A Random Reference Model

Özgür Kıbrıs[*] Yusufcan Masatlioglu[†] Elchin Suleymanov[‡]

August 22, 2021

Abstract

We provide two nested models of random reference-dependent choice in which the reference point is endogenously determined by random processes. Random choice behavior is due to random reference points even though, from the viewpoint of the decision maker, choices are deterministic. We first propose and characterize a parametric random reference model. The components of this model can be identified from observed choices. We then illustrate that similar revelations hold in a much general model as well. Hence, our work is within the revealed preference paradigm.

Keywords: Random Choice, Random Reference, Reference Point Formation, Reference Dependence, Revealed Preference
JEL Codes: D01, D11

[*]Faculty of Arts and Social Sciences, Sabancı University; e-mail: ozgur.kibris@sabanciuniv.edu.
[†]Department of Economics, University of Maryland; e-mail: yusufcan@umd.edu.
[‡]Department of Economics, Purdue University; e-mail: esuleyma@purdue.edu.

1 Introduction

Reference dependence is widely accepted as a fundamental feature of decision-making in behavioral economics.[1] In response to mounting evidence, a variety of theories of reference-dependent choice were proposed. Their common feature is that a single alternative serves as a reference point for each decision problem. However, many real life situations present a multitude of possible candidates for a reference point (e.g. Baucells et al. [2011], Koop and Johnson [2012], Baillon et al. [2020]). Indeed, the seminal work of Kahneman [1992] emphasizes that for each decision problem there might be multiple potential reference points: *"There are many situations in which people are fully aware of the multiplicity of relevant reference points, and the question of how they experience such outcomes and think about them must be raised. There appears to have been little discussion of this issue in behavioral decision research."*

To study multiple reference points, we consider a single individual making repeated decisions.[2] We assume that each feasible alternative serves as a reference point with a certain probability. In our model, this probability is assumed to be unobservable and, more importantly, will be endogenously derived from observed choice behavior. In line with the literature, each reference point induces a specific reference-dependent preference. Furthermore, given a reference point, the agent chooses the alternative which maximizes the corresponding reference-dependent preference. The reference-dependent preferences are also assumed to be unobservable, and will be endogenously derived from choice data.

In our model, choices are deterministic from the point of view of the individual. Yet, she might pick different alternatives in repetitions of a decision problem due to changes in her reference point. The analyst observes probabilistic choice data due to randomness in determination of the reference point.

The first component of our model is a family of reference-dependent preferences. Each alternative, x, induces a reference-dependent preference (relation) denoted by $\succ_x$. It is well known that reference-dependent preferences must be tied together through reasonable re-

[1] Some of applications of reference-dependent preferences include attitudes towards risk (Rabin [2000], Wakker [2010]), the equity premium puzzle (Benartzi and Thaler [1995]), the annuitization puzzle (Benartzi et al. [2011]), the disposition effect (Odean [1998], Genesove and Mayer [2001]), default bias in pension and insurance choice (Samuelson and Zeckhauser [1988], Thaler and Benartzi [2004], Sydnor [2010]), selection of internet privacy policies (Johnson et al. [2002]), and organ donation (Johnson and Goldstein [2003]), as well as behavior of professional golf players (Pope and Schweitzer [2011]), poker players (Eil and Lien [2014]), cab drivers (Camerer et al. [1997]), and physicians (Rizzo and Zeckhauser [2003]).

[2] Individual choices in repeated decisions were first reported in Tversky [1969], which showed that many decision makers make different choices when faced with the same choice problem. Numerous other experimental studies have replicated this choice pattern (*e.g.* see Agranov and Ortoleva [2017] and the references therein). Repeated decisions are also observable in naturally occurring data, such as scanner data from supermarkets and online data from digital platforms.

strictions so that the resulting model is well-behaved (e.g. see Masatlioglu and Ok [2005], Sagi [2006], Kőszegi and Rabin [2006]). To relate reference-dependent preferences $\{\succ_x\}$, we impose the classical status quo bias condition which says that an alternative is most desirable when it is the reference point. In other words, if x is preferred to y when x is not the reference point, then x must also be preferred to y when x itself is the reference point. This condition eliminates unwanted behavioral patterns such as cyclical choice and status quo *aversion*, see Sagi [2006], Masatlioglu and Ok [2014].

The second key ingredient of our model is the stochastic process through which the random reference points are formed. Our rule satisfies two desirable requirements: (i) in a choice set, distribution of reference points need not be uniform (it is alternative specific), and (ii) the distribution of the reference points depends on the choice set (it is context-dependent). We provide a simple, though non-trivial, rule which satisfies these requirements and is tractable enough to use in applications. Namely, each alternative x is endowed with a weight w_x measuring the salience of the alternative as a reference point. In line with Luce's logit model, the probability of an alternative being the reference point is determined by its own weight relative to the total weight of all available alternatives. Note that these weights are not observable and will be derived from observed random choices. The probability of x being chosen in S can then be expressed as

$$p(x|S) = \sum_{y \in S} \underbrace{\left(\frac{w_y}{\sum\limits_{z \in S} w_z} \right)}_{\substack{\text{probability of } y \text{ being} \\ \text{the reference point}}} \underbrace{\mathbb{1}(x \text{ is } \succ_y\text{-best in } S)}_{x \text{ is the maximizer of } \succ_y} .$$

We call this the Random Reference Model (RAR). RAR is a canonical model in the sense that it does apply to any individual choice problem, such as choice among consumption bundles, lotteries, acts, consumption streams, and distributions of wealth. Consequently, we focus on an abstract domain where the distribution of the reference point is only determined by the choice problem.[3]

RAR offers a new perspective for stochastic choices. Previous literature interpreted stochastic choices of a single individual as the outcome of fluctuating tastes (Thurstone [1927], Luce [1959], Marschak [1959])[4], random attention (Manzini and Mariotti [2014], Brady and Rehbeck [2016], Aguiar [2017], Cattaneo et al. [2020]), learning (Baldassi et al. [2020]), ran-

[3]While we consider an arbitrary domain of alternatives, in application to specific domains, the reference probability could be a function of different attributes of an alternative. For example, if the objects are risky prospects, then the reference probability could depend on the prize dimension only or on both the prize and the probability dimension.

[4]See also Apesteguia et al. [2017], Ahumada and Ulku [2018], Echenique and Saito [2019], Kovach and Tserenjigmid [2019], Filiz-Ozbay and Masatlioglu [2020], Horan [2021].

dom stopping (Dutta [2020]), or deliberate randomization (Machina [1985], Fudenberg et al. [2015], Cerreia-Vioglio et al. [2019]). In our model, on the other hand, the source of randomness is stochastic reference points.

RAR differs from the classical random utility model (RUM) in two important ways. First, while there are multiple preferences in RAR, these are related through the status quo bias condition. In RUM, the set of preferences is arbitrary. Second, in RAR the set of reference-dependent preferences is context-dependent and the number of preferences applied to a choice set is bounded by the number of alternatives in it. In RUM, the set of preferences is independent of context. Finally, RAR violates the well-known regularity condition of RUM. Furthermore, the intersection of RAR and RUM only contains the deterministic rational-choice model and the Luce model. Appendix C contains a more detailed discussion of the relationship between RAR and the random choice literature.

RAR can accommodate well-known context effects such as the attraction effect (Huber et al. [1982]). The attraction effect refers to the finding that the relative choice proportion of two alternatives is affected by the availability of a third option that is asymmetrically dominated by one of the alternatives. In particular, it is frequently observed that the addition of an asymmetrically dominated alternative improves the choice share of the dominating option. While RAR can accommodate such choice patterns, since the attraction effect involves a regularity violation, it cannot be accommodated within random utility models. In addition, although the attraction effect has often been demonstrated in the marketing and economics literature using between-subjects designs, recently, Berkowitsch et al. [2014] and Mohr et al. [2017] show that the attraction effect may also be observed for the same individual when making repeated decisions. While deterministic models of reference dependence are not capable of explaining these experimental observations, our model can easily accommodate them.

