Self-progressive choice models*

KEMAL YILDIZ

Bilkent University and Princeton University

the latest version: https://arxiv.org/abs/2212.13449

February 11, 2024

Abstract

Consider a population of heterogenous agents whose choice behaviors are partially *comparable* according to a given *primitive ordering*. An analyst seeks to select a *choice model*–a set of choice functions–to explain the observed random choice behavior. As a criterion to guide the model selection process, we introduced *self-progressiveness* ensuring that each aggregate choice behavior explained by the model has a unique orderly representation within the model. We establish an equivalence between self-progressive choice models and well-known algebraic structures called *lattices*. This equivalence provides for a precise recipe to restrict or extend any choice model for unique orderly representation. To prove out, we characterize the minimal self-progressive extension of rational choice functions. Then, we provide necessary and sufficient conditions for (unique) identification of the underlying primitive ordering.
Keywords: Random choice, heterogeneity, identification, unique orderly representation, lattice, multiple characteristics, ternary relations,integer-programming.

*I am grateful to Christopher Chambers, Battal Dogan, Faruk Gul and Ariel Rubinstein for helpful discussions and feedback. I thank seminar participants at University of Maryland, Georgetown University, Duke University, Princeton University, and participants of several conferences and workshops for valuable comments and suggestions. This research is supported by the Marie Curie International Reintegration Grant (# 837702) within the European Community Framework Programme. I am grateful to Princeton University Department of Economics for their hospitality during my visit.
First posted version: December 27, 2022. Email: kemal.yildiz@bilkent.edu.tr

Contents

1 Introduction

Random choice models are used successfully to identify heterogeneity in the aggregate choice behavior of a population. The success is achieved despite prominent choice models, such as the random utility model, are underidentified in the sense that the observed choice behavior renders diverse representations within the model. The typical remedy to this challenging matter has been structuring the model to obtain a unique representation and achieve point-identification.[1] Here, instead of focusing on a specific choice model, we adopt a novel approach in which we start with *choice models* as our objects of analysis, and assume an "orderliness" in the population that enables partial comparison of agents' choice behaviors.[2] We formulate and analyze *self-progressive choice models* that guarantee a unique orderly representation within the model for each aggregate choice behavior explained by the model. Our motivation stemmed from the potential value of self-progressive choice models in organizing random choice data.

We first establish an equivalence between self-progressive choice models and well-known algebraic structures called *lattices* (Theorem 1). It follows from this equivalence that self-progressive models allow for specification of multiple behavioral characteristics that is critical in explaining economically relevant phenomena. Additionally, we obtain a precise recipe and a tool to restrict or extend any choice model to be self-progressive. To prove out, we characterize the minimal self-progressive extension of rational choice functions, which offers an intuitive explanation for the *choice overload* phenomena (Theorem 2). We then investigate how to identify the orderliness in the population that renders our choice overload representation to a choice model (Theorem 3).

[1] See for example Gul & Pesendorfer (2006) and Dardanoni, Manzini, Mariotti, Petri & Tyson (2022).
[2] See the discussions by Apesteguia, Ballester & Lu (2017) and Filiz-Ozbay & Masatlioglu (2023).

Our consideration revolves around an analyst aiming to select a suitable "choice model" to elucidate observed random choice behavior. A *choice model*–such as the rational choice model–is simply a set of (*admissible*) choice functions that may be adopted by any agent in the population. As a criterion to guide the analyst to select a choice model, we will introduce *self-progressiveness*.

The analyst aims to deduce population heterogeneity through a *primitive ordering* over alternatives. A primitive ordering can incorporate elements such as risk attitudes (Chiappori et al. 2019), social preferences (Dillenberger & Sadowski 2012),[3] or rational assessments–such as consumers' valuations for commodities free from firm obfuscation (Spiegler 2016). A pair of choice functions are *comparable* if the alternative chosen by one of the choice functions is ranked higher than the alternative chosen by the other for every choice set. In our illustrative Example 2, lotteries are ranked by expected monetary payoffs, while choice functions reflect agent types with varied bounded rationality.

To introduce *self-progressiveness*, suppose that the analyst represents the aggregate choice behavior of a population as a probability distribution over a set of admissible choice functions. The same aggregate choice behavior renders a unique representation as a probability distribution over–possibly different–choice functions that are *comparable* to each other.[4] *Self-progressiveness* requires these comparable choice functions to be admissible–to be contained in the model–as well. Thus, a self-progressive model provides a language to the analyst that allows for orderly representing any aggregate choice behaviour that is explained by the model via using the elements of the same model.

[3]Other instances include tax policies ordered by the total revenue (Roberts 1977), payments ordered by the present value (Manzini & Mariotti 2006), acts ordered by ambiguity level (Chew et al. 2017). Our Theorem 1 holds even if we permit primitive ordering to depend on the available alternatives, called a *choice set*. Thus, we accommodate, for instance, the temptation or information processing costs that depend on the availability of more tempting or memorable alternatives in a choice set.

[4]This fact is shown by Filiz-Ozbay & Masatlioglu (2023) that we will discuss in detail.

In our main result (Theorem 1), by using a simple probabilistic decomposition procedure, we establish an equivalence between self-progressive choice models and *lattices*. A choice model forms a *lattice* if for each pair of admissible choice functions, their *join*–formed by collecting the better choices–and *meet*–formed by collecting the worse choices–are admissible as well. Hence, self-progressive choice models extend beyond the scope of previously analyzed models in which agents' choices are ordered according to a single characteristic. To demonstrate the relevance of this generality, we present examples of choice models in which multiple behavioral characteristics are parameterized.

Theorem 1 provides for a precise recipe to restrict or extend any choice model for unique orderly representation. To prove out, Theorem 2 characterizes the *minimal self-progressive extension* of rational choice functions via two choice axioms. The resulting model offers an intuitive explanation for why agents might exhibit *choice overload*.[5] In that, the axioms require a more valuable (or the same) alternative be chosen whenever we remove alternatives that are less valuable than the chosen one, or add alternatives that are more valuable than the chosen one. Proposition 1 characterizes the random counterpart of the model by applying classical integer-programming techniques.

Until now, we have assumed the analyst specifies the primitive ordering. Can we, however, effectively infer the primitive ordering for a choice model? Theorem 3 presents necessary and sufficient conditions for the existence and uniqueness of a primitive ordering that renders our choice overload representation to a choice model. We use classical and modern results from foundational geometry to identify the primitive ordering. We conclude with a simple observation on choice models rendering unique orderly representations, irrespective of primitive orderings.

[5]Choice overload refers to the phenomena that agents tend to deviate from their accurate preferences in complex environments. See Chernev, Böckenholt & Goodman (2015) for a recent meta-analysis.

The following example illustrates self-progressiveness and our main result.

Example 1. Consider a population of agents choosing from subsets of the alternatives a, b, and c. The analyst contemplates using a model μ to understand the choice behaviour of the agents based on the observed `RCF` (random choice function) and the primitive ordering $a > b > c$, which is of particular interest to the analyst. Figure 1 specifies the observed `RCF` (on the right) and the four choice functions comprising the choice model μ, where the choice functions are arranged according to the domination relation $\rhd$ obtained from $>$ (on the left).

RCF	a	b	c
$\{a,b,c\}$	1	0	0
$\{a,b\}$	$\frac{2}{3}$	$\frac{1}{3}$	—
$\{a,c\}$	1	—	0
$\{b,c\}$	—	$\frac{2}{3}$	$\frac{1}{3}$

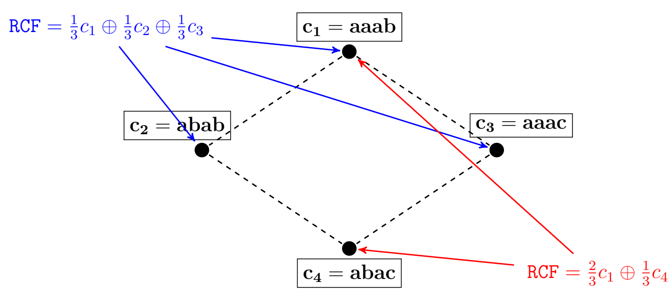


Figure 1: Each node of the graph on the left specifies the chosen alternatives from choice sets $\{a,b,c\}$, $\{a,b\}$, $\{a,c\}$, $\{b,c\}$ respectively. Dotted lines correspond to the domination relation $\rhd$ among choice functions obtained from $>$.

The observed `RCF` can be represented as the probability distribution that assigns equal weights to choice functions c_1, c_2, and c_3. However, this representation is not progressive, since neither c_2 nor c_3 dominates the other.[6] Alternatively, the same `RCF` can be represented by assigning a weight of $\frac{2}{3}$ to the choice function c_1–which chooses the best alternative from each choice set–and a weight of $\frac{1}{3}$ to the choice function c_4–which chooses the worst ones. Since c_1 dominates c_4, the latter representation is progressive.[7] Given that the choice model μ contains c_1 and c_4, we can not refute that μ is self-progressive. However, to conclude that μ is self-progressive, this exercise should be

[6]In that, c_2 chooses b, the worse alternative from choice set $\{a,b\}$, and the better one from the choice set $\{b,c\}$, while c_3 does the opposite.

