we need to show that they have mutual consistency and will therefore allow us to obtain a single underlying distribution over preferences.

Our axiomatic foundations are based on a single, intuitive property, which can be operationalized by means of the usual cumulative logits. For a given alternative a in decision problem j, the cumulative logit is

$$\ell_j(a) = \log \frac{F_j(a)}{1 - F_j(a)},$$

where $F_j(a)$ is the observed cumulative choice probability in j of the alternatives below or equal to a. Cumulative logit is a standard tool in the treatment of ordered logit, where it corresponds to the normalized threshold preference $(t^a_j - \tau)/\sigma$. Notice then that, for every two pairs of threshold preferences in two, possibly different, decision problems, $\{t^a_j, t^b_j\}$ and $\{t^{a'}_{j'}, t^{b'}_{j'}\}$, such that $t^a_j + t^b_j = t^{a'}_{j'} + t^{b'}_{j'}$, it follows immediately that ordered logit implies that $\ell_j(a) + \ell_j(b) = \ell_{j'}(a') + \ell_{j'}(b')$. In our study of the continuous setting, we show that this property is not only necessary, but, remarkably, sufficient. That is, any choice dataset satisfying the property can be rationalized by a logistic distribution over preferences. The corresponding analysis of the discrete setting requires us to consider a generalization of this property that operates over any two equally-sized collections of threshold types. Interestingly, the proofs of the two characterization results are susceptible to the same statistical treatment as used by Galambos and Kotz (1978), but deal in unique ways with the specific natures of their censored choice data.

We close this introduction by placing our contribution within the literature. First and foremost, as already mentioned, this paper provides choice-based foundations to the extensive empirical literature using ordered logit in both continuous and discrete settings.[2] This paper also contributes to the strand of literature that provides choice-based foundations to various stochastic choice models. The classic works are those of Luce (1959) and Block and Marshak (1960).[3] Another relevant strand of the literature is that formed by 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). The current paper is the first in obtaining the revealed preference foundations of ordered logit.

2. Continuous choice

The ordered set of preference types is the real line, $\mathbb{R}$. Let $\mathbb{J} = \{1, 2, \ldots, J\}$ be a collection of continuous decision problems.[4] For every decision problem $j \in \mathbb{J}$ and type $t \in \mathbb{R}$, we denote by a_j^t the alternative chosen by type t in j, assuming it to be

[2] Apart from the references cited above, influential examples of papers using empirically ordered logit in fields as diverse as political economy, finance, labour, political science, welfare, management, gender and networks are, respectively, Besley and Persson (2011), Kaplan and Zingales (1997), Blau and Hagy (1998), Baum (2002), Campante and Yanagizawa-Drott (2015), Cummings, (2004), Carlana (2019) and Bailey, Cao, Kuchler and Stroebel (2018).

[3] 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).

[4] The archetypical examples of the continuous setting in Economics are linear budget sets.

unique (except for a non-measurable set of types). We assume that the set of maximal alternatives within a decision problem, $\bigcup_{t\in\mathbb{R}}\{a_j^t\}$, is a connected subset of alternatives on the frontier of the decision problem. Thus, we opt to represent each continuous decision problem by the interval $[0,1]$. Ordered choice is formalized by assuming that for every $j \in \mathbb{J}$:

(1) $\lim\limits_{t\to-\infty} a_j^t = 0$,
(2) $t_1 < t_2$ implies $a^{t_1} \leq a_j^{t_2}$ with strict inequality whenever $a_j^{t_1}, a_j^{t_2} \in (0,1)$,
(3) $\lim\limits_{t\to+\infty} a_j^t = 1$.

Notice that the ordered choice structure allows us to invert the (subset of the) map of types to (interior) choices and, whenever this is the case, we can denote by t_j^a the unique type for which alternative $a \in (0,1)$ is maximal in menu j.