RAR includes interesting special cases. First, if all the reference-dependent preferences are identical, RAR reduces to the classical model of deterministic rational choice, where the decision maker maximizes a reference-free preference. Second, if each reference-dependent preference exhibits extreme bias towards its reference point, that is, if the reference point is always the best alternative, then RAR becomes the standard logit model of Luce. In that case, since an alternative is chosen only when it is the reference point, the choice probability of an alternative is equal to the probability of it being the reference point. This choice behavior resembles the idea of personal equilibrium of Kőszegi and Rabin [2006]. That is, the distribution of reference points matches with the distribution of choices. Note that this equivalence is independent of the parametric structure we have on reference point formation; it continues to hold for any stochastic reference formation process. In other words, personal equilibrium is equivalent to extreme status quo bias in our framework.

In Section 3, we provide a set of behavioral postulates that characterize the empirical content of RAR. They can be classified into two groups as (i) ordinal axioms dealing with issues such as choice with zero probability or violations of regularity, and (ii) axioms on cardinal probability values. Among our ordinal axioms, one imposes that if an alternative is chosen with zero probability in a binary comparison, addition of new alternatives should not increase its choice probability. Other ordinal axioms impose conditions under which a regularity violation might be observed. A regularity violation is said to occur when elimination of an alternative reduces the choice probability of another alternative. One of the key axioms states that a regularity violation can occur only if the eliminated alternative was chosen with zero probability. Another key axiom imposes asymmetry on regularity violations. The axiom states that if removing an alternative z causes a regularity violation for another alternative x, then removing z cannot cause a regularity violation for $y \neq x$ as long as x is available. Our cardinal axioms provide conditions under which Luce's Independence of Irrelevant Alternatives (IIA) property should be satisfied, and they regulate changes in probabilities in cases when it is violated.

In Section 4, we illustrate how to identify the reference-dependent preferences $\{\succ_x\}$ and the reference weights from stochastic choice data. The inference about preferences relies on three different observations. The first one is being chosen with probability 1. In this case, the chosen alternative is better than any alternative for any reference point in that decision problem as well as any other reference point. The second one is being chosen with positive probability in some choice problem. Then, we can say that this alternative is the best alternative in the choice problem when it is the reference point. The final one is about regularity violations. If the elimination of x from S causes a regularity violation for y, then y must be better than any other alternative in S when x is the reference point. These three observations completely characterize all the revealed preference implications of our model.

In the Luce model, the relative weight of x and y can be inferred from any choice set. This is not the case in our model. First, choice probabilities might be 0 or 1 in some choice problems. Second, even for intermediate probability values, the relative choice probabilities are context dependent and thus can change. Nevertheless, we show that all relevant relative reference weights can be revealed by looking at choice sets of size two and three. There are three different observations through which the reference weight of x relative to y can be revealed. First, the relative choice probabilities of x and y in $\{x, y\}$ reveal the reference weight of x relative to y unless one of the alternatives is chosen with probability one. Second, if x is not chosen from $\{x, y, z\}$ and if its elimination induces a regularity violation on z, the difference between the relative choice probabilities of z and y in $\{x, y, z\}$ versus $\{y, z\}$ reveals the reference weight of x relative to y. The third way in which the reference weight of x relative to y can be revealed is when x and y are consecutive members of a cycle in which the reference weights for every other consecutive pair are known.

In random utility models, the set of "rationalizing" preferences is not unique in general. While revealed reference-dependent preferences in RAR are also not unique, the relative ranking of any two alternatives under any reference point can be revealed whenever it matters for choice. Alternatively, we may not reveal the relative ranking of two alternatives y and z under the reference point x only if x is preferred to both y and z when it is the reference point.

One might wonder whether the strong (reference-dependent) preference revelations we obtain are due to the particular parametric assumption we make on the reference probabilities. In Section 5 we answer this question by studying a general model of reference formation. The reference formation process of this generalized RAR (henceforth, GAR) satisfies two basic conditions: the reference probabilities (i) are strictly positive, and (ii) satisfy strict regularity (*i.e.* the reference probability of an alternative is decreasing in the menu size). These basic conditions are satisfied by a wide range of "rational" reference formation processes including that of RAR.[5] We show that the revealed (reference-dependent) preferences of GAR are exactly the same as that of RAR. This result has two implications. First, the revealed preference result is not driven from our parametric modelling choice of reference probabilities. Second, any stochastic model of reference dependence satisfying these two assumptions will have exactly the same revealed preference as RAR. In Section 5, we also provide a set of behavioral postulates that characterize the empirical content of GAR. On top of the four ordinal axioms that characterize RAR, we introduce a critical axiom. It is closely related to Motzkin's transposition theorem, a member of the well-known theorems of the alternatives. This axiom guarantees that the reference point formation rule satisfies the aforementioned requirements.

Our paper is foremost related to the growing literature on reference-dependent choice. The earliest strand of this literature treats the reference point as exogenous (e.g. Kahneman and Tversky [1979], Tversky and Kahneman [1991], Munro and Sugden [2003], Sugden [2003], Masatlioglu and Ok [2005], Sagi [2006], Salant and Rubinstein [2008], Masatlioglu and Ok [2014], Dean et al. [2017], Guney and Richter [2018], Kovach and Suleymanov [2021]). Our paper is distinct from these papers in endogenizing the reference formation process. In addition, all of these papers except for Kovach and Suleymanov [2021] address deterministic choices. A second strand of the literature studies endogenous reference point formation. In models of Bodner and Prelec [1994], Kivetz et al. [2004], Orhun [2009], Bordalo et al. [2013], and Tserenjigmid [2019], the reference point depends on the structure of the choice set but it

[5] For example, random utility representations with full support on preference rankings, the random consideration model of Manzini and Mariotti [2014], or the weighted linear stochastic choice model of Chambers et al. [2021] all satisfy both requirements.

is independent of individual characteristics. Kőszegi and Rabin [2006], Rubinstein and Salant [2006], Ok et al. [2015], Freeman [2017], Kıbrıs et al. [2018], and Lim [2020] consider models where the endogenous reference point might differ across individuals. Maltz [2020] consider a hybrid model which combines an exogenous status quo with an endogenous reference point. These papers, however, only address deterministic choice behavior.[6] To the best of our knowledge, RAR is the first stochastic model of endogenous reference dependence.

The paper is organized as follows. In Section 2, we introduce our model. In Section 3, we introduce the behavioral postulates that characterize RAR and present our representation theorem. In Section 4, we show how the primitives of RAR are revealed from observed choices. In Section 5, we generalize our model to GAR and discuss its identification and characterization. We conclude in Section 6.

2 Model

Let X be a non-empty finite set of alternatives, and let $\mathcal{X}$ be the set of all nonempty subsets of X. A *choice problem* is a set of alternatives $S \in \mathcal{X}$ from which the decision maker needs to make a choice. A *choice rule* is a map $p : X \times \mathcal{X} \to [0, 1]$ such that for all $S \in \mathcal{X}$, $p(x|S) > 0$ only if $x \in S$ and $\sum_{x \in S} p(x|S) = 1$. The choice rule p represents data on the choice behavior of the decision maker (hereafter, *DM*). The expression $p(x|S)$ represents the probability of x being chosen from the choice problem S. Note that if $p(x|S) \in \{0, 1\}$ for every x and S, then choices are deterministic. Hence, our formulation encompasses both stochastic and deterministic choice rules.

Our model has two components: (i) a family $\{\succ_x\}_{x \in X}$ of *reference-dependent preferences* where each $\succ_x$ is a strict linear order that represents the DM's preferences under the reference point x,[7] and (ii) a family $\{w_x\}_{x \in X}$ of *reference weights*, where each $w_x > 0$ influences the likelihood of alternative x being realized as the reference point.

We assume that the reference-dependent preferences $\{\succ_x\}_{x \in X}$ satisfy the following assumption: if x is preferred to y when z is the reference point, then x must also be preferred to y when x itself is the reference point. This assumption relates two reference-dependent preferences and, in line with the concept of status quo bias, requires that being the reference point cannot hurt any alternative.

[6] To clarify, what we refer to is deterministic choice behavior rather than the possibility of having alternatives that are lotteries. For example, Kőszegi and Rabin [2007] takes alternatives to be lotteries, yet studies deterministic choices from them.

[7] A binary relation R on X is a strict linear order if it is (i) weakly connected: for every $x, y \in X$, $x \neq y$ implies either xRy or yRx, (ii) irreflexive: for every $x \in X$, it is not the case that xRx, and (iii) transitive: for every $x, y, z \in X$, xRy and yRz imply xRz.