[7]There are infinitely many other representations for the given `RCF`, but none of them is progressive.

repeated for each `RCF` that can be represented as a probability distribution over choice functions c_1, c_2, c_3, and c_4. Put differently, each representation of the former type (highlighted in blue) should be paired with a corresponding progressive representation (highlighted in red). In contrast, our Theorem 1 directly establishes that μ is self-progressive, since $\langle \mu, \rhd \rangle$ forms a lattice.

1.1 Related literature

We pursue a novel approach offering a tool for analysts to tackle with underidentification issue that is commonly observed for random choice models. The findings of two recent studies that use the orderliness in the population are precursory for our formulation of self-progressiveness. Here, we aim to highlight the conceptual significance and economic relevance of our contribution in this context.

Apesteguia, Ballester & Lu (2017) are the first who use the "orderliness" in the population to refine the random utility model for unique representation. In addition to their axiomatic characterization, they observed that if a random utility model is represented as a probability distribution over comparable rational choice functions, called *single crossing random utility model* (SCRUM), then the representation must be unique.[8] Extending this observation, Filiz-Ozbay & Masatlioglu (2023) show that each random choice function can be uniquely represented as a probability distribution over choice functions that are comparable to each other. These findings motivated us for employing "progressiveness" to select a representation for a random choice function from what is typically an infinite array of possibilities as demonstrated in Example 1.

[8] See Costa, Ramos & Riella (2020) and Petri (2023) for extension to random choice correspondences.

To highlight the difference, in our terminology, Apesteguia, Ballester & Lu (2017) and Filiz-Ozbay & Masatlioglu (2023) show that SCRUM and the entire set of choice functions are two examples of self-progressive choice models. *Our aim is to determine the characteristics that define the comprehensive family of self-progressive choice models*–a pursuit that holds conceptual and technical significance.

We next discuss the economic relevance of this generality. Both Apesteguia, Ballester & Lu (2017) and Filiz-Ozbay & Masatlioglu (2023) present intriguing examples in which agents' choices are ordered according to a single characteristic. However, it remains unclear whether orderly representations exist for models that capture how agents' choices vary with different behavioral characteristics. A classical example is the equity premium puzzle (Mehra & Prescott 1985) which defies explanation through the maximization of CRRA or CARA utilities parameterized by the risk aversion coefficient. In response, Epstein & Zin (1989) proposed utility functions that disentangle the coefficient of risk aversion from the elasticity of substitution.[9] Our findings suggest that self-progressive models allow for specifying multiple behavioral characteristics separately. Our following example illustrates another economic scenario in a similar vein.

Example 2. *(Similarity-based choice)* Let (m, p) denote a *lottery* giving a monetary prize $m \in (0, 1]$ with probability $p \in (0, 1]$ and the prize 0 with the remaining probability. Consider a population of agents choosing from binary lottery sets[10] such that each agent i has a *perception of similarity* described by (ϵ^i, δ^i) with $\delta^i \geq \epsilon^i$ as follows: for each $t_1, t_2 \in (0, 1]$, "t_1 is similar to t_2" if $|t_1 - t_2| < \epsilon^i$ and "t_1 is different from t_2" if $|t_1 - t_2| > \delta^i$.

[9] Another explanation based on agents' choices is Benartzi & Thaler (1995)'s *myopic loss aversion* that combines *loss aversion*–a greater sensitivity to losses than to gains–and a tendency to evaluate outcomes more frequently. Since two parameters should be specified separately, population heterogeneity explained by these models may not be consistent with a fixed set of choice functions ordered according to a single characteristic.

[10] One can consider a smallest monetary unit of account and probability differences resulting in a finite domain.

Then, in the vein of Rubinstein (1988), to choose between two lotteries (m_1, p_1) and (m_2, p_2), agent i first checks if "m_1 is similar to m_2 and p_1 is different from p_2", or vice versa.[11] If one of these two statements is true, for instance, m_1 is similar to m_2 and p_1 is different from p_2, then the probability dimension becomes the decisive factor, and i chooses the lottery with the higher probability. Otherwise, each agent chooses the lottery with a higher *expected monetary payoff*, which derives the primitive ordering $>$. By taking the rational assessment as the primitive ordering, the analyst seeks to describe the population heterogeneity emanating from different levels of bounded rationality. The question then arises: Does the similarity-based choice model always provide for progressive representations? Next, we show that the answer is affirmative.

The set of similarity-based choice functions, μ, together with the domination relation, $\rhd$, generated from $>$ is a lattice. In that, for each pair of similarity-based choice functions described by (ϵ^i, δ^i) and (ϵ^j, δ^j), their join and meet are the choice functions that can be described by perceptions of similarity $(min(\epsilon^i, \epsilon^j), max(\delta^i, \delta^j))$ and $(max(\epsilon^i, \epsilon^j), min(\delta^i, \delta^j))$. Then, it follows from our Theorem 1 that the set of similarity-based choice functions is self-progressive. Thus the model always provides for progressive representations.

2 Self-progressive choice models

Let X be the **alternative set** with n elements. A **choice set** S is a subset of X containing at least two alternatives. The **choice domain** Ω is a nonempty collection of choice sets

[11] Rubinstein (1988) additionally requires one of these two statements be true. The slight difference is that our "t_1 is different from t_2" statement implies the negation of "t_1 is similar to t_2", while the converse does not necessarily hold. Both versions of the procedure provide explanations to the Allais paradox.

allowing for limited data sets. A **choice function** is a mapping $c : \Omega \to X$ such that for each $S \in \Omega$, we have $c(S) \in S$. A **choice model** μ is a nonempty set of choice functions. We consider two choice procedures with possibly different formulations as equivalent if these procedures are observationally indistinguishable in the revealed preference framework, that is, two choice procedures rationalize the same set of choice functions.

A **random choice function** (RCF) ρ assigns each choice set $S \in \Omega$ a probability measure over S. We denote by $\rho_x(S)$ the probability that alternative x is chosen from choice set S. A (deterministic) choice function can be represented by an $|\Omega| \times |X|$ matrix with rows indexed by the choice sets and columns indexed by the alternatives, and entries in $\{0, 1\}$ such that each row has exactly one 1. For each $(S, x) \in \Omega \times X$, having 1 in the entry corresponding to row S and column x indicates that x is chosen in S. Similarly, an RCF can be represented by an $|\Omega| \times |X|$ matrix having entries in $[0, 1]$ such that the sum of the entries in each row is 1. For each RCF and each pair $(S, x) \in \Omega \times X$, the associated entry indicates the probability that x is chosen in S.

It follows from Birkhoff-von Neumann Theorem (Birkhoff 1946, Von Neumann 1953) that each RCF can be represented as a probability distribution over a set of deterministic choice functions. However, this representation is not necessarily unique. Let $\Delta(\mu)$ be the **random choice model** associated with a choice model μ, which is the set of RCFs that can be represented as a probability distribution over the elements of μ.

For each choice set $S \in \Omega$, a **primitive ordering** $>_S$ is a complete, transitive, and asymmetric binary relation over S. We write $\geq_S$ for its union with the equality relation. Then, we obtain the partial order $\rhd$ from the primitive orderings such that for each pair of choice functions c and c', we have $c \rhd c'$ if and only if $c(S) \geq_S c'(S)$ for each $S \in \Omega$, and $c(S) \neq c'(S)$ for some $S \in \Omega$. We write $c \unrhd c'$ if $c \rhd c'$ or $c = c'$.

Definition. Let $\rhd$ be the partial order over choice functions obtained from the primitive orderings $\{>_S\}_{S\in\Omega}$. Then, a choice model μ is **self-progressive** with respect to $\rhd$ if each RCF $\rho \in \Delta(\mu)$ can be uniquely represented as a probability distribution over a set of choice functions $\{c^1, \ldots, c^k\} \subset \mu$ such that $c^1 \rhd c^2 \cdots \rhd c^k$.

As formulated by Filiz-Ozbay & Masatlioglu (2023), an RCFρ has a **progressive representation** if it can be represented as a probability distribution over a set of choice functions $\{c^1, \ldots, c^k\} \subset \mu$ such that $c^1 \rhd c^2 \cdots \rhd c^k$. To see that a progressive representation is unique whenever it exists, consider the $\rhd$-best choice function c^1 in a progressive representation. Note that c^1 chooses the $>_S$-best alternative that is assigned positive probability by ρ in each $S \in \Omega$. Therefore, the probability weight of c^1 is determined uniquely as the lowest probability of $c^1(S)$ being chosen from any S. Repeating this argument shows that the progressive representation is unique.