2.1. **Ordered-logistic choices.** In a continuous setting, an ordered-logistic distribution with median type $\tau \in \mathbb{R}$ and scale parameter $\sigma \in \mathbb{R}_{++}$ determines a collection of menu-dependent, continuous CDFs describing choices. Interior alternatives are selected by a unique type and have no mass. For every $j \in \mathbb{J}$ and every alternative $a \in (0,1)$, the choice probability, within menu j, of the measurable set of alternatives $[0,a]$ is

$$F_j^{ol}(a) = \frac{1}{1+e^{-(t_j^a-\tau)/\sigma}}.$$

The value of the choice probability of any interval $(a_1,a_2]$, with $a_1,a_2 \in (0,1)$, is $F_j^{ol}(a_2) - F_j^{ol}(a_1)$. Notice that we allow for corner alternatives of a decision problem, $a \in \{0,1\}$, to be maximal for a measurable subset of types. Whenever this is the case, these corner alternatives will obviously be atoms that are chosen with strictly positive mass in the corresponding decision problem. If $a=0$ is chosen by a measurable subset of types within decision problem j, ordered choice guarantees that there is a largest type making that choice, i.e., $t_j^0 = \lim\limits_{a\to0^+} t_j^a$. Hence, the choice probability of the corner alternative $a=0$ in ordered-logistic choice is equal to:

$$\lim_{a\to0^+} F_j^{ol}(a) = \frac{1}{1+e^{-(t_j^0-\tau)/\sigma}}.$$

Similarly, if $a=1$ is selected by a measurable subset of types in decision problem j, we can denote by t_j^1 the smallest type which selects that particular alternative, and the choice probability of this corner alternative in ordered-logistic choice will be equal to:

$$1 - \lim_{a \to 1^-} F_j^{ol}(a) = 1 - \frac{1}{1 + e^{-(t_j^1 - \tau)/\sigma}}.$$

The interpretation of the parameters is straightforward. Parameter τ captures the median type, which is equal to the mean type, in the distribution. Notice that, when τ points to an interior alternative $a_j^\tau \in (0,1)$, it is $F_j^{ol}(a_j^\tau) = .5$.[5] Parameter σ captures the variability of the distribution around this median type. When σ goes to zero, the model resembles deterministic rational choice, with the choice probability of any interval containing type τ approaching 1, and the choice probability of any interval not containing τ approaching zero. When σ goes to ∞, the density of τ becomes closer to that of any other type for any given interval (and thus, the CDF resembles uniform choices over larger intervals of types).

2.2. **A characterization.** Choice data are represented by a collection of menu-dependent CDFs, $\{F_j\}_{j \in \mathbb{J}}$, each of which is defined over the space of alternatives $[0,1]$. We assume that each of these CDFs is continuous and strictly increasing in $(0,1)$. We also assume that, for every decision problem, a corner alternative, $a \in \{0,1\}$, has strictly positive mass if, and only if, it is maximal for some measurable subset of types. Finally, we consider the following richness assumption, which states that, for every two decision problems $j, j' \in \mathbb{J}$, there exists a sequence of decision problems $j^0 = j, j^1, \ldots, j^k, \ldots, j^K = j'$ such that, for every $k \in \{0, \ldots, K-1\}$, the set of types, $\bigcup_{a \in (0,1)} \{t_{j^k}^a\}$, leading to interior alternatives in decision problem j^k has measurable intersection with the set of types, $\bigcup_{a \in (0,1)} \{t_{j^{k+1}}^a\}$, leading to interior alternatives in decision problem j^{k+1}. That is, we can link the types that are informative in decision problem j to the types that are informative in decision problem j', without necessarily assuming that these two decision problems have measurable intersection between informative types, but simply that this occurs through some chain of decision problems. This is a minimal richness assumption satisfied by all relevant datasets.

We now introduce a simple property on choice data $\{F_j\}_{j \in \mathbb{J}}$ and show that it is necessary and sufficient for choice data to be rationalized by a logistic distribution over preferences. The property considers a pair $\{a, b\}$ of interior alternatives in decision problem j and a pair $\{a', b'\}$ of interior alternatives in decision problem j', and it states

[5]Obviously, if $a_j^\tau = 0$ or $a_j^\tau = 1$, we know that $\lim_{a \to 0^+} F_j^{ol}(a) \geq .5$ and $\lim_{a \to 1^-} F^{ol}(a) \leq 0.5$, respectively.

that, if the sum of types $t_j^a + t_j^b$ coincides with the sum of types $t_{j'}^{a'} + t_{j'}^{b'}$, the corresponding sums of cumulative logits coincide. That is:

Cumulative Logit Additivity (CLA). For every $j, j' \in \mathbb{J}$ and every $a, b, a', b' \in (0,1)$, $t_j^a + t_j^b = t_{j'}^{a'} + t_{j'}^{b'}$ implies that $\ell_j(a) + \ell_j(b) = \ell_{j'}(a') + \ell_{j'}(b')$.