Status Quo Bias (SQB). If $x \succ_z y$ then $x \succ_x y$.

In our model, the reference point is stochastically determined à la Luce [1959]. That is, we assume that the probability of x being the reference point in S is equal to its own reference weight relative to the total weight of all alternatives in S. Once a reference point x is determined from the choice problem S, the DM maximizes the associated reference-dependent preference $\succ_x$ on S to make a choice. The following definition formally states the choice process in our model.

Definition 1. *A choice rule p is consistent with the* ***random reference model (RAR)*** *if there exist a family $\{\succ_x\}_{x\in X}$ of reference-dependent preferences satisfying SQB and a family $\{w_x\}_{x\in X}$ of reference weights such that for each $S \in \mathcal{X}$ and $x \in S$,*

$$p(x|S) = \sum_{y\in S} \left(\frac{w_y}{\sum\limits_{z\in S} w_z} \right) \mathbb{1}(x \textit{ is } \succ_y\textit{-best in } S).$$

We also say $\{\succ_x, w_x\}_{x\in X}$ represents p, or p admits a RAR representation.

RAR includes two well-known special cases. At one extreme, if the DM has a reference-independent preference and all reference-dependent preferences coincide with this preference, RAR reduces to the classical model of deterministic rational choice. At the other extreme, if the DM exhibits extreme status quo bias where each reference-dependent preference ranks its reference point at the top, RAR coincides with the Luce model with weights $\{w_x\}_{x\in X}$. Hence, the rational choice model and the Luce model are extreme cases of RAR with no status quo bias and extreme status quo bias, respectively.

One important characteristic of the Luce model and all other random utility models is the *regularity axiom* (Suppes and Luce [1965]). It requires that the choice probability of an alternative does not decrease as the choice set gets smaller, that is, $p(x|S) \leq p(x|T)$ for $x \in T \subset S$. As discussed in the introduction, one of the common findings in the literature is the attraction (asymmetric dominance) effect, which involves a regularity violation. RAR can accommodate such violations. For example, consider three alternatives x, y, z where z is dominated by x but is not dominated by y. The asymmetric dominance effect states that adding z to the choice set $\{x, y\}$ should increase the choice probability of x (see Table 1), a choice pattern in contrast to random utility models. This choice pattern can be accommodated in RAR by assuming that the preference under the reference point z is $x \succ_z z \succ_z y$ and x and y are ranked at the top when they are reference points.

In this example, RAR also predicts the magnitude of the regularity violation. It requires that the relative choice probability of x and y in $\{x, y, z\}$ is equal to the summation of relative choice probabilities of x and y in $\{x, y\}$ and z and y in $\{y, z\}$. Note that the choice rule in

Table 1 satisfies this property ($\frac{1/2}{1/2} = \frac{1/3}{2/3} + \frac{1/3}{2/3}$). The next section illustrates additional restrictions RAR imposes on observed choices.

Table 1: RAR accommodates choice data exhibiting regularity violations.

$p(\cdot\|S)$	$\{x, y, z\}$	$\{x, y\}$	$\{x, z\}$	$\{y, z\}$
x	1/2	1/3	1	–
y	1/2	2/3	–	2/3
z	0	–	0	1/3

3 Behavioral Postulates

In this section, we discuss the behavioral postulates that characterize the empirical content of RAR. Our postulates can be classified into two groups. Axioms 1-4 consider ordinal properties of observed choice rule, dealing with issues such as choice with zero probability or regularity violations. Axioms 5-8, on the other hand, consider cardinal probability values. Axiom 5 imposes a condition akin to Independence of Irrelevant Alternatives (IIA) property on revealed reference probabilities. Axiom 6 discusses conditions under which IIA is satisfied for observed choice probabilities, and Axioms 7-8 regulate changes in probabilities when it is violated.

The following terminology will be helpful. We say x *is chosen from* S if $p(x|S) > 0$. If $p(x|S) = 0$, we say x is not chosen from S. Our first axiom states that if x is not chosen against y in a binary comparison, it cannot be chosen from any choice problem contains y. Hence, it is a significant relaxation of the well-known regularity condition which requires $p(x|S) \leq p(x|T)$ whenever $x \in T \subset S$.

Axiom 1. *If $p(x|\{x, y\}) = 0$, then $p(x|S) = 0$ for every S that contains y.*

In our model, if x is not chosen against y, that means $y \succ_x x$. Since reference-dependent preferences that we consider satisfy the SQB property, it must be the case that y is ranked above x regardless of the reference point. Hence, x can never be chosen in the presence of y.

Our model allows regularity violations. The following three axioms regulate what type of regularity violations can be observed. The next axiom states that if an alternative x is not chosen in S, there must be an alternative y which beats x in a binary comparison, and if this y does not beat all other alternatives in S, then eliminating x must induce a regularity violation for y.

Axiom 2. *If $p(x|S) = 0$, then there is $y \in S$ such that $p(y|\{x, y\}) = 1$ and either $p(y|S) = 1$ or $p(y|S) > p(y|S \setminus x)$.*

This axiom necessarily holds in our model. If x is not chosen from S, that means another alternative y must be $\succ_x$-best in S, and by SQB, must also be $\succ_y$-best. Since y is better than x under both reference-dependent preferences, it must be that x is not chosen from $\{x, y\}$. Furthermore, unless y similarly beats every other alternative in S (in which case y must be the only chosen alternative in S), it must be that elimination of x from S decreases the choice probability of y, inducing a regularity violation. To see why, let $A(y|S)$ be the set of reference points in S according to which y is the best alternative in S. By $p(y|\{x, y\}) = 1$, we know that x belongs to that set. With the elimination of x from S, $A(y|S \setminus x)$ gets smaller. If $A(y|S) \neq S$, then this induces $p(y|S) > p(y|S \setminus x)$. If there are no other chosen alternatives (that is, if $A(y|S) = S$), on the other hand, $p(y|S) = p(y|S \setminus x) = 1$ remains the same.

Axiom 2 states that elimination of an unchosen alternative should induce a regularity violation, unless only one alternative is chosen from S. The following axiom completes this picture by stating that elimination of a chosen alternative cannot induce a regularity violation. Hence, in our model regularity violations happen only due to zero probability choices.

Axiom 3. *If $p(x|S) > 0$, then $p(y|S) \leq p(y|S \setminus x)$ for any $y \in S \setminus x$.*

To see why this axiom holds in our model, note that if x is chosen in S, it must be $\succ_x$-best in S. Hence, if y is $\succ_z$-best for some alternative $z \in S$, then z is distinct from x and y is also $\succ_z$-best in $S \setminus x$, that is, $A(y|S) \subseteq A(y|S \setminus x)$. Since $p(y|S) = \frac{\sum_{A(y|S)} w_z}{\sum_S w_z}$, this implies $p(y|S) \leq p(y|S \setminus x)$ for every $y \in S \setminus x$.

The next axiom imposes a form of asymmetry on regularity violations. It considers a situation where elimination of z induces a regularity violation on x when y is available. Our axiom then states that, in the presence of x, elimination of z cannot induce a regularity violation for y.

Axiom 4. *If $p(x|S) > p(x|S \setminus z)$ and $x, y, z \in T \cap S$, then $p(y|T) \leq p(y|T \setminus z)$.*

To see why Axiom 4 is satisfied our model, note that z induces a regularity violation on x in a set that contains y only if $x \succ_z y$. Due to asymmetry of $\succ_z$, there cannot be another case where z induces a regularity violation on y when x is available.

Axiom 4 is related to the "single reversal" axiom on deterministic choice (Kıbrıs et al. [2018]). Their axiom states that if elimination of x induces a choice reversal in a choice set containing y (i.e., $c(S) \neq c(S \setminus x)$ when $x \neq c(S)$ and $y \in S$), then there cannot be another choice set in which, now, elimination of y induces a choice reversal when x is available. The stochastic analog of the single reversal axiom would require that in any choice set there is at most one alternative elimination of which can cause a regularity violation. While this does not hold in our model, Axiom 4 restricts the number of possible regularity violations in another way by stating that eliminating an alternative can cause a regularity violation for a unique alternative.