2.1 Equivalence between self-progressive models and lattices

Let $\{>_S\}_{S\in\Omega}$ be the primitive orderings and $\rhd$ be the associated partial order over choice functions. For each pair of choice functions c and c', their *join* (*meet*) is the choice function $c \vee c'$ $(c \wedge c')$ that chooses from each choice set S, the $>_S$-best(worst) alternative among the ones chosen by c and c' at S. Then, for each choice model μ, the pair $\langle \mu, \rhd \rangle$ is a **lattice** if for each pair of choice functions c and c' in μ, their join $c \vee c'$ and meet $c \wedge c'$ are in contained μ as well.

Theorem 1. *Let μ be a choice model and $\rhd$ be the partial order over choice functions obtained from the primitive orderings $\{>_S\}_{S\in\Omega}$. Then, μ is self-progressive with respect to $\rhd$ if and only if the pair $\langle \mu, \rhd \rangle$ is a lattice.*

To see that the *only if* part holds, let $c, c' \in \mu$. Then, consider the $\texttt{RCF}\rho$ such that for each $S \in \Omega$, $c(S)$ or $c'(S)$ is chosen evenly. Note that ρ has a unique progressive representation in which only $c \vee c'$ and $c \wedge c'$ receive positive probability. Since μ is self-progressive, it follows that $c \vee c' \in \mu$ and $c \wedge c' \in \mu$.

As for the *if* part, suppose that $\langle \mu, \rhd \rangle$ is a lattice, and let $\rho \in \Delta(\mu)$. Next, we present our **uniform decomposition procedure**, which yields the progressive random choice representation for ρ with respect to $\rhd$. Figure 2 demonstrates the procedure.

Step 1: For each choice set S, let $\rho^+(S) = \{x \in S : \rho(x, S) > 0\}$. Partition the $(0, 1]$ interval into intervals $\{I_{Sx}\}_{\{x \in \rho^+(S)\}}$ such that each interval $I_{Sx} = (l_{Sx}, u_{Sx}]$ is half open with length $\rho(x, S)$, and for each $x, y \in \rho^+(S)$ if $x >_S y$, then l_{Sx} is less than l_{Sy}.

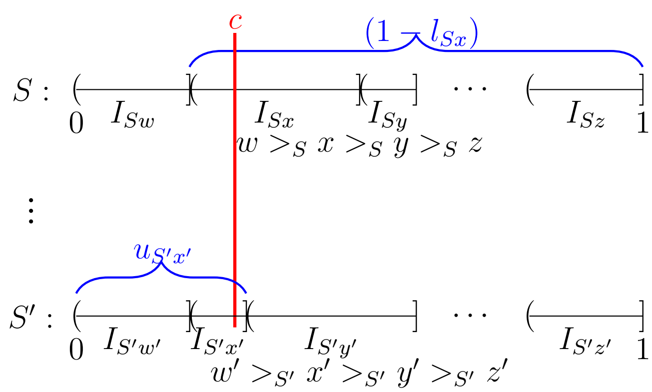


Figure 2

Step 2: Pick a real number $r \in (0, 1]$ according to the Uniform distribution on $(0, 1]$. Then, for each choice set and alternative pair (S, x), let $c(S) = x$ if and only if $r \in I_{Sx}$. It is clear that this procedure gives us a unique probability distribution over a set of choice functions $\{c^i\}_{i=1}^k$ such that $c^1 \rhd c^2 \cdots \rhd c^k$.[12] Next, we will show that $\{c^i\}_{i=1}^k \subset \mu$.

[12] See Theorem 1 by Filiz-Ozbay & Masatlioglu (2023) for an elaborate proof of this fact. It is easy to see that this procedure is applicable even if the choice space is infinite. In a contemporary study, Petri (2023) independently extends Theorem 1 by Filiz-Ozbay & Masatlioglu (2023) to infinite choice spaces.

Lemma 1. *Let μ be a choice model such that $\langle \mu, \rhd \rangle$ is a lattice. Let $c \in \mathcal{C}$ be a choice function. If for each given $S, S' \in \Omega$, there is a choice function $c^* \in \mu$ such that $c^*(S) = c(S)$ and $c^*(S') = c(S')$, then $c \in \mu$.*

Proof. The result is obtained by applying the following observation inductively. Consider any $\mathbb{S} \subset \Omega$ containing at least three choice sets. Let $c_1, c_2, c_3 \in \mu$ be such that for each $i \in \{1, 2, 3\}$, there exists at most one $S_i \in \mathbb{S}$ with $c_i(S_i) \neq c(S_i)$. Suppose that for each $i, j \in \{1, 2, 3\}$, if such S_i and S_j exist, then $S_i \neq S_j$. Now, for each $S \in \mathbb{S}$, we have $c(S)$ is chosen by the choice function $(c_1 \wedge c_2) \vee (c_1 \wedge c_3) \vee (c_2 \wedge c_3) \in \mu$. To see this, let $S \in \mathbb{S}$, and note that there exist at least two $i, j \in \{1, 2, 3\}$ such that $c_i(S) = c_j(S) = c(S)$. Assume without loss of generality that $i = 1$ and $j = 2$. Now, if $c(S) \geq_S c_3(S)$, then we get $c(S) \vee c_3(S) \vee c_3(S) = c(S)$; if $c_3(S) >_S c(S)$, then we get $c(S) \vee c(S) \vee c(S) = c(S)$. $\square$

Proof of Theorem 1. We proved the *only if* part. For the *if* part, let c^r be a choice function that is assigned positive probability in the uniform decomposition procedure. We show that $c^r \in \mu$ by using Lemma 1. To see this, let $S, S' \in \Omega$ such that $x = c^r(S)$ and $x' = c^r(S')$. We will show that there exists $c^* \in \mu$ such that both $c^*(S) = x$ and $c^*(S') = x'$.

First, as demonstrated in Figure 2, we have $(1 - l_{Sx}) + u_{S'x'} > 1$. Thinking probabilistically, this means that making a choice that is worse than x in S and better than x' in S' are not mutually exclusive events. Since $\rho \in \Delta(\mu)$, it follows that there exists $c_1 \in \mu$ such that $c_1(S) \leq_S x$ and $c_1(S') \geq_{S'} x'$. Symmetrically, since $(1 - l_{S'x'}) + u_{Sx} > 1$, there exists $c_2 \in \mu$ such that $c_2(S) \geq_S x$ and $c_2(S') \leq_{S'} x'$.

Next, consider the set $\{c \in \mu : x \geq_S c(S)\}$ and let c_x be its join. Similarly, consider $\{c \in \mu : x' \geq_{S'} c(S')\}$ and let $c_{x'}$ be its join. Since $\rho \in \Delta(\mu)$ and c^r is assigned positive probability in the uniform decomposition procedure, μ must contain a choice function choosing x from S and possibly a different one choosing x' from S'. Since $\langle \mu, \rhd \rangle$ is a lattice, it follows that $c_x(S) = x$ and $c_{x'}(S') = x'$. Moreover, c_1 is a member of the former set, while c_2 is a member of the latter one. Now, define $c^* = c_x \wedge c_{x'}$. Then, $c^*(S) = x$, since $c_x(S) = x$ and $c_{x'}(S) \geq_S c_2(S) \geq_S x$. Similarly, $c^*(S') = x'$, since $c_{x'}(S') = x'$ and $c_x(S') \geq_{S'} c_1(S') \geq_{S'} x'$. Finally, $c^* \in \mu$ since $\langle \mu, \rhd \rangle$ is a lattice containing c_x and $c_{x'}$. $\square$

3 Examples and discussion

3.1 Rational choice and chain lattices

We first observe that the rational choice model fails to be self-progressive. To see this, let $X = \{a, b, c\}$ and $\Omega = \{X, \{a, b\}, \{a, c\}, \{b, c\}\}$. Suppose that each primitive ordering is obtained by restricting the ordering $a > b > c$ to a choice set. Figure 3 demonstrates the associated choice functions lattice in which each array specifies the chosen alternatives respectively. The rational choice functions (dark-colored ones) fail to form a lattice. In that, each light-colored choice function is a join or meet of a rational choice function.