We can now establish the following result.

Theorem 1. *In a continuous setting, $\{F_j\}_{j \in \mathbb{J}}$ is ordered logistic iff it satisfies CLA.*

Proof of Theorem 1: In a continuous setting, it is immediate to see that any choice data that are ordered logistic must satisfy CLA. We then need to prove the sufficiency part of the result. For this, we start by constructing, for every decision problem $j \in \mathbb{J}$, 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 a value $k < 1$ (respectively, strictly bounded from below by a 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 $\log \frac{G_j^n(t_1)}{1-G_j^n(t_1)} + \log \frac{G_j^n(t_2)}{1-G_j^n(t_2)} = \log \frac{G_j^n(t_1')}{1-G_j^n(t_1')} + \log \frac{G_j^n(t_2')}{1-G_j^n(t_2')}$.

For every decision problem $j \in \mathbb{J}$, the first interval of types, I_j^0, is the set of types that select interior alternatives in j, i.e.,

$$I_j^0 = \bigcup_{a \in (0,1)} \{t_{j^a}\}.$$

The first function, G_j^0, corresponds to the function induced by choice data over these types, i.e., for every $t \in I_j^0$,

$$G_j^0(t) = F_j(a_j^t).$$

G_j^0 is a well-defined function thanks to the uniqueness assumption made on optimal alternatives. It is obviously strictly increasing and takes values in $(0,1)$, given the assumptions made over F_j. Moreover, if the interval I_j^0 is bounded from above (respectively, from below), there is a measurable subset of types selecting $a = 1$ (respectively,

$a = 0$) and hence, $\lim_{a\to 1^-} F_j(a) < 1$ (respectively, $\lim_{a\to 0^+} F_j(a) > 0$) and the boundedness conditions hold for G_j^0. In addition, G_j^0 must satisfy property 4. To see this, notice that we can apply CLA with $j = j'$ using the collection of interior alternatives $a_j^{t_1}, a_j^{t_2}, a_j^{t_1'}, a_j^{t_2'}$, and the equality over choice data corresponds exactly to property 4 over G_j^0.

The remaining intervals and functions are now defined recursively. Given the collections $\{I_j^0, I_j^1, \dots, I_j^n\}$ and $\{G_j^0, G_j^1, \dots, G_j^n\}$ satisfying all the properties, we define interval I_j^{n+1} and function G_j^{n+1} such as to guarantee that $\{I_j^0, I_j^1, \dots, I_j^{n+1}\}$ and $\{G_j^0, G_j^1, \dots, G_j^{n+1}\}$ also satisfy the properties. Our first consideration is the definition of the new interval of types, I_j^{n+1}, which depends on the parity of n. If n is an even (respectively, an odd) integer, we define interval I_j^{n+1} as follows:

- If I_j^n is not bounded from above (respectively, from below) define $I_j^{n+1} = I_j^n$.
- If I_j^n is bounded from above (respectively, from below), define I_j^{n+1} as the union of: (i) the previous interval I_j^n, (ii) the lowest upper bound (respectively, the largest lower bound) z_j^n of interval I_j^n, and (iii) the set of types t for which there exists $t' \in I_j^n$ such that $t = 2z_j^n - t'$.[6]

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 alternative z_j^n, define $G_j^{n+1}(z_j^n) = \lim_{s\to z_j^n} G_j^n(s)$, where a right-hand or left-hand limit 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}.

We now show that I_j^{n+1} and G_j^{n+1} satisfy the properties above. First, it is immediate to see that $I_j^n \subseteq I_j^{n+1}$ and, hence, property 1 holds.

[6]Intuitively, we are simply duplicating the original bounded interval I_j^n to its right (respectively, to its left) and adding the boundary points between the two. This step is not needed when there are no corner choices, and hence the initial interval I_j^0 is the set of all types $\mathbb{R}$. When choices are observed in only one of the corners, or, equivalently, I_j^0 is bounded on one side, a unique duplication is needed, which already forms the entire real line. If choices are observed in both corners, 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.