The next two axioms are related to Luce's well-known IIA axiom which states that the relative choice probability of two alternatives is independent of the choice problem they are considered in, that is, for $x, y \in S \cap T$,

$$\frac{p(x|S)}{p(y|S)} = \frac{p(x|T)}{p(y|T)}.$$

RAR may not satisfy IIA in general. First, RAR does not require that an alternative is chosen from every choice problem once it is chosen from one choice problem. Second, even when the two alternatives are both chosen from the two choice problems, RAR can exhibit IIA violations (recall the example in Table 1). Hence, accommodating the type of choice behavior we are interested in requires formulation of more qualified versions of IIA. This is what we will do next.

Our first "IIA axiom" ensures that the weights of alternatives revealed from binary and trinary choice sets are consistent. First, note that for any binary choice problem $\{x, y\}$, if $p(x|\{x, y\}) \in (0, 1)$, then $p(x|\{x, y\})$ must reflect the reference probability of x in $\{x, y\}$. Hence, in any binary choice problem where both alternatives are chosen, reference probabilities are fully revealed. We use this to construct a function $q(\cdot|\cdot)$ where $q(x|S)$ reflects the probability that x is the reference point in S. For any $\{x, y\}$ with $p(x|\{x, y\}) \in (0, 1)$ and $a \in \{x, y\}$, let

$$q(a|\{x, y\}) = p(a|\{x, y\}).$$

Note that $q(\cdot|\cdot)$ cannot be defined for all binary choice sets, since we might have $p(x|\{x, y\}) \in \{0, 1\}$ for some $\{x, y\}$. However, when q is defined, it must be strictly between 0 and 1.

Next, consider a trinary choice set. First, if all alternatives are chosen, then it must be the case that observed choice probabilities correspond to reference probabilities. Hence, for any $\{x, y, z\}$ where $p(a|\{x, y, z\}) > 0$ for all $a \in \{x, y, z\}$, we have

$$q(a|\{x, y, z\}) = p(a|\{x, y, z\}).$$

Now suppose that $p(x|\{x, y, z\}) > 0$, $p(y|\{x, y, z\}) > 0$, and $p(z|\{x, y, z\}) = 0$. Interestingly, if p is consistent with RAR, we can also fully reveal $q(\cdot|\cdot)$ in this case. To see this, note that if p is consistent with RAR, we must have either $p(x|\{x, y, z\}) > p(x|\{x, y\})$ or $p(y|\{x, y, z\}) > p(y|\{x, y\})$ (Axiom 2). Assume $p(x|\{x, y, z\}) > p(x|\{x, y\})$. This reveals that

$z \in A(x|\{x, y, z\})$, and hence the observed choice probability of y from $\{x, y, z\}$ must be the same as its reference probability:

$$q(y|\{x, y, z\}) = p(y|\{x, y, z\}).$$

In addition, if p is consistent with RAR, we must also have $p(x|\{x, y\}) > 0$ and $p(y|\{x, y\}) > 0$ (Axiom 1). Hence, $q(x|\{x, y\})$ and $q(y|\{x, y\})$ are defined and equal to $p(x|\{x, y\})$ and $p(y|\{x, y\})$, respectively. Since reference probabilities satisfy IIA in our model, we should then have

$$\frac{q(x|\{x, y, z\})}{q(y|\{x, y, z\})} = \frac{q(x|\{x, y\})}{q(y|\{x, y\})}.$$

Therefore, we can define $q(x|\{x, y, z\})$ as

$$q(x|\{x, y, z\}) = \frac{p(x|\{x, y\})p(y|\{x, y, z\})}{p(y|\{x, y\})}.$$

Lastly, since reference probabilities add up to 1, we have $q(z|\{x, y, z\})$ as

$$q(z|\{x, y, z\}) = p(x|\{x, y, z\}) - \frac{p(x|\{x, y\})p(y|\{x, y, z\})}{p(y|\{x, y\})}.$$

Notice that for singleton choice sets $p(x|\{x\}) = q(x|\{x\}) = 1$ must hold. However, as discussed above, if S is not a singleton set, we cannot reveal $q(\cdot|S)$ when there is an alternative $x \in S$ that satisfies $p(x|S) = 1$. We will let $\mathcal{T}$ denote all singleton, binary, and trinary choice sets for which q is defined as above. The next axiom ensures that q is generated by the Luce rule on $\mathcal{T}$.

Axiom 5. *For any $S_1, S_2, \dots, S_N \in \mathcal{T}$ and any $x_1, \dots, x_N \in X$ such that $\{x_i, x_{i+1}\} \subseteq S_i$ for $i < N$ and $\{x_1, x_N\} \subseteq S_N$,*

$$\frac{q(x_1|S_N)}{q(x_N|S_N)} = \prod_{i=1}^{N-1} \frac{q(x_i|S_i)}{q(x_{i+1}|S_i)}.$$

Since reference probabilities satisfy the IIA property in our model, we should expect that observed choices should also satisfy IIA under certain conditions. For example, if all alternatives are chosen with positive probability from every set, then our model reduces to the Luce rule and, hence, satisfies IIA. Our next axiom generalizes this observation. Consider a chosen alternative $z \in S$ such that elimination of no alternative in S induces a regularity violation for z. Axiom 6 then states that choices must satisfy IIA when such a "well-behaved" alternative z is eliminated from S.

Axiom 6. *If $p(x|S)p(y|S)p(z|S) > 0$ and $p(z|S) \leq p(z|S \setminus t)$ for all $t \in S \setminus z$, then*

$$\frac{p(x|S)}{p(y|S)} = \frac{p(x|S \setminus z)}{p(y|S \setminus z)}.$$

To see that our model satisfies this axiom, note that $p(x|S)p(y|S)p(z|S) > 0$ implies that for these alternatives, the associated reference-dependent preference ranks the reference point as the best in S. Furthermore, elimination of no alternative in S induces a regularity violation for z. That means $A(z|S) = \{z\}$, that is, no other reference-dependent preference in S is maximized at z. Hence, elimination of z does not affect the sets of reference-dependent preferences maximized at x or y: $A(x|S) = A(x|S \setminus z)$ and $A(y|S) = A(y|S \setminus z)$. Thus, the relative choice probability of x and y remains unchanged.

It is useful to note that if all alternatives are chosen with positive probability in S, then Axiom 3 guarantees that observed choices satisfy regularity when any alternative is eliminated from S, and hence Axioms 3 and 6 jointly guarantee that IIA holds in this case.

While RAR allows IIA violations, these violations have a certain structure. Axiom 5 provides one such structure in binary and trinary sets by imposing a condition on revealed reference probabilities. The next two axioms describe the structure of these violations in all other sets. Consider a choice problem S where x and y are chosen, and z and t are not chosen. Furthermore, suppose elimination of z from S induces a regularity violation on x. Since in our model z can induce a regularity violation for only one alternative, this guarantees that elimination of z from S will change the relative choice probabilities of x and y. Similarly, eliminating z from $S \setminus t$ will also change the relative choice probabilities of x and y. The following two axioms describe how this change on S and $S \setminus t$ are related. It turns out that the relationship depends on whether elimination of t from S induces a regularity violation on x or not.

For the first axiom, assume that elimination of t from S does not induce a regularity violation for x. In this case, we look at the ratio

$$\frac{p(x|S)/p(y|S)}{p(x|S \setminus z)/p(y|S \setminus z)}$$

which measures how much more likely x is to be chosen relative to y from S compared to $S \setminus z$. Our axiom states that the above ratio must be the same on S and $S \setminus t$. Hence, z improves the likelihood that x is chosen in a consistent way.

Axiom 7. *If $p(x|S \setminus t) \geq p(x|S) > p(x|S \setminus z)$, $p(y|S) > 0$, and $p(t|S) = 0$, then*

$$\frac{p(x|S)/p(y|S)}{p(x|S \setminus z)/p(y|S \setminus z)} = \frac{p(x|S \setminus t)/p(y|S \setminus t)}{p(x|S \setminus \{t, z\})/p(y|S \setminus \{t, z\})}.$$

To see that our model satisfies this axiom, note that under the given assumptions, $z \in A(x|S)$. Hence, $A(y|S) = A(y|S \setminus z)$. This implies

$$\frac{p(x|S)/p(y|S)}{p(x|S \setminus z)/p(y|S \setminus z)} = \left(\frac{\sum\limits_{A(x|S)} w_a}{\sum\limits_{A(y|S)} w_a}\right)\left(\frac{\sum\limits_{A(y|S\setminus z)} w_a}{\sum\limits_{A(x|S\setminus z)} w_a}\right) = \frac{\sum\limits_{A(x|S)} w_a}{\sum\limits_{A(x|S\setminus z)} w_a}.$$

Since $t \notin A(x|S)$, we have $A(x|S) = A(x|S \setminus t)$ and $A(x|S \setminus z) = A(x|S \setminus \{z,t\})$. Hence, the above ratio remains the same when we replace S with $S \setminus t$.