We can use the equivalence between self-progressiveness and lattices as a guide to restrict or extend rational choice model to be self-progressive. In this vein, a particularly simple lattice is a *chain lattice,* which is a set of choice functions $\{c_i\}_{i=1}^k$ that are comparable: $c_1 \rhd c_2 \cdots \rhd c_n$. Suppose that each primitive ordering $>_S$ is obtained by restricting the ordering $>_X$ to the choice set S. Then, there is a one-to-one correspondence between the chain lattices of rational choice model and the preferences with *single-crossing property* defined by Apesteguia, Ballester & Lu (2017).[13] To see this, let $\mu = \{c_i\}_{i=1}^k$ be a

[13]Additionally, if the choice domain Ω contains every choice set, then every lattice $\langle \mu, \rhd \rangle$ is a chain lattice. This is not true for a general domain of choice sets. For a simple example, suppose that the choice

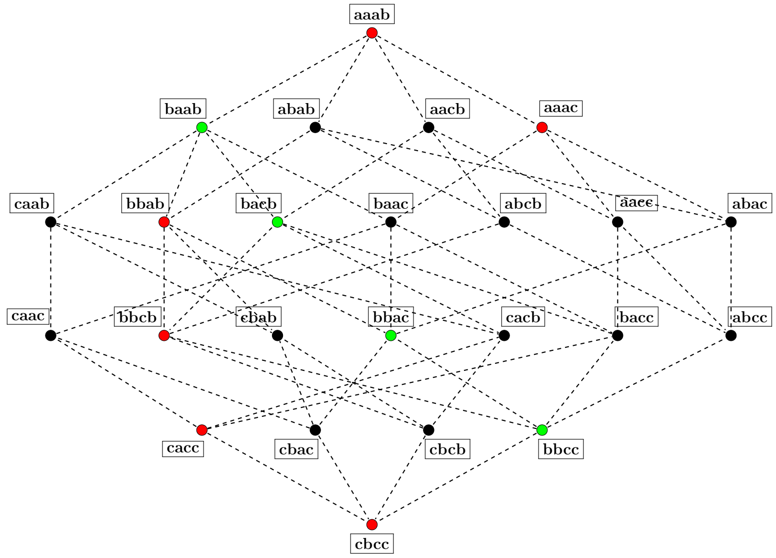


Figure 3: The choice functions lattice.

choice model consisting of choice functions rationalized by maximization of preferences $\{\succ_i\}_{i=1}^k$. Then, $\{\succ_i\}_{i=1}^k$ is *single-crossing with respect to* $>_X$ means: for each alternative pair $x >_X y$, if $x \succ_i y$, then $x \succ_j y$ for every $i > j$. It is easy to see that $\langle \mu, \rhd \rangle$ is a chain lattice if and only if $\{\succ_i\}_{i=1}^k$ is single-crossing with respect to $>_X$.[14]

Apesteguia, Ballester & Lu (2017) present economic examples of rational choice functions that form chain lattices. Filiz-Ozbay & Masatlioglu (2023) present choice models that are not rational. However, these choice models also form chain lattices because the choice functions are ordered according to a single behavioral characteristic.

domain consists of disjoint binary choice sets. Then, every choice function is rational, thus every sublattice of choice functions is a set of rational choice functions.

[14]See also Lemma 1 by Filiz-Ozbay & Masatlioglu (2023).

3.2 Beyond single characteristic and chain lattices

It follows from our Theorem 1 that self-progressive choice models are not limited to chain lattices, thus capture multiple behavioral characteristics of agents. We demonstrate this point with additional examples similar to our Example 2.

Example 3. Let $\mathcal{P}$ be a set of faulty preferences that are single-crossing with respect to the accurate preference $>$. Then, a choice function $c \in \mu$ if for each choice set S, the alternative $c(S)$ is the $\succ_S$-maximal one in S for some $\succ_S \in \mathcal{P}$. If S is a subset of S', then $\succ_S$ is more aligned with $>$ (less faulty) than $\succ_{S'}$. Note that μ is self-progressive with respect to the comparison relation obtained from $>$, since the join and meet of each $c^i, c^j \in \mu$ are the choice functions obtained by maximization of the preferences $max(\{\succ^i_S, \succ^j_S\}, \geq)$ and $min(\{\succ^i_S, \succ^j_S\}, \geq)$.

Example 4. Consider a population with primitive orderings $\{>_S\}_{S \in \Omega}$ in which each agent i has the same preference relation $\succ^*$, but a possibly different *threshold alternative* x^i_S for each choice set S. Then, for given choice set S, agent i chooses the $\succ^*$-best alternative in the consideration set $\{x \in S : x \geq_S x^i_S\}$. Let μ be the set of associated choice functions. Then, $\langle \mu, \rhd \rangle$–where $\rhd$ is obtained from $\succ^*$–is a lattice, since the join and meet of each $c^i, c^j \in \mu$ are the choice functions described by threshold alternatives $max(\{x^i_S, x^j_S\}, \geq_S)$ and $min(\{x^i_S, x^j_S\}, \geq_S)$.[15]

[15] As a special case, consider agents who faces temptation with limited willpower formulated as by Masatlioglu, Nakajima & Ozdenoren (2020). Each agent i chooses the alternative that maximizes the common *commitment ranking* u from the set of alternatives where agent i overcomes *temptation*, represented by v^i, with his *willpower stock* w^i. Suppose that the primitive orderings are aligned with the commitment ranking u. Then, for each choice set S, let the threshold alternative x^i_S be the $>_S$-worst one such that $v^i(x) - max_{z \in S} v^i(z) \leq w^i$. As demonstrated by Filiz-Ozbay & Masatlioglu (2023) if we only allow agents' willpower stock to differ, then we obtain a choice model forming a chain lattice.

4 Minimal self-progressive extension of rational choice

We will follow the guide provided by Theorem 1 to discover the "minimal" self-progressive extension of the rational choice model. We assume that there is a single primitive ordering $>$ rankings of which reflect alternatives' "accurate values" and Ω contains every choice set. The comparison relation $\triangleright$ over choice functions is obtained from $>$ as usual.

An extension is minimal if we are parsimonious in adding nonrational choice functions so that each choice model containing all rational choice functions and is contained in the extension fails to be self-progressive with respect to $\triangleright$.[16] Next, we characterize the minimal self-progressive extension of the rational choice model in terms of two choice axioms. Figure 4 demonstrates the minimal extension when there are three alternatives.

Theorem 2. *Let μ^θ be the minimal self-progressive extension of the rational choice model with respect to $\triangleright$. Then, a choice function $c \in \mu^\theta$ if and only if for each $S \in \Omega$ and $x \in S$,*

$\theta 1.$ *if $c(S) > x$ then $c(S \setminus \{x\}) \geq c(S)$, and*

$\theta 2.$ *if $x > c(S)$ then $c(S) \geq c(S \setminus \{x\})$.*[17]

Proof. Please see Section 7.1 in the Appendix. □

Axioms $\theta 1$ and $\theta 2$ require a more valuable (or the same) alternative be chosen whenever we remove alternatives that are less valuable than the chosen one, or add alternatives that are more valuable than the chosen one. Along these lines–in an attempt to unravel the choice overload phenomena–Chernev & Hamilton (2009) experimentally demonstrate that consumers' selection among choice sets is driven by the value of the

[16]It follows from Theorem 1 that the minimal extension of any choice model is unique.

[17]*Independence from preferred alternative* formulated by Masatlioglu, Nakajima & Ozdenoren (2020) similarly require choice remain unchanged whenever unchosen better options are removed.

alternatives constituting the choice sets. In that, the smaller choice set is more likely to be selected when the value of the alternatives is high than when it is low. The proof of Theorem 2 demonstrates how to use Theorem 1 and Lemma 1 to obtain similar results.

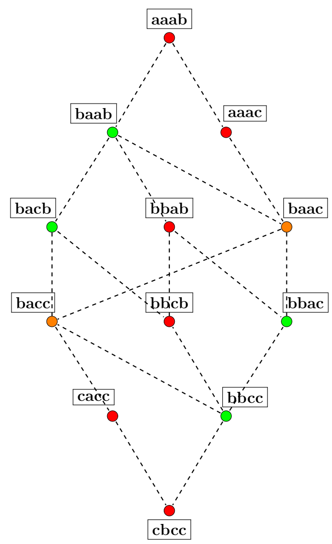


Figure 4: A demonstration of $\langle \mu^{\theta}, \rhd \rangle$, where $X = \{a, b, c\}$, $\Omega = \{X, \{a, b\}, \{a, c\}, \{b, c\}\}$, and each array specifies the respective choices. The rational choice functions are colored in red, their joins and meets are colored in green, and the additional ones–obtained as a join or meet of the previous ones–are colored in orange.

4.1 Identification from random choice

Let $\mu^{\theta}(>)$ be the minimal self-progressive extension of the rational choice model with respect to the primitive ordering $>$. Next, we ponder how to identify if an RCF ρ is representable as a probability distribution over choice functions contained in $\mu^{\theta}(>)$. This question is of economic interest for at least two reasons. Firstly, the resulting axioms facilitate identification of the model from random choice data. Secondly, it is important from a normative perspective to determine if the underlying axioms carry over the choice overload interpretation found in the deterministic model.

The characterizing axioms turn out to be the probabilistic counterparts of θ^1 and θ^2. Axioms $r\theta 1$ and $r\theta 2$ require the probability of choosing an alternative that is as valuable as a fixed alternative y goes up (or remains the same), whenever we remove alternatives that are less valuable than y, or add alternatives that are more valuable than y.