Second, it is also immediate that the construction guarantees that the function G_j^{n+1} extends G_j^n, therefore property 2 is satisfied.

Third, 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 an immediate 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 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 $\log \frac{G_j^{n+1}(t_1)}{1-G_j^{n+1}(t_1)} = \log \frac{G_j^n(t_1)}{1-G_j^n(t_1)} < \log \frac{G_j^n(z_j^n)}{1-G_j^n(z_j^n)} < 2\log \frac{G_j^n(z_j^n)}{1-G_j^n(z_j^n)} - \log \frac{G_j^n(t_2')}{1-G_j^n(t_2')} = \log \frac{G_j^{n+1}(t_2)}{1-G_j^{n+1}(t_2)}$, 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 of them 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 merely 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 due to 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 in the lower part of the interval and hence the property is satisfied, as G_j^{n+1} extends G_j^n. For the upper part of the interval, suppose that I_j^{n+1} is bounded from above. Therefore, it must be that I_j^n is bounded from below (say, with largest lower bound k). It then becomes obvious 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.

Fourth, 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 $t_1 < t_1' \leq t_2' < t_2$.[7] 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 $\overline{I}_j^n$, thanks to the recursive assumption on G_j^n, the fact that

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

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 $\overline{I}_j^n$. In this case, 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: $t_1 \in \overline{I}_j^n$. In this case, 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 clearly $t_1 + s_1' + s_2' = s_2 + 2z_j^n$. Define $\hat{t} = s_2 + z_j^n - t_1$, which clearly 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: $t_1, t_1' \in \overline{I}_j^n$. In this case, 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 clearly $t_1 + s_2' = t_1' + s_2$ 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: $t_1, t_1', t_2' \in \overline{I}_j^n$. In this case, there must exist $s_2 \in I_j^n$ such that $t_2 = 2z_j^n - s_2$. It must clearly be that $t_1 + 2z_j^n = t_1' + t_2' + s_2$. Define $\hat{t} = t_1 + z_j^n - t_1'$, which clearly 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 property 4 holds and hence, we have shown 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:

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

We claim that H_j is a continuous, strictly increasing CDF over the reals that is symmetric with respect to the origin. We need to show symmetry. For this, consider $t_1 = \tau_j - x$, $t_2 = \tau_j + x$ 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, $H_j(-x) = G_j(t_1) = 1 - G_j(t_2) = 1 - H_j(x)$, and the symmetry of H_j has been proved.

Consider now the following function defined over the positive reals:

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

Since H_j is a continuous, strictly increasing CDF over the reals with $H_j(0) = .5$, it is immediate that $1 - O_j(x)$ must be a continuous, strictly increasing CDF over the positive reals with no mass at zero. Moreover, given that G_j satisfies property 4 above, the definition of H_j and O_j guarantees that $O_j(x)O_j(z) = O_j(x+z)$ must hold for every pair of positive real values x and z. One can then reproduce the standard argument that goes back to Cauchy (1821), and is described in Galambos and Kotz (1978; Theorem 1.3.1), which guarantees that O_j must be of the exponential type with no mass at the origin.[8] That is, there exists $\sigma_j \in \mathbb{R}_{++}$ such that

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

That is, for every $x \geq 0$, it is true that $H_j(x) = \frac{1}{1+e^{-x/\sigma_j}}$. Moreover, given the symmetry of H_j with respect to the origin, for every $x < 0$, it must also be true that $H_j(x) = 1 - H_j(-x) = 1 - \frac{1}{1+e^{x/\sigma_j}} = \frac{1}{1+e^{-x/\sigma_j}}$. That is, H_j is a logistic distribution with median zero and parameter σ_j and, evidently, G_j is ordered logistic with median τ_j and 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 $j, j' \in \mathbb{J}$. By our richness assumption, there exists a sequence of menus $j^0 = j, j^1, \ldots, j^k, \ldots, j^K = j'$ such that, for every $k \in$

[8] The property is satisfied by exponential types with and without mass at zero. Since we know that O has no mass at zero, it must be of the latter type.