For the second axiom, now assume elimination of t from S does induce a regularity violation for x. In this case, we consider the difference

$$\frac{p(x|S)}{p(y|S)} - \frac{p(x|S \setminus z)}{p(y|S \setminus z)},$$

which is an alternative measure of how much more likely x is to be chosen relative to y from S compared to $S \setminus z$. Our axiom states that the above difference must be the same on S and $S \setminus t$.

Axiom 8. *If* $p(x|S) > \max\{p(x|S \setminus z), p(x|S \setminus t)\}$ *and* $p(y|S) > 0$*, then*

$$\frac{p(x|S)}{p(y|S)} - \frac{p(x|S \setminus z)}{p(y|S \setminus z)} = \frac{p(x|S \setminus t)}{p(y|S \setminus t)} - \frac{p(x|S \setminus \{z,t\})}{p(y|S \setminus \{z,t\})}.$$

To see that our model satisfies this axiom, note that under the given assumptions, $z, t \in A(x|S)$. Hence, $A(y|S) = A(y|S \setminus z)$. This implies

$$\frac{p(x|S)}{p(y|S)} - \frac{p(x|S \setminus z)}{p(y|S \setminus z)} = \frac{\sum\limits_{A(x|S)} w_a}{\sum\limits_{A(y|S)} w_a} - \frac{\sum\limits_{A(x|S\setminus z)} w_a}{\sum\limits_{A(y|S\setminus z)} w_a} = \frac{w_z}{\sum\limits_{A(y|S)} w_a}.$$

Since $A(y|S) = A(y|S \setminus t)$, this expression remains the same when we replace S with $S \setminus t$.

To get an intuition for the last two axioms, consider an attraction effect example with four alternatives where z is dominated by x and t is either dominated by x or y. Axioms 7 and 8 impose some form consistency on how addition of z to a choice set will improve the relative likelihood that x is chosen. Clearly, the impact of adding z will depend on whether t is dominated by x or not. First, suppose t is dominated by y. Then, Axiom 7 states adding z to either $\{x, y\}$ or $\{x, y, t\}$ must increase the relative choice probability of x and y by the same percentage. Note that since t is dominated by y, x must be chosen with a smaller probability from $\{x, y, t\}$ compared to $\{x, y\}$, and hence an increase by the same percentage implies a smaller increase in magnitude when z is added to $\{x, y, t\}$.

Now, suppose t is dominated by x. Intuitively, one might expect that in this case adding z to $\{x, y, t\}$ should make a smaller impact than adding z to $\{x, y\}$, since there is already an alternative dominated by x in the former choice set. In fact, it can be seen that in this case the relative choice probability of x and y will increase by a smaller percentage when z is added to $\{x, y, t\}$ compared to the case when z is added to $\{x, y\}$. Axiom 8 then imposes an alternative consistency requirement that the magnitude of the increase in the relative choice probabilities must be the same.

We now state the characterization result.

Theorem 1. *A random choice rule p satisfies Axioms 1-8 if and only if it has a RAR representation.*

4 Revelations

Our model has two components. Namely, reference-dependent preferences and reference weights. In this section, we discuss how to infer information about them from choice data that is consistent with RAR.

Before the formal discussion, let us present two extreme examples. First, consider choice data satisfying Luce's IIA. It can have multiple RAR representations. In every one of them, the reference weight of an alternative x relative to an alternative y (that is, w_x/w_y) can be uniquely determined as the choice probability of x relative to y in the grand set (or, by Luce's IIA, in any set that contains them). In addition, each reference point must be ranked as the top alternative in all these RAR representations. However, we cannot infer the relative ranking of two non-reference alternatives. This is an unavoidable non-uniqueness in our model.

At the other extreme, consider deterministic choice data generated by maximization of a single reference-independent preference relation. It also has multiple RAR representations. Unlike the other extreme case though, all reference-dependent preferences are uniquely identified. Indeed, they must be the same as the corresponding (reference-independent) preference relation generating the data. But unlike the other extreme case, reference weights cannot be identified: any positive vector of weights would work.

As will be demonstrated further in this section, existence of multiple representations is not restricted to the extreme cases discussed above. If two representations disagree on the relative ranking of two alternatives under some reference point, then we cannot make any conclusions regarding the ranking of these two alternatives. Similarly, if two representations disagree on the relative reference weights of alternatives, then we cannot make any conclusions. Before making any conclusions regarding revealed preference and reference weights, we will require that all representations agree on these revelations. Formally, assume that p admits k different RAR representations $(\{\succ_x^i, w_x^i\}_{x\in X})_{i\in\{1,\dots,k\}}$. Then we say

1. x is revealed to be preferred to y under reference point z if $x \succ_z^i y$ for each $i \in \{1, ..., k\}$, and

2. the reference weight of x relative to y is revealed to be α_{xy} if $\alpha_{xy} = w_x^i/w_y^i$ for each $i \in \{1, ..., k\}$.

Note that, while the first item is standard in the literature, the second item is novel. In particular, it refers to the identification of the relative reference weights of two alternatives rather than the reference weight of a single alternative. This is because absolute reference weights are essentially non-unique. For example, even in the case of the Luce model, both w and aw' produce the same choice data whenever $a > 0$. Our definition bypasses this rather trivial non-uniqueness by considering relative rather than absolute weights. But there still remains another nontrivial source of non-uniqueness in our model. If an alternative dominates every other alternative (for example, see Table 3), its weight relative to others is not revealed by choice data, as will be discussed below.

According to our definition, to make any conclusions about revealed preference and reference weights, one needs to construct all possible RAR representations and verify the cases in which these representations agree. It is clearly not practical to use this method to obtain revelations. Thus, we next provide a practical method to obtain revealed reference-dependent preferences and reference weights. To do this, we focus on three different types of observations revealing information about the reference-dependent preference under x. For $y \neq z$ in S, we learn that $y \succ_x z$ if

1. $p(y|S) = 1$,

2. $p(y|S) > 0$ and $x = y$, or

3. $p(y|S) > p(y|S \setminus x)$.

For the first type of observation, note that an alternative chosen with probability 1 must be the best alternative in S for any reference point. That is, $y \succ_z z$ for any $z \in S \setminus y$. SQB then implies $y \succ_x z$ for any $x \in X$. The second type of observation is about an alternative chosen with positive probability. In that case, this alternative must be the best alternative for at least one reference point: there exists $z \in S$ such that x is $\succ_z$-best in S. Then, by SQB, x must also be $\succ_x$-best in S. Hence, $x \succ_x z$. The third type of observation is a regularity violation caused by elimination of x, $p(y|S) > p(y|S \setminus x)$. In any RAR representation of p, this means y is the $\succ_x$-best in S, implying $y \succ_x z$. To see this, note that if y is not the $\succ_x$-best in S, the set of reference-dependent preferences it maximizes, i.e., the set $A(y|S) = \{z \in S \mid y = \arg\max(\succ_z, S)\}$), cannot get smaller after elimination of x: $A(y|S) \subseteq A(y|S \setminus x)$. Since $p(y|S) = \frac{\sum_{A(y|S)} w_z}{\sum_S w_z}$, we would then have $p(y|S) \leq p(y|S \setminus x)$, satisfying regularity.

Given choice data p that is consistent with the RAR model, let

$$yP_xz \text{ if one of the three patterns above is observed.}$$

We show that P_x must be transitive. However, it might not be complete (see the first extreme case above). Since any RAR representation of p must be consistent with the above revelations, P_x must be part of the revealed reference-dependent preferences. Furthermore, there is no other revelation: any reference-dependent preference $\{\succ_x\}_{x\in X}$ that respect $\{P_x\}_{x\in X}$ represents p. In other words, if two alternatives are not ranked according to P_x, then we can always find two representations where the relative ranking of these alternatives is opposite of each other. The following remark, the proof of which is identical to the proof of Theorem 1, establishes this point.