To state the result, we first define the **cumulative random choice function** (`CRCF`) $\rho^{\uparrow}$ associated to a given `RCF` ρ as follows: for each $S \in \Omega$ and $y \in S$, $\rho^{\uparrow}(y, S)$ is the total probability of choosing an alternative that is more valuable than y in S, i.e. $\rho^{\uparrow}(y, S) = \sum_{x \in S : x > y} \rho(x, S)$.

Proposition 1. *An RCF ρ is representable as a probability distribution over choice functions contained in $\mu^{\theta}(>)$ if and only if for each $S \in \Omega$ and $x, y \in S$ we have*

$r\theta 1$. *if $y > x$ then $\rho^{\uparrow}(y, S \setminus \{x\}) \geq \rho^{\uparrow}(y, S)$, and*

$r\theta 2$. *if $x > y$ then $\rho^{\uparrow}(y, S) \geq \rho^{\uparrow}(y, S \setminus \{x\})$.*

Our next observation paves the way for proving Proposition 1. For this observation, we use two classical results from the integer-programming literature, namely the ones developed by Hoffman & Kruskal (2010) and Heller & Tompkins (1956).[18] We state these results in Appendix 1.

Lemma 2. *The set of CRCFs $\rho^{\uparrow}$ that satisfy $r\theta 1$ and $r\theta 2$ forms a polytope whose extreme points are $\{0, 1\}$-valued.*

Proof. Please see Section 1 in the Appendix. □

Proof of Proposition 1. Only if part: Suppose that ρ is an `RCF` that is representable as a probability distribution over choice functions contained in $\mu^{\theta}(>)$. Let $\mathbb{X} = \{(x, S) :$

[18] This technique has been previously used by Dogan & Yildiz (2023) to obtain a similar result in choice theory.

$S \in \Omega \;\&\; x \in S\}$. Then, for each choice function c let $c^{\upharpoonleft} : \mathbb{X} \to \{0,1\}$ denote the associated CRCF. Note that a choice function $c \in \mu^{\theta}(>)$–satisfies $\theta 1$ and $\theta 2$–if and only if the associated CRCF $c^{\upharpoonleft}$ satisfies $r\theta 1$ and $r\theta 2$. Since, the set of CRCFs that satisfy $r\theta 1$ and $r\theta 2$ is a convex set, it follows that ρ satisfies $r\theta 1$ and $r\theta 2$.

If part: Suppose that ρ is an RCF with an CRCF $\rho^{\upharpoonleft}$ that satisfies $r\theta 1$ and $r\theta 2$. Then, it follows from Lemma 2 that $\rho^{\upharpoonleft}$ can be represented as a convex combination of the CRCFs $c^{\upharpoonleft} : \mathbb{X} \to \{0,1\}$ that satisfy $r\theta 1$ and $r\theta 2$. Next, let $c^{\upharpoonleft}$ be such an CRCF. For each $S \in \Omega$, either (i) $c^{\upharpoonleft}(x, S) = 0$ for every $x \in S$ or (ii) there exists $x^* \in S$ such that for each $x \in S$, we have $c^{\upharpoonleft}(x, S) = 1$ if and only if $x^* \geq x$. Therefore, there is a one-to-one correspondence between the CRCF $c^{\upharpoonleft}$ and the choice function c defined as follows: for each $S \in \Omega$, if (i) holds then $c(S)$ is the $>$-best alternative in S, and if (ii) holds then $c(S) = x^*$. Now, since $\rho^{\upharpoonleft}$ is representable as a convex combination of the CRCFs $c^{\upharpoonleft}$ that satisfy $r\theta 1$ and $r\theta 2$, we conclude that ρ is representable as a convex combination of the corresponding choice functions c that satisfy $\theta 1$ and $\theta 2$. $\square$

4.2 Identification of the primitive ordering

So far, we assumed that the analyst has specified the primitive ordering. We next focus on how to identify a primitive ordering that renders a choice overload representation to a choice model. Formally, a primitive ordering $>$ *renders a choice overload representation* to a choice model μ if $\mu \subset \mu^{\theta}(>)$. Thus, the observed choice functions can be interpreted as the sample choice behavior of a population whose choices comply with $\theta 1$ and $\theta 2$ according to the primitive ordering $>$.

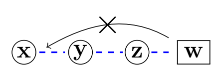
We first show how to infer from a given choice model that an alternative lies "between" two other alternatives according to every primitive ordering that renders a choice overload representation to the choice model. Let μ be a given choice model and $x, y, z \in X$ be a triple. Then, y is **revealed to be between** x and z–denoted by $y\ \mathcal{B}_\mu\ \{x, z\}$–if there exists a choice function $c \in \mu$ such that $c(S) = y$ and $c(S \setminus z) = x$ for some choice set S. Then, we have $y\ \mathcal{B}_\mu\ \{x, z\}$ if and only if $x > y > z$ or $x < y < z$ for every primitive ordering $>$ that renders a choice overload representation to μ. We refer to $\mathcal{B}_\mu$ as the **betweenness relation** associated to the choice model μ.

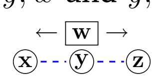
We next introduce conditions structuring the betweenness relation associated with a choice model. We show that these conditions are necessary and sufficient for the existence and uniqueness of a primitive ordering that renders a choice overload representation to the choice model. As a corollary, we observe that the primitive ordering associated to the minimal extension of rational choice functions is identified unique up to its inverse. Let μ be a choice model and $\mathcal{B}_\mu$ be the associated betweenness relation.

$B1$. Each triple $x, y, z \in X$ appear in at most one $\mathcal{B}_\mu$-comparison.

$sB1$. Each triple $x, y, z \in X$ appear in exactly one $\mathcal{B}_\mu$-comparison.

For axioms $B2$ and $B3$, let $x, y, z, w \in X$ be distinct and $y\ \mathcal{B}_\mu\ \{x, z\}$.

$B2$. If $z\ \mathcal{B}_\mu\ \{x, w\}$, then it is not $w\ \mathcal{B}_\mu\ \{x, y\}$. 

$B3$. If x, y, w and y, z, w appear in $\mathcal{B}_\mu$-comparison, then $y\ \mathcal{B}_\mu\ \{x, w\}$ or $y\ \mathcal{B}_\mu\ \{z, w\}$ but not both. 

We can interpret $B1$ as an "asymmetry" and $B2$ as an "3-acylicity" requirement for the betweenness relation. In the vein of *negative transitivity* (Kreps 1988), $B3$ requires that if y lies between x and z, then each w should lie either on the x- or z-side of y.

Finally, $sB1$ strengthes $B1$ by requiring that $x\ \mathcal{B}_\mu\ \{y,z\}$, $y\ \mathcal{B}_\mu\ \{x,z\}$ or $z\ \mathcal{B}_\mu\ \{y,x\}$ for each triple x, y, z. Since this condition is to be satisfied by a choice model, different choice functions may provide for different triples related according to the betweenness relation. Thus, we can interpret $sB1$ as a "richness" requirement for the choice model, which may be hard for a single choice function to satisfy.

Theorem 3. *Let μ be a choice model and $\mathcal{B}_\mu$ be the associated betweenness relation. Then,*

i. $\mathcal{B}_\mu$ satisfies $B1 - B3$ if and only if there is a primitive ordering $>$ such that μ is contained in $\mu^\theta(>)$.

ii. $\mathcal{B}_\mu$ satisfies $sB1$ and $B3$ if and only if there is a unique (up to its inverse) primitive ordering $>$ such that μ is contained in $\mu^\theta(>)$.

For the proof, it is critical to identify an ordering that **agrees with the betweenness relation**, in the sense that if $y\ \mathcal{B}_\mu\ \{x,z\}$ then $x > y > z$ or $x < y < z$ for each triple $x, y, z \in X$. Betweenness is a ternary relation, interest in which stems from their use in axiomatizations of geometry. For example, Huntington & Kline (1917) proposed eleven different sets of axioms to characterize the usual betweenness on a real line; most of which can be translated to replace $sB1$ and $B3$ jointly. Our $sB1$ appears almost directly in these axiomatizations, whereas $B3$ is most similar to the axioms used in more succinct characterizations provided by Huntington (1924) and Fishburn (1971).[19]

To prove part i of Theorem 3, we have $B1$ instead of $sB1$. To fill the gap, we use a recent result by Biró, Lehel & Tóth (2023) who provide a unified view to existing results. They show that if there is an agreeing ordering on every four elements, then there is an agreeing ordering for the whole set.[20] To use their result, in Lemma 3, we show that

[19]See axiom 9 used by Huntington (1924) and axiom $A3$ used by Fishburn (1971).

[20]This result is in line with the fact that no axiom appeared in the aforementioned characterizations uses more than four elements.

$B1 - B3$ suffice for the existence of orderings that "locally" agree with our betweenness relation. We use a characterization by Fishburn (1971) to prove Lemma 3.