$\{0, \ldots, K-1\}$, $I^0_{j^k} \cap I^0_{j^{k+1}} \neq \emptyset$. 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-logistic structure of G_{j^k} and $G_{j^{k+1}}$, it follows immediately that they must both have a common median type τ and a common parameter σ. The recursive application of this argument shows that G_j and $G_{j'}$ must have the same common median type τ and parameter σ, which concludes the proof. ■

The proof of the theorem comprises a number of steps. First, we need to account for the censoring generated by choices of the corner alternatives. Alternatives $a \in (0, 1)$ are associated with a unique type and hence, the choice distribution over these alternatives immediately induces a distribution over the corresponding types. However, alternatives $\{0, 1\}$ may be associated with a measurable set of types generating these choices. Whenever this is the case, the respective masses of the corner alternatives must be appropriately distributed among these types, in such a way as to ensure that the distribution over types thus constructed satisfies the assumed CLA constraint on the interior choices. We use a recursive construction to address this requirement. Second, in order to adopt the ordered-logistic form, we build upon Galambos and Kotz (1978; Theorem 2.1.5), which provides a necessary and sufficient condition over triplets of real numbers for a single CDF that is symmetric with respect to the origin. Importantly, notice that in our revealed preference setting, the induced distribution over types can have any mean, therefore its symmetry needs to be proved. We show that our CLA property using quadruplets provides sufficient proof. Third, we need to show that all the induced distributions over types, one for each decision problem, share the location and scale parameters of the logistic distribution. We do this by applying the CLA property to different decision problems.

We close this section by noting that the parameters of the logistic distribution (τ, σ) rationalizing the choice data are unique. This is guaranteed by the assumed existence of interior choices.

3. Discrete Choice

The ordered set of preference types is the real line, $\mathbb{R}$. Let $\mathbb{J} = \{1, 2, \ldots, J\}$ be a collection of discrete decision problems. We assume ordered choice over the maximal alternatives, which, in the present discrete setting, implies that, in any given menu j, types $(-\infty, t^1_j)$ uniquely select one alternative, which we label as 1, types (t^1_j, t^2_j)

uniquely select another alternative labelled 2, and so on and so forth, with types $(t_j^{I_j}, \infty)$ uniquely selecting another alternative labelled $I_j + 1$.[9] Denote the set of maximal alternatives in decision problem j by $\mathbb{I}_j = \{1, 2, \ldots, i, \ldots, I_j + 1\}$.

3.1. **Ordered-logistic choices.** In a discrete choice setting, an ordered-logistic distribution with median type $\tau \in \mathbb{R}$ and scale parameter $\sigma \in \mathbb{R}_{++}$ determines a collection of menu-dependent (discrete) probability distributions. We can work with the corresponding (discrete) CDFs; for every $j \in \mathbb{J}$ and every $i < I_j + 1$, the cumulative choice probability of alternatives $\{1, 2, \ldots, i\}$ in decision problem j is determined by the threshold type t_j^i as

$$F_j^{ol}(i) = \frac{1}{1 + e^{-(t_j^i - \tau)/\sigma}}.$$

Again, the interpretation of the parameters is straightforward. Parameter τ captures the median type and hence, when $\tau = t_j^i$, it is $F_j^{ol}(i) = .5$. Parameter σ captures the variability of the distribution around this median type. When σ goes to zero, the model resembles deterministic rational choice, with the mass of any interval containing type τ approaching 1 (and hence, the alternative i in menu j such that $t_j^{i-1} < \tau < t_j^i$ concentrates all the choice probability). When σ goes to ∞, the density of τ becomes closer, for any given interval, to that of any other type (and hence, alternatives 1 and $I_j + 1$, which are selected by unbounded intervals of types, concentrate all the choice probability, each to the same degree).

3.2. **A characterization.** Choice data are represented by a collection of discrete menu-dependent CDFs $\{F_j\}_{j \in \mathbb{J}}$ each of which is defined on $\mathbb{I}_j$. F_j describes the observed, cumulative, choice probabilities in decision problem j, with $F_j(i)$ representing the cumulative choice probability of alternatives $\{1, 2, \ldots, i\} \subseteq \mathbb{I}_j$. We assume positivity, that is, $0 < F_j(1) < \cdots < F_j(i) < \cdots < F_j(I_j) < 1 = F_j(I_j + 1)$ for every $j \in \mathbb{J}$. In addition, we make the following technical assumption: there exist threshold types $t_{j_1}^{i_1}, t_{j_2}^{i_2}, t_{j_3}^{i_3}, t_{j_4}^{i_4}$, and $t_{j^*}^{i^*}$ such that $F_{j_1}(i_1) < F_{j^*}(i^*) = .5 < F_{j_2}(i_2)$, and the value $\frac{t_{j_3}^{i_3} - t_{j^*}^{i^*}}{t_{j_4}^{i_4} - t_{j^*}^{i^*}}$ is an irrational number.