Remark 1. ***(Revealed Preference)*** *Suppose p admits a RAR representation. Then y is revealed to be preferred to z under reference point x if and only if yP_xz.*

We next discuss how relative reference weights are revealed from choice data. To this end, we will only make use of binary and trinary sets, even though similar revelations hold for larger sets as well. Three types of observations reveal information about the reference weight of x relative to y

1. if $p(x|\{x, y\}) \in (0, 1)$,
$$\alpha_{xy} = \frac{p(x|\{x, y\})}{p(y|\{x, y\})},$$
2. if $p(x|\{x, y, z\}) = 0$ and $p(z|\{x, y, z\}) > p(z|\{y, z\})$,
$$\alpha_{xy} = \frac{p(z|\{x, y, z\})}{p(y|\{x, y, z\})} - \frac{p(z|\{y, z\})}{p(y|\{y, z\})}.[8]$$
3. for any $\{x_1, ..., x_n\}$, if $(n-1)$ of the n alpha values $\{\alpha_{x_ix_{i+1}}\}_{i=1}^n$ are already known (with

[8]Note that the above conditions imply $p(y|\{x, y, z\}) > 0$ and $p(y|\{y, z\})p(z|\{y, z\}) > 0$. Hence, the expression is well-defined.

the abuse of notation $x_{n+1} = x_1$), the last one can be identified through the equality

$$\prod_{i=1}^{n} \alpha_{x_i x_{i+1}} = 1.$$

We now show that α_{xy} is well-defined and equal to $\frac{w_x}{w_y}$ whenever p has a RAR representation with $\{w_x\}_{x \in X}$. For the first type of observation, note that if $x \neq y$ and both are chosen from $\{x, y\}$ with positive probability, then the corresponding reference point is the best under both $\succ_x$ and $\succ_y$. Hence, the relative choice probability of x and y reflects their relative reference weight and, thus, $p(x|\{x, y\})/p(y|\{x, y\}) = w_x/w_y$, uniquely identifying α_{xy}. The second type of observation concerns an alternative z that is both $\succ_x$ and $\succ_z$ best in $\{x, y, z\}$.[9] Since $p(y|\{x, y, z\}) > 0$, the relative choice probabilities of z and y in $\{x, y, z\}$ must then be $\frac{p(z|\{x,y,z\})}{p(y|\{x,y,z\})} = \frac{w_x + w_z}{w_y}$. Additionally, Axiom 1 implies that $p(y|\{y, z\}) \in (0, 1)$. Hence, as argued in case of the first observation, $p(z|\{y, z\})/p(y|\{y, z\}) = w_z/w_y$. Together, these imply

$$\frac{p(z|\{x, y, z\})}{p(y|\{x, y, z\})} - \frac{p(z|\{y, z\})}{p(y|\{y, z\})} = \frac{w_x}{w_y},$$

uniquely identifying α_{xy}. The third type of observation is indirect. It follows from $\alpha_{x_i x_{i+1}} = \frac{w_{x_i}}{w_{x_{i+1}}}$ for each $i = 1, ..., n$. Hence, the $\alpha_{x_i x_{i+1}}$ values multiply to 1 and knowing $n - 1$ of them reveals the last one. An implication of this observation is that if α_{yx} is known, α_{xy} can be revealed as $\alpha_{xy} = 1/\alpha_{yx}$. This establishes that any RAR representation of p must be consistent with the above revelations.

Given choice data p that is consistent with the RAR model, let

$$xWy \text{ if } \alpha_{xy} \text{ is defined by one of the three patterns above.}$$

First note that W is transitive and symmetric, as stated in the last observation. However, it may not be complete. For example, in case of deterministic choice data generated by maximization of a single reference-independent preference relation, $W = \emptyset$. As we argued above, W must be part of the revealed relative reference weights. Furthermore, there is no other revelation: any weight vector $\{w_x\}_{x \in X}$ that respects α_{xy} for all $(x, y) \in W$ represents p. In other words, if $(x, y) \notin W$, then for any positive real number $\sigma \in \mathbb{R}_+$, we can find a representation where the reference weight of x relative to y is σ. The following remark establishes this point. The proof is identical to the proof of Theorem 1.

Remark 2. ***(Revealed Reference Weights)*** *Suppose p admits a RAR representation. Then the reference weight of x relative to y is revealed to be α_{xy} if and only if xWy.*

[9] As argued in the previous paragraph, the observation $p(z|\{x, y, z\}) > p(z|\{y, z\})$ reveals z is $\succ_x$-best in $\{x, y, z\}$ and SQB implies that it is $\succ_z$-best as well.

In the following examples, we illustrate how to find such revelations.

Example 1. *Consider the choice data on $X = \{x, y, z\}$, as given in Table 2.*

Revealed Preference: *To reveal P_z, note that $p(y|\{y,z\}) = 1$ implies yP_zz and $p(x|\{x,y,z\}) > p(x|\{x,y\})$ implies xP_zy. Together, we have xP_zyP_zz. To reveal P_y, note that $p(y|\{x,y,z\}) > 0$ implies yP_yx and yP_yz; and $p(x|\{x,z\}) = 1$ implies xP_yz. Together, we have yP_yxP_yz. To reveal P_x, note that $p(x|\{x,y,z\}) > 0$ implies xP_xy and xP_xz; and $p(y|\{y,z\}) = 1$ implies yP_xz. Together, we have xP_xyP_xz.*

Revealed Relative Reference Weights: *From the set $\{x,y\}$, we observe $\alpha_{xy} = \frac{p(x|\{x,y\})}{p(y|\{x,y\})} = \frac{1}{2}$. Then, from the set $\{x,y,z\}$,*

$$\alpha_{zy} = \frac{p(x|\{x,y,z\})}{p(y|\{x,y,z\})} - \frac{p(x|\{x,y\})}{p(y|\{x,y\})} = \frac{2}{1} - \frac{1}{2} = \frac{3}{2}.$$

Hence, $\alpha_{yz} = \frac{2}{3}$ and $\alpha_{xz} = \alpha_{xy}\alpha_{yz} = \frac{1}{3}$. For example, normalizing $w_x = 1$, this gives us $w_y = 2$ and $w_z = 3$.

Table 2: The choice data in Example 1.

$p(\cdot\|S)$	$\{x,y,z\}$	$\{x,y\}$	$\{x,z\}$	$\{y,z\}$
x	2/3	1/3	1	–
y	1/3	2/3	–	1
z	0	–	0	0

Example 2. *Consider the choice data on $X = \{x, y, z\}$, as given in Table 3.*

Revealed Preference: *To reveal P_z, note that $p(z|\{y,z\}) > 0$ implies zP_zy and $p(x|\{x,y,z\}) = 1$ implies xP_zz. Together, we have xP_zzP_zy. Similarly, xP_yyP_yz. To reveal P_x, note that note that $p(x|\{x,y,z\}) = 1$ implies xP_xy and xP_xz. There are no more revelations. Thus, both xP_xyP_xz and xP_xzP_xy are possible.*

Revealed Relative Reference Weights: *From the set $\{y,z\}$ we learn $\alpha_{yz} = \frac{2}{3}$. There is no other revelation. For example, fixing $w_y = 2$, we have $w_z = 3$, but w_x can be chosen to be any value.*

5 A General Model

In the previous section, we showed that RAR is a very tractable model, that is, its components can almost be fully identified. In this section, we analyze whether these strong revelations

Table 3: The choice data in Example 2

$p(\cdot\|S)$	$\{x,y,z\}$	$\{x,y\}$	$\{x,z\}$	$\{y,z\}$
x	1	1	1	–
y	0	0	–	2/5
z	0	–	0	3/5

are due to RAR's parametric structure in reference point formation. To this end, we study a generalization of RAR where the reference formation process is only assumed to satisfy two basic conditions: the reference probabilities (i) are strictly positive, and (ii) satisfy strict regularity (*i.e.*, the reference probability of an alternative is decreasing in the menu size). Both conditions are satisfied by RAR, as well as a range of other reference formation processes.

Formally, the decision maker has a set of reference-dependent preferences $\{\succ_x\}_{x\in X}$ and context-dependent reference weights $\rho(\cdot|S)$. The expression $\rho(x|S)$ represents the likelihood that alternative x will be the reference point in S. Assume that ρ satisfies (i) $\rho(x|S) > 0$ for all $x \in S$ (positivity), and (ii) $\rho(x|S) < \rho(x|T)$ for all $x \in T \subsetneq S$ (strict regularity). As before, we assume that reference-dependent preferences do not allow for indifferences, a standard assumption in the discrete random choice literature. We also continue to assume that reference-dependent preferences $\{\succ_x\}_{x\in X}$ satisfy SQB. The following definition formally states the choice process in our model.

Definition 2. *A choice rule p is consistent with the* ***generalized random reference model (GAR)*** *if there exist a family $\{\succ_x\}_{x\in X}$ of reference-dependent preferences satisfying SQB and context-dependent reference weights ρ satisfying positivity and strict regularity such that for each $S \in \mathcal{X}$ and $x \in S$,*

$$p(x|S) = \sum_{y\in S} \rho(x|S)\,\mathbb{1}(x \textit{ is } \succ_y\textit{-best in } S).$$

When the above definition is satisfied, we also say $(\{\succ_x\}_{x\in X}, \rho)$ represents p, or p admits a GAR representation. Note that RAR is a special case of GAR where $\rho(x|S) = \frac{w(x)}{\sum_{y\in S} w(y)}$.

As in Section 4, we say y is revealed preferred to z under x if for every GAR representation $(\{\succ_x\}_{x\in X}, \rho)$ of p, we have $y \succ_x z$. We now illustrate that the revealed reference-dependent preference of this general model is the same as the parametric model. To this end, let P_x be defined as in the previous section.

Remark 3. ***(Revealed Preference)*** *Suppose p admits a GAR representation. Then y is* *revealed to be preferred to z under reference point x if and only if yP_xz.*

To see this, first assume that $p(y|S) = 1$. Since every $x \in S$ becomes the reference point with positive probability, that is $\rho(x|S) > 0$, we must have y is better than z when the reference point is x, $y \succ_x z$. Note that then $y \succ_z z$ as well. Hence, by SQB, for any $x \notin S$, $y \succ_x z$ also holds. This is the first condition in our revealed preference definition. Second, assume that $p(y|S) > 0$. Then there is a reference point t where $y \succ_t z$ for every $z \in S \setminus y$. By SQB, we must then have $y \succ_y z$. Finally, assume $p(y|S) > p(y|S \setminus x)$. Assume there exists $z \in S$ such that $z \succ_x y$. Therefore, removing x does not shrink the set of alternatives which place y as the top ranked alternative in S. Since $\rho(z|S \setminus x) > \rho(z|S)$ for every z in S, we must have $p(y|S) < p(y|S \setminus x)$, a contradiction. Hence, $y \succ_x z$. This discussion illustrates that these three conditions are necessary for the revealed preference. To see why they are sufficient, assume that we have not revealed the relative ranking of $x, y \in X$ under the reference z using our definition. Then, it must be the case that $z \succ_z x$, $z \succ_z y$, $x \succ_x y$ and $y \succ_y x$. But if all these conditions are satisfied, the relative ranking of x and y under z cannot impose any restrictions on observed choices. Hence, our definition reveals us everything that can be revealed about the preference, so it has to be sufficient as well.

This result is important for two reasons. First, our reference-dependent preference revelations are not driven by the parametric structure of RAR. Second, any generalization of RAR that satisfies the two basic assumptions of GAR shares the same revelations about the reference-dependent preferences.

Next, we discuss the behavioral postulates that characterize GAR. First, it is routine to show that GAR still satisfies the ordinal Axioms 1-4. Unfortunately, these axioms are not sufficient. Our next example illustrates this point. The choice rule presented in Table 5 satisfies Axioms 1-4, but cannot be represented by GAR. To see this, suppose not. By the previous discussion on revealed preferences, we must have $x \succ_y y \succ_y z$, $x \succ_x y, z$ and $z \succ_z x, y$. Since $p(y|\{y,z\}), p(z|\{y,z\}) \in (0,1)$, $\rho(y|\{y,z\}) = p(y|\{y,z\}) = 0.1$. Similarly, $\rho(x|\{x,z\}) = p(x|\{x,z\}) = 0.5$. By strict regularity, we must have $\rho(x|\{x,y,z\}) \leq 0.5$ and $\rho(y|\{x,y,z\}) \leq 0.1$. This is a contradiction since $p(x|\{x,y,z\}) > 0.1 + 0.5$.

Table 4: A choice rule that satisfies Axioms 1-4, but cannot be represented by GAR

$p(\cdot, S)$	$\{x,y,z\}$	$\{x,y\}$	$\{x,z\}$	$\{y,z\}$
x	0.7	1	0.5	–
y	0	0	–	0.1
z	0.3	–	0.5	0.9

The example in Table 5 shows that, for choice data to be representable by GAR, the size of regularity violations it exhibits must be bounded. The following axiom guarantees this is the case. Together with the previous axioms, it characterizes GAR. This axiom is closely related to Motzkin's transposition theorem [Motzkin, 1936], a member of the well-known theorems of the alternatives. It guarantees that the reference point formation rule satisfies positivity and strict regularity.

Before stating the axiom, we introduce the following notation. For any $x, y \in S \subseteq X$, let $\lambda_{xy}(S)$ denote a positive constant. Let λ stand for the vector consisting of all constants $\lambda_{xy}(S)$. For notational simplicity, let $\lambda_x(S)$ stand for $\lambda_{xx}(S)$. For any such λ, we define

$$V(\lambda, p) = \sum_{S \in \mathcal{X}} \sum_{x \in S \subseteq X} \lambda_x(S) p(x, S) + \sum_{S \in \mathcal{X}} \sum_{x,y \in S \subseteq X \text{ s.t. } x \neq y} \lambda_{xy}(S)(p(x, S \setminus y) - p(x, S)).$$

For any $x \in S$, let

$$\Gamma_\lambda(x, S) = \lambda_x(S) + \sum_{y \notin S} \lambda_{xy}(S \cup y) - \sum_{z \in S \setminus x} \lambda_{xz}(S)$$

Axiom 9. *For any $\lambda \geq 0$ that satisfies $\Gamma_\lambda(x, S) = \Gamma_\lambda(y, S)$ for all $x, y \in S$ such that either $p(x|S) = 1$ or $p(x, S) > p(x, S \setminus y)$, we have $V(\lambda, p) \geq 0$, with strict inequality if $\lambda \neq 0$.*

We can now state the characterization result.

Theorem 2. *A random choice rule p satisfies Axioms 1-4 and Axiom 9 if and only if it has a GAR representation.*

6 Conclusion

We provide a simple model of random reference-dependent choice (RAR) in which the reference point is endogenously determined à la Luce. The key innovation of the paper is the idea that observations of stochastic choice behavior might be a product of stochastic reference points even though, from the point of view of the decision maker, choices are deterministic.

We provide a set of behavioral postulates that characterize the empirical content of RAR. We also illustrate how to identify RAR's basic parameters from choice data. The underlying preferences can be uniquely identified whenever they affect choice and this strong identification is not due to the Luce structure on reference point formation. We show that a similar revelation holds in a much general model that we call GAR, as well as any model that lies "in between" RAR and GAR. We also provide a characterization of GAR.

Unlike the classical model of Luce, RAR can accommodate violations of regularity, as often displayed in experimental and empirical studies. In RAR regularity violations only happen when an alternative that is chosen with zero probability transfers its reference weight onto the maximizer of its reference-dependent preferences. Hence, elimination of an alternative chosen with zero probability induces a regularity violation on another. Having said that, it is rather straightforward to create a modification of RAR where all alternatives are chosen with positive probability, yet, regularity violations continue to occur. Consider a simple modified version of RAR model

$$p(x|S) = \sum_{y \in S} \left(\frac{w(y)}{\sum_{z \in S} w(z)} \right) (\varepsilon \mathbb{1}(x = y) + (1-\varepsilon)\mathbb{1}(x \text{ is } \succ_y\text{-best in } S))$$

where $\varepsilon > 0$. As $\varepsilon \to 0$, this model converges to RAR. It replaces the zero probabilities in RAR with ε, yet continues to accommodate regularity violations. This model has a bounded rationality interpretation where the decision maker maximizes her reference dependent preferences with probability $1-\varepsilon$ and sticks with her existing reference point with the remaining ε probability.

Another possible extension of RAR allows an exogenous reference point r to probabilistically affect the reference formation process:

$$p_r(x|S) = \sum_{y \in S} \left(\frac{w_r(y)}{\sum_{z \in S} w_r(z)} \right) \mathbb{1}(x \text{ is } \succ_y\text{-best in } S)$$

where $w_r(r) = w(r) + b(r)$ and $w_r(z) = w(z)$ for all $z \neq r$. In this model, the exogenously given reference point r receives a boost $b(r)$ to its original reference weight, and hence it becomes the reference point with a higher probability. As the boost $b(r) \to \infty$, this model converges to a deterministic choice model where the DM maximizes $\succ_r$ when her reference point is r. On the other hand, as $b(r) \to 0$, the model converges to RAR. Further analysis of such extensions is left for future research.

References

M. Agranov and P. Ortoleva. Stochastic choice and preferences for randomization. *Journal of Political Economy*, 125(1):40–68, 2017.

V. H. Aguiar. Random categorization and bounded rationality. *Economics Letters*, 159: 46–52, 2017.

A. Ahumada and L. Ulku. Luce rule with limited consideration. *Mathematical Social Sciences*, 93:52–56, 2018.

J. Apesteguia, M. A. Ballester, and J. Lu. Single-crossing random utility models. *Econometrica*, 85(2):661–674, 2017.

A. Baillon, H. Bleichrodt, and V. Spinu. Searching for the reference point. *Management Science*, 66(1):93–112, 2020.

C. Baldassi, S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and M. Pirazzini. A behavioral characterization of the drift diffusion model and its multialternative extension for choice under time pressure. *Management Science*, 66(11):5075–5093, 2020.

M. Baucells, M. Weber, and F. Welfens. Reference-point formation and updating. *Management Science*, 57(3):506–519, 2011.

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

S. Benartzi, A. Previtero, and R. H. Thaler. Annuitization puzzles. *Journal of Economic Perspectives*, 25(4):143–164, 2011.

N. A. Berkowitsch, B. Scheibehenne, and J. Rieskamp. Rigorously testing multialternative decision field theory against random utility models. *Journal of Experimental Psychology: General*, 143(3):1331, 2014.

R. Bodner and D. Prelec. The centroid model of context dependent choice. *Unpublished Manuscript, MIT*, 1994.

P. Bordalo, N. Gennaioli, and A. Shleifer. Salience and consumer choice. *Journal of Political Economy*, 121(5):803–843, 2013.

R. L. Brady and J. Rehbeck. Menu-dependent stochastic feasibility. *Econometrica*, 84(3): 1203–1223, 2016.

C. Camerer, L. Babcock, G. Loewenstein, and R. Thaler. Labor supply of New York City cabdrivers: One day at a time. *Quarterly Journal of Economics*, 112(2):407–441, 1997.

M. D. Cattaneo, X. Ma, Y. Masatlioglu, and E. Suleymanov. A random attention model. *Journal of Political Economy*, 128(7):2796–2836, 2020.

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

C. Chambers, Y. Masatlioglu, and C. Raymond. Weighted linear stochastic choice. Working paper, 2021.

M. Dean, Ö. Kıbrıs, and Y. Masatlioglu. Limited attention and status quo bias. *Journal of Economic Theory*, 169:93–127, 2017.

R. Dutta. Gradual pairwise comparison and stochastic choice. *Theoretical Economics*, 15(4): 1335–1364, 2020.

F. Echenique and K. Saito. General Luce model. *Economic Theory*, 68:811–826, 2019.

D. Eil and J. W. Lien. Staying ahead and getting even: Risk attitudes of experienced poker players. *Games and Economic Behavior*, 87:50–69, 2014.

E. Filiz-Ozbay and Y. Masatlioglu. Progressive Random Choice. *Working Paper*, 2020.

D. J. Freeman. Preferred personal equilibrium and simple choices. *Journal of Economic Behavior & Organization*, 143:165–172, 2017.

D. Fudenberg, R. Iijima, and T. Strzalecki. Stochastic choice and revealed perturbed utility. *Econometrica*, 83(6):2371–2409, 2015.

D. Genesove and C. Mayer. Loss aversion and seller behavior: Evidence from the housing market. *Quarterly Journal of Economics*, 116(4):1233–1260, 2001.

B. Guney and M. Richter. Costly switching from a status quo. *Journal of Economic Behavior & Organization*, 156:55–70, 2018.

S. Horan. Stochastic semi-orders. *Journal of Economic Theory*, 192:105171, 2021.

J. Huber, J. W. Payne, and C. Puto. Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, 9(1): 90–98, 1982.

E. J. Johnson and D. Goldstein. Do defaults save lives? *Science*, 302(5649):1338–1339, 2003.

E. J. Johnson, S. Bellman, and G. L. Lohse. Defaults, framing and privacy: Why opting in-opting out. *Marketing Letters*, 13(1):5–15, 2002.

D. Kahneman. Reference points, anchors, norms, and mixed feelings. *Organizational behavior and human decision processes*, 51(2):296–312, 1992.

D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–292, 1979.

Ö. Kıbrıs, Y. Masatlioglu, and E. Suleymanov. A theory of reference point formation. *Unpublished Paper*, 2018.

R. Kivetz, O. Netzer, and V. Srinivasan. Alternative models for capturing the compromise effect. *Journal of Marketing Research*, 41(3):237–257, 2004.

G. J. Koop and J. G. Johnson. The use of multiple reference points in risky decision making. *Journal of Behavioral Decision Making*, 25(1):49–62, 2012.

B. Kőszegi and M. Rabin. A model of reference-dependent preferences. *Quarterly Journal of Economics*, 121(4):1133–1165, 2006.

B. Kőszegi and M. Rabin. Reference-dependent risk attitudes. *American Economic Review*, 97(4):1047–1073, 2007.

M. Kovach and E. Suleymanov. Reference dependence and random attention. *arXiv 2106.13350*, 2021.

M. Kovach and G. Tserenjigmid. Behavioral foundations of nested stochastic choice and nested logit. *Available at SSRN 3437165*, 2019.

X. Z. Lim. Ordered reference dependent choice. *Working Paper*, 2020.

R. D. Luce. *Individual Choice Behavior*. John Wiley and Sons, Inc, 1959.

M. J. Machina. Stochastic choice functions generated from deterministic preferences over lotteries. *The Economic Journal*, 95(379):575–594, 1985.

A. Maltz. Exogenous endowment - endogenous reference point. *The Economic Journal*, 130 (625):160–182, 2020.

P. Manzini and M. Mariotti. Stochastic choice and consideration sets. *Econometrica*, 82(3): 1153–1176, 2014.

J. Marschak. Binary choice constraints on random utility indicators. Technical report, Cowles Foundation for Research in Economics, Yale University, 1959.

Y. Masatlioglu and E. A. Ok. Rational choice with status quo bias. *Journal of Economic Theory*, 121(1):1–29, 2005.

Y. Masatlioglu and E. A. Ok. A canonical model of choice with initial endowments. *Review of Economic Studies*, 81(2):851–883, 2014.

P. N. Mohr, H. R. Heekeren, and J. Rieskamp. Attraction effect in risky choice can be explained by subjective distance between choice alternatives. *Scientific Reports*, 7(1):1–10, 2017.

T. S. Motzkin. *Beiträge zur Theorie der linearen Ungleichungen*. Azriel Press, 1936.

A. Munro and R. Sugden. On the theory of reference-dependent preferences. *Journal of Economic Behavior & Organization*, 50(4):407–428, 2003.

T. Odean. Are investors reluctant to realize their losses? *The Journal of Finance*, 53(5): 1775–1798, 1998.

E. A. Ok, P. Ortoleva, and G. Riella. Revealed (p)reference theory. *American Economic Review*, 105(1):299–321, 2015.

A. Y. Orhun. Optimal product line design when consumers exhibit choice set-dependent preferences. *Marketing Science*, 28(5):868–886, 2009.

D. G. Pope and M. E. Schweitzer. Is Tiger Woods loss averse? Persistent vias in the face of experience, competition, and high stakes. *American Economic Review*, 101(1):129–157, 2011.

M. Rabin. Risk aversion and expected-utility theory: A calibration theorem. *Econometrica*, 68(5):1281–1292, 2000.

J. A. Rizzo and R. J. Zeckhauser. Reference incomes, loss aversion, and physician behavior. *Review of Economics and