Lemma 3. *Let μ be a choice model such that the associated betweenness relation $\mathcal{B}_\mu$ satisfies $B1 - B3$. Then, for each distinct $x, y, z, w \in X$, there is an ordering $>_L$ such that for each triple $a, b, c \in \{x, y, z, w\}$, if $b \ \mathcal{B}_\mu \ \{a, c\}$ then $a >_L b >_L c$ or $a <_L b <_L c$.*

Proof. Please see Section 3 in the Appendix. □

Proof of Theorem 3. Part i: It directly follows from $\theta1$ and $\theta2$ that the if part holds. To prove the only if part, suppose that $\mathcal{B}_\mu$ satisfies $B1 - B3$. Then, it follows from our Lemma 3 and Theorem 1 by Biró, Lehel & Tóth (2023) that there is an ordering $>$ over X such that for each triple $x, y, z \in X$, if $y \ \mathcal{B}_\mu \ \{x, z\}$ then $x > y > z$ or $x < y < z$. Thus, we conclude that μ is contained in $\mu^\theta(>)$.

Part ii: Since by $sB1$, each triple $x, y, z \in X$ appears in an $\mathcal{B}_\mu$-comparison, it follows from the proof of part i that there is an ordering $>$ over X such that for each triple $x, y, z \in X$, we have $y \ \mathcal{B}_\mu \ \{x, z\}$ if and only if $x > y > z$ or $x < y < z$. Then, by the only if part, μ is contained in $\mu^\theta(>)$. By the if part, $>$ and its inverse are the only such orderings. □

A choice model can comprise a single choice function as well as a collection of choice functions representing the revealed choice behavior of a population. To best of our knowledge, identifying primitives from a choice model in this way is novel. Pursuing this approach further, suppose that a choice model μ coincides with the minimal extension of rational choice functions with respect to a primitive ordering $>$, i.e. $\mu = \mu^\theta(>)$. Then, it follows from our Theorem 3 that we can identify the underlying primitive ordering unique up to its inverse.

Corollary 1. *Let μ be a choice model. Then, $\mu = \mu^\theta(>)$ and $\mu = \mu^\theta(>')$ if and only if $>'$ is the inverse of $>$.*

Proof. Suppose that $\mu = \mu^\theta(>)$ for some primitive ordering $>$. Then, we show that μ satisfies $sB1$. To see this, let $x, y, z \in X$ be a triple such that $x > y > z$. Consider the choice function c such that $c(\{x, y, z\}) = y$ and $c(\{x, y\}) = x$, and $c(S)$ is the $>$-best alternative in S for every other choice set S. Since c satisfies $\theta1$ and $\theta2$ according to $>$, we have $c \in \mu$ and $y\ \mathcal{B}_\mu\ \{x, z\}$. Thus, μ satisfies $sB1$ and the conclusion follows from part ii of Theorem 3. □

5 Universally self-progressive choice models

A natural question is whether there are choice models that yield unique orderly representations for any primitive orderings. We next define and examine this strong condition.

Definition. *A choice model μ is **universally self-progressive** if μ is self-progressive with respect to any partial order $\rhd$ obtained from any set of primitive orderings $\{>_S\}_{S\in\Omega}$.*

To characterize the universally self-progressive choice models, we first offer a fresh perspective about choice functions. A choice function can be interpreted as a complete contingent plan to be implemented upon observing available alternatives.[21] Then, suppose that a population of agents evaluate choice functions via a common *value function*, which can be thought of as an indirect utility function associated with the problem of optimally adopting a choice function. The population is homogeneous in the sense that each agent evaluates choice functions via the same value function. The unique source

[21] Here, a choice function is analogous to a "worldview" as described by Bernheim, Braghieri, Martínez-Marquina & Zuckerman (2021) who offer a dynamic model of endogenous preference formation.

of heterogeneity is the maximizers' multiplicity. Then the question arises: What sort of choice heterogeneity allows for universal self-progressiveness? We show that additive separability of the value function over *set contingent utilities* is the answer. Since this characterization may be less direct for identification, Proposition 2 additionally presents an equivalent ordinal condition strengthening the lattice requirement.

For each $S \in \Omega$ and $x \in S$, let $U(x, S)$ be the **set contingent utility** of choosing x. In addition to the intrinsic utility of alternative x that may be menu independent, $U(x, S)$ can accommodate the likelihood of S being available or the temptation cost due to choosing x in the presence of more tempting alternatives.[22]

Proposition 2. *The following assertions are equivalent.*

I. A choice model μ is universally self-progressive.

II. If c^* is obtained as a ***mixture*** *of some $c, c' \in \mu$ in the sense that $c^*(S) \in \{c(S), c'(S)\}$ for every $S \in \Omega$, then $c^* \in \mu$ as well.*

III. There is a set contingent utility function $U(\cdot, S)$ for each $S \in \Omega$ such that μ is the set of choice functions that maximize their sum, that is $\mu = \text{argmax}_{c \in \mathcal{C}} \sum_{S \in \Omega} U(c(S), S)$.

Proof. If I then II: By contradiction, let $c_1, c_2 \in \mu$ such that $c(S) \in \{c_1(S), c_2(S)\}$ for every $S \in \Omega$, but $c \notin \mu$. Then, for each $S \in \Omega$, define the primitive ordering $>_S$ such that $c(S)$ is highest-ranked. Thus, we have $c = c_1 \vee c_2$, but $\langle \mu, \rhd \rangle$ is a not a lattice. By Theorem 1, this contradicts that I holds. If II then I: Since meet and join are special mixtures, $\langle \mu, \rhd \rangle$ is a lattice for any partial order $\rhd$ obtained from a set of primitive orderings. Then, it follows from Theorem 1 that I holds. If II then III: Define the set contingent utilities for each $S \in \Omega$ such that $U(x, S) = 1$ if there exists $c \in \mu$ with $c(S) = x$, and $U(x, S) = 0$

[22] For example, in the vein of Gul & Pesendorfer (2001), one can set $U(x, S) = u(x) + v(x) - max_{z \in S} v(z)$, where u represents the *commitment* ranking and v represents the *temptation ranking*.

otherwise. Since μ satisfies II, a choice function $c \in \mu$ if and only if $U(x, S) = 1$ for each $S \in \Omega$. It follows that μ is the set of choice functions that maximize $\sum_{S\in\Omega} U(c(S), S)$. Thus, III holds. If III then II: If two choice functions c_1 and c_2 maximize the sum of a collection of set contingent utilities, so does any mixture of c_1 and c_2. Thus, II holds. $\square$

Proposition 2 shows how to modify a choice model for universal self-progressiveness, while reflecting its demanding nature. To see this, consider a choice model μ consisting of two choice functions rationalized by maximizing preference relations $\succ_1$ and $\succ_2$. For fixed primitive orderings, we can make μ self-progressive by adding at most two choice functions. In contrast, to extend μ as being universally self-progressive we must add every choice function choosing the $\succ_1$- or $\succ_2$-maximal alternative in each choice set. More generally, if the choice domain contains every choice set, then to extend the rational choice model into a universally self-progressive one, we must add every choice function. In contrast, Theorem 2 showed that the minimal self-progressive extension of the rational choice is a structured model. We finally present Example 5 demonstrating that Proposition 2 facilitates verifying if a choice model is universally self-progressive.

Example 5. *Kalai, Rubinstein & Spiegler (2002)* Let $\{\succ_k\}_{k=1}^K$ be a K-tuple of strict preference relations on X. A choice function $c \in \mu$ if for each $S \in \Omega$, the alternative $c(S)$ is the $\succ_k$-maximal one in S for some k. To see that μ is universally self-progressive, define $U(x, S) = 1$ if x is the $\succ_k$-maximal alternative in S for some k; and $U(x, S) = 0$ otherwise. It follows that μ is the set of choice functions that maximize $\sum_{S\in\Omega} U(c(S), S)$. To see that every universally self-progressive choice model is not representable in this way, let $U(x, S) = 1$ and $U(x, T) = 0$ for a pair of choice sets S and T with $x \in T \subset S$. Then, there is a strict preference relation $\succ_k$ such that x is the $\succ_k$-maximal alternative in S. Thus, we obtain $c \in \mu$ with $c(T) = x$, contradicting that c maximizes $\sum_{S\in\Omega} U(c(S), S)$.

6 Final comments

We have explored a novel approach to analyze heterogeneity in the aggregate choice behavior of a population. Our emphasis is on an analyst who seeks to select a choice model to explain observed random choice behavior. As a criterion to guide the model selection process, we introduced *self-progressiveness* ensuring that each aggregate choice behavior explained by the model has a unique orderly representation within the model.

As an advantage of our model-free approach, we obtained a foundational tool for restricting or extending any choice model to be self-progressive. To demonstrate, we characterized the minimal self-progressive extension of the rational choice model. This model provides an experimentally supported explanation for choice overload phenomena, and enables intuitive identification of the primitive ordering. Hence, we observed that, beyond their analytical properties, self-progressive choice models can prove valuable in formulating choice models that elucidate economically relevant phenomena.