We introduce a generalization 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,

[9] Since types have no mass in our analysis, the choice at threshold types is irrelevant and ties can be avoided.

if, and only if, the corresponding sum of cumulative logits in the former is larger than that in the latter. That is,

Generalized Cumulative Logit Additivity (GCLA). For every positive integer M, and for every two collections of threshold types $\{t_{j_1}^{i_1}, \ldots, t_{j_m}^{i_m}, \ldots, t_{j_M}^{i_M}\}$ and $\{t_{j'_1}^{i'_1}, \ldots, t_{j'_m}^{i'_m}, \ldots, t_{j'_M}^{i'_M}\}$, $\sum_{m=1}^{M} t_{j_m}^{i_m} \geq \sum_{m=1}^{M} t_{j'_m}^{i'_m}$ iff $\sum_{m=1}^{M} \ell_{j_m}(i_m) \geq \sum_{m=1}^{M} \ell_{j'_m}(i'_m)$.

We can then establish the following result.

Theorem 2. *In a discrete setting, $\{F_j\}_{j \in \mathbb{J}}$ is ordered logistic iff it satisfies GCLA.*

Proof of Theorem 2: In a discrete setting, it is immediate to see that any choice data that are ordered logistic must satisfy GCLA. 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 $t_{j^*}^{i^*}$ such that $F_{j^*}(i^*) = .5$. Consider then the subsets of real numbers

$$\mathcal{T} = \{x : x = t_j^i - t_{j^*}^{i^*} \text{ for some } j \in \mathbb{J} \text{ and } i < I_j + 1\},$$

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

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.[10] We can then find, for every $x \in \mathbb{R}$, a sequence of elements $(x_1, x_2, \ldots, x_k, \ldots)$ 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 $\{t_{j_1}^{i_1}, \ldots, t_{j_v}^{i_v}, \ldots, t_{j_V}^{i_V}\}$ and $\{s_{j_1}^{i_1}, \ldots, s_{j_w}^{i_w}, \ldots, s_{j_W}^{i_W}\}$ such that $x_k = \sum_{v=1}^{V} n_v (t_{j_v}^{i_v} - t_{j^*}^{i^*}) - \sum_{w=1}^{W} n_w (s_{j_w}^{i_w} - t_{j^*}^{i^*})$, 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. GCLA guarantees that the cumulative choice probability is the same in both decision problems, and hence the function H is well defined.[11] We now prove that H is increasing. Let $x < x'$. We know

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

[11] 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.

that there exist sequences of elements $(x_1, x_2, \ldots, x_k, \ldots)$ and $(x'_1, x'_2, \ldots, x'_k, \ldots)$ 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 $\{n'_w, s^{i'_w}_{j'_w}\}_{w=1}^{W'}$. 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 GCLA guarantees that the sum of cumulative logits 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 $F_{j_1}(i_1) < F_{j^*}(i^*) = .5 < F_{j_2}(i_2)$ 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 1.

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 1 can be used to show the ordered-logistic nature of H, as desired. ■

The proof of the theorem is different from that of Theorem 1. In the case of continuous choice, choice data from a decision problem provide information on an interval of types that can be extended to the real line yet still satisfy CLA. Then, we use the intersection of the intervals of types across menus and the CLA property to guarantee that the extensions are of the same CDF. However, each discrete choice problem provides CDF 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, GCLA, 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 1.[12]

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), "Random Utility Models with Ordered Types and Domains," mimeo.

[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] Baum M. (2002), "Sex, Lies, and War: How Soft News Brings Foreign Policy to the Inattentive Public," *American Political Science Review*, 96(1):91–109.

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

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

[10] 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.

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

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

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

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

[29] 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.

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

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

[32] 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.