References

Apesteguia, J., Ballester, M. A. & Lu, J. (2017), 'Single-crossing random utility models', *Econometrica* **85**(2), 661–674. 3, 7, 8, 14, 15

Benartzi, S. & Thaler, R. H. (1995), 'Myopic loss aversion and the equity premium puzzle', *Quarterly Journal of Economics* **110**(1), 73–92. 8

Bernheim, B. D., Braghieri, L., Martínez-Marquina, A. & Zuckerman, D. (2021), 'A theory of chosen preferences', *American Economic Review* **111**(2), 720–54. 24

Birkhoff, G. (1946), 'Three observations on linear algebra', *Univ. Nac. Tacuman, Rev. Ser. A* **5**, 147–151. 10

Biró, C., Lehel, J. & Tóth, G. (2023), 'Helly-type theorems for the ordering of the vertices of a hypergraph', *Order* pp. 1–18. 22, 23

Chernev, A., Böckenholt, U. & Goodman, J. (2015), 'Choice overload: A conceptual review and meta-analysis', *Journal of Consumer Psychology* **25**(2), 333–358. 5

Chernev, A. & Hamilton, R. (2009), 'Assortment size and option attractiveness in consumer choice among retailers', *Journal of Marketing Research* **46**(3), 410–420. 17

Chew, S. H., Miao, B. & Zhong, S. (2017), 'Partial ambiguity', *Econometrica* **85**(4), 1239–1260. 4

Chiappori, P.-A., Salanié, B., Salanié, F. & Gandhi, A. (2019), 'From aggregate betting data to individual risk preferences', *Econometrica* **87**(1), 1–36. 4

Costa, M., Ramos, P. H. & Riella, G. (2020), 'Single-crossing choice correspondences', *Social Choice and Welfare* **54**(1), 69–86. 7

Dardanoni, V., Manzini, P., Mariotti, M., Petri, H. & Tyson, C. (2022), 'Mixture choice data: Revealing preferences and cognition', *Journal of Political Economy* . 3

Dillenberger, D. & Sadowski, P. (2012), 'Ashamed to be selfish', *Theoretical Economics* **7**(1), 99–124. 4

Dogan, S. & Yildiz, K. (2023), 'Every choice function is pro-con rationalizable', *Operations Research* **71**(5), 1857–1870. 19

Epstein, L. G. & Zin, S. E. (1989), 'Substitution, risk aversion, and the temporal behavior of consumption', *Econometrica* **57**(4), 937–969. 8

Filiz-Ozbay, E. & Masatlioglu, Y. (2023), 'Progressive random choice', *Journal of Political Economy* **131**(3), 716–750. 3, 4, 7, 8, 11, 12, 15, 16

Fishburn, P. C. (1971), 'Betweenness, orders and interval graphs', *Journal of Pure and Applied Algebra* **1**(2), 159–178. 22, 23, 34

Gul, F. & Pesendorfer, W. (2001), 'Temptation and self-control', *Econometrica* **69**(6), 1403–1435. 25

Gul, F. & Pesendorfer, W. (2006), 'Random expected utility', *Econometrica* **74**(1), 121–146. 3

Heller, I. & Tompkins, C. (1956), 'An extension of a theorem of Dantzig's', *Linear inequalities and related systems* **38**, 247–252. 19, 33

Hoffman, A. J. & Kruskal, J. B. (2010), Integral boundary points of convex polyhedra, *in* '50 Years of Integer Programming 1958-2008', Springer, pp. 49–76. 19, 33

Huntington, E. V. (1924), 'A new set of postulates for betweenness, with proof of complete independence', *Transactions of the American Mathematical Society* **26**(2), 257–282. 22

Huntington, E. V. & Kline, J. R. (1917), 'Sets of independent postulates for betweenness', *Transactions of the American Mathematical Society* **18**(3), 301–325. 22

Kalai, G., Rubinstein, A. & Spiegler, R. (2002), 'Rationalizing choice functions by multiple rationales', *Econometrica* **70**(6), 2481–2488. 26

Kreps, D. M. (1988), 'Notes on the theory of choice', *Taylor & Francis eBooks DRM Free Collection* . 21

Manzini, P. & Mariotti, M. (2006), 'Two-stage boundedly rational choice procedures: Theory and experimental evidence'. 4

Masatlioglu, Y., Nakajima, D. & Ozdenoren, E. (2020), 'Willpower and compromise effect', *Theoretical Economics* **15**(1), 279–317. 16, 17

Mehra, R. & Prescott, E. C. (1985), 'The equity premium: A puzzle', *Journal of Monetary Economics* **15**(2), 145–161. 8

Petri, H. (2023), 'Characterizing ordered models of behavioral heterogeneity', *mimeo* . 7, 12

Roberts, K. W. (1977), 'Voting over income tax schedules', *Journal of Public Economics* **8**(3), 329–340. 4

Rubinstein, A. (1988), 'Similarity and decision-making under risk (is there a utility theory resolution to the Allais paradox?)', *Journal of Economic Theory* **46**(1), 145–153. 9

Spiegler, R. (2016), 'Choice complexity and market competition', *Annual Review of Economics* **8**, 1–25. 4

Von Neumann, J. (1953), 'A certain zero-sum two-person game equivalent to the optimal assignment problem', *Contributions to the Theory of Games* **2**(0), 5–12. 10

7 Appendix

7.1 Proof of Theorem 2

Since μ^θ is self-progressive, it follows from Theorem 1 that $\langle \mu^\theta, \rhd \rangle$ is a lattice such that there is no $\mu \subsetneq \mu^\theta$ that contains every rational choice function and $\langle \mu, \rhd \rangle$ is a lattice. Let μ^* be the choice model comprising choice functions that satisfy $\theta 1$ and $\theta 2$.

We first show that $\mu^\theta \subset \mu^*$. To see this, first note that each rational choice function $c \in \mu^*$, since for each $S \in \Omega$ and $x \in S$, rationality of c implies that $c(S) \neq c(S \setminus \{x\})$ only if $x = c(S)$. Next, we show that $\langle \mu^*, \rhd \rangle$ is a lattice. Let $c^1, c^2 \in \mu^*$ and $c = c^1 \vee c^2$. Then, to see that c satisfies $\theta 1$ and $\theta 2$, assume w.l.o.g. that $c(S) = c^1(S)$. Now, if $c^1(S) > x$ then, since c^1 satisfies $\theta 1$, we have $c^1(S \setminus \{x\}) \geq c^1(S)$. It follows that $c(S \setminus \{x\}) \geq c(S)$. If $x > c^1(S)$, then $x > c^2(S)$. Since c^1 and c^2 satisfy $\theta 2$, we have $c(S) \geq c(S \setminus \{x\})$. Thus, we conclude that $c^1 \vee c^2 \in \mu^*$. Symmetric arguments show that $c^1 \wedge c^2 \in \mu^*$ as well.

Next, we show that $\mu^* \subset \mu^\theta$. To see this, let $c \in \mu^*$. Since $\langle \mu^\theta, \rhd \rangle$ is a lattice, by Lemma 1, it suffices to show that for each $S, S' \in \Omega$, there exists $c^* \in \mu^\theta$ such that $c^*(S) = c(S)$ and $c^*(S') = c(S')$. Let $S, S' \in \Omega$ such that $c(S) = a$ and $c(S') = a'$. If $a = a'$, then $c(S)$ and $c(S')$ are obtained by maximizing a preference relation that top-ranks a. If $a \neq a'$, then assume w.l.o.g. that $a > a'$. Now, there are two cases.

Case 1: Suppose that $\{a, a'\} \not\subset S \cap S'$. Then, let c_1 be a choice function maximizing a preference relation that top-ranks first a then a', and c_2 be a choice function maximizing a preference relation that top-ranks first a' then a. Next, if $a \notin S'$ then let $c^* = c_1 \vee c_2$, if $a' \notin S$ then let $c^* = c_1 \wedge c_2$. For both cases, $c^*(S) = a$ and $c^*(S') = a'$, and $c^* \in \mu^\theta$ since $\langle \mu^\theta, \rhd \rangle$ is a lattice containing every rational choice function.

Case 2: Suppose that $\{a, a'\} \subset S \cap S'$. First, we show that either (i) there exists $x \in S \setminus S'$ with $x > a$ or (ii) there exists $y \in S' \setminus S$ with $a' > y$. If not, then consider $S \cap S'$. Suppose that we remove each $x \in S \setminus S'$ from S one-by-one. Since $c \in \mu^\theta$, by applying $\theta 1$ at each step, we get $c(S \cap S') \geq c(S)$. Similarly, suppose that we remove each $y \in S' \setminus S$ from S' one-by-one. Then, by applying $\theta 2$ at each step, we get $c(S') \geq c(S \cap S')$. Therefore, we must have $a' \geq a$, a contradiction. Thus, we conclude that (i) or (ii) holds.

Suppose that (i) holds. Then, let $c^* = c_1 \wedge c_2$, where c_1 maximizes a preference relation that top-ranks first x then a', and c_2 maximizes a preference relation that top-ranks a. Suppose that (ii) holds. Then, let $c^* = c_1 \vee c_2$, where c_1 maximizes a preference relation that top-ranks first y then a, and c_2 maximizes a preference relation that top-ranks a'. For both cases, $c^*(S) = a$ and $c^*(S') = a'$, and $c^* \in \mu^\theta$ since $\langle \mu^\theta, \rhd \rangle$ is a lattice such that μ^θ contains every rational choice function.

7.2 Proof of Lemma 2

Recall that $\mathbb{X} = \{(x, S) : S \in \Omega \;\&\; x \in S\}$. For each $S \in \Omega$, let $\bar{x}_S$ ($\underline{x}_S$) be the $>$-best(-worst) alternative in S. We denote the element that is immediately $>$-worse than an alternative $x \in S \setminus \{\underline{x}_S\}$ by $x - 1$ (we suppress the reference to S, since it will be clear from the context). Then, let $q : \mathbb{X} \to [0, 1]$ that satisfies the following inequalities

$$q(y, S) - q(y, S \setminus \{x\}) \leq 0 \quad \forall (y, S) \in \mathbb{X} \;\; \& \; \forall x \in S \text{ such that } y > x \tag{1}$$

$$q(y, S \setminus \{x\}) - q(y, S) \leq 0 \quad \forall (y, S) \in \mathbb{X} \;\; \& \;\; \forall x \in S \text{ such that } x > y \tag{2}$$

$$q(x, S) - q(x - 1, S) \leq 0 \quad \forall S \in \Omega \;\&\; \forall x \in S \setminus \{\underline{x}_S\} \tag{3}$$

$$q(\underline{x}_S, S) \leq 1 \quad \forall S \in \Omega \tag{4}$$

Note that this system of linear inequalities can be written as: $\Lambda q \leq \mathbb{I}$ where $\Lambda = [\lambda_{rc}]$ is a matrix with entries $\lambda_{rc} \in \{-1, 0, 1\}$, and $\mathbb{I}$ is a column vector whose entries are 0 or 1. Each column of Λ is associated to some $(x, S) \in \mathbb{X}$. Let Q denote the associated polytope $\{q \in [0,1]^{|\mathbb{X}|} : \Lambda q \leq \mathbb{I}\}$. The matrix Λ is called **totally unimodular** if the determinant of each square submatrix of Λ is 0, 1 or -1. It follows from Theorem 2 by Hoffman & Kruskal (2010) that if Λ is totally unimodular then the vertices of Q are $\{0, 1\}$- valued. Heller & Tompkins (1956) provide the following sufficient condition for a matrix being totally unimodular.

Theorem 4 (Heller & Tompkins (1956))**.** *A matrix Λ' is totally unimodular if its rows can be partitioned into two disjoint sets R_1 and R_2 such that:*

1. *Each entry in Λ' is 0, 1, or -1;*
2. *Each column of Λ' contains at most two non-zero entries;*
3. *If two non-zero entries in a column of Λ' have the same sign, then the row of one is in R_1, and the other is in R_2;*
4. *If two non-zero entries in a column of Λ' have opposite signs, then the rows of both are in R_1, or both in R_2.*

To use this result, let Λ' be the transpose of Λ. As it is well-known, and immediately follows from the definition of total unimodularity, a matrix is totally unimodular if and only if its transpose totally unimodular. Next, we show that Λ' satisfies the premises of Theorem 4. First, note that each column in Λ' contains at most two nonzero entries which can be 1 or -1. Therefore, (1) and (2) hold for each column in Λ'. Second, let R_1 be the whole row set while R_2 is the empty set. Note that if a column of Λ'contains two nonzero entries, then one of them is 1 while the other one is -1. Therefore, (3) and (4) hold for each column in Λ'.

7.3 Proof of Lemma 3

If there no triple among $x, y, z, w \in X$ appear in $\mathcal{B}_\mu$, then let $>_L$ be any ordering of these alternatives. For what follows, assume w.l.o.g that $y\ \mathcal{B}_\mu\ \{x, z\}$. If no other triple appear in $\mathcal{B}_\mu$, then let $>_L$ be any ordering such that $x >_L y >_L z$. If neither x, y, w nor y, z, w appear in $\mathcal{B}_\mu$, then let $>_L$ be any ordering such that $x >_L y >_L z$ and w is ordered depending on how x, z, w appear in $\mathcal{B}_\mu$. It is easy to see that for these cases the selected $>_L$ agrees with $\mathcal{B}_\mu$.

Suppose that x, y, w and y, z, w appear in $\mathcal{B}_\mu$. Then, it follows from $B3$ that w lies either on the x- or z-side of y. If x, z, w fail to appear in $\mathcal{B}_\mu$, then we can choose $>_L$ such that $x >_L y >_L z$ and w is ordered depending on the side of y in which w is located. If x, z, w appear in $\mathcal{B}_\mu$, then $\mathcal{B}_\mu$ satisfies $sB1$. Then, by Theorem 4 of Fishburn (1971), there is an ordering that agrees with $\mathcal{B}_\mu$.

Finally, suppose that only one of the triples x, y, w or y, z, w fail to appear in $\mathcal{B}_\mu$. Assume w.l.o.g. that it is y, z, w. If x, z, w also fail to appear in $\mathcal{B}_\mu$, then we can choose $>_L$ such that $x >_L y >_L z$ and w is ordered depending on how x, y, w appear in $\mathcal{B}_\mu$. If x, z, w appear in $\mathcal{B}_\mu$, then there are three cases that we will consider separately.

Case 1: Suppose that $z\ \mathcal{B}_\mu\ \{x, w\}$. Since we also have $y\ \mathcal{B}_\mu\ \{x, z\}$, we can construct an ordering $>_L$ that agrees with $\mathcal{B}_\mu$ only if $y\ \mathcal{B}_\mu\ \{x, w\}$. To see that $y\ \mathcal{B}_\mu\ \{x, w\}$, by contradiction suppose that $w\ \mathcal{B}_\mu\ \{x, y\}$ or $x\ \mathcal{B}_\mu\ \{w, y\}$. First, since $y\ \mathcal{B}_\mu\ \{x, z\}$ and $z\ \mathcal{B}_\mu\ \{x, w\}$, it directly follows from $B2$ that it is not $w\ \mathcal{B}_\mu\ \{x, y\}$. If $w\ \mathcal{B}_\mu\ \{x, y\}$, then since x, z, w

and y, z, w appear in $\mathcal{B}_\mu$, it follows from $B3$ that $w\ \mathcal{B}_\mu\ \{x, z\}$ or $w\ \mathcal{B}_\mu\ \{y, z\}$ but not both. However, since we supposed $z\ \mathcal{B}_\mu\ \{x, w\}$, by $B1$, it is not $w\ \mathcal{B}_\mu\ \{x, z\}$. Since we supposed y, z, w fail to appear in $\mathcal{B}_\mu$ it is not $w\ \mathcal{B}_\mu\ \{x, z\}$ either.

Case 2: Suppose that $w\ \mathcal{B}_\mu\ \{x, z\}$. Since we also have $y\ \mathcal{B}_\mu\ \{x, z\}$, we can construct an ordering $>_L$ that agrees with $\mathcal{B}_\mu$ unless $x\ \mathcal{B}_\mu\ \{w, y\}$. To see that it is not $x\ \mathcal{B}_\mu\ \{w, y\}$, by contradiction, suppose that $x\ \mathcal{B}_\mu\ \{w, y\}$. Then, since x, y, z and x, w, z appear in $\mathcal{B}_\mu$, it follows from $B3$ that $x\ \mathcal{B}_\mu\ \{y, z\}$ or $x\ \mathcal{B}_\mu\ \{w, z\}$ but not both. But, by $B1$, this is not possible since we already have $y\ \mathcal{B}_\mu\ \{x, z\}$ and $x\ \mathcal{B}_\mu\ \{w, z\}$.

Case 3: Suppose that $x\ \mathcal{B}_\mu\ \{w, z\}$. Since $y\ \mathcal{B}_\mu\ \{x, z\}$, an ordering $>_L$ agrees with $\mathcal{B}_\mu$ only if $x\ \mathcal{B}_\mu\ \{w, y\}$. To see that $x\ \mathcal{B}_\mu\ \{w, y\}$, first notice x, y, z and x, y, w appear in $\mathcal{B}_\mu$. Then, since $x\ \mathcal{B}_\mu\ \{w, z\}$, it follows from $B3$ that $x\ \mathcal{B}_\mu\ \{y, z\}$ or $x\ \mathcal{B}_\mu\ \{w, y\}$ but not both. Since we already have $y\ \mathcal{B}_\mu\ \{x, z\}$, by $B1$, it is not $x\ \mathcal{B}_\mu\ \{y, z\}$, thus we must have $x\ \mathcal{B}_\mu\ \{w, y\}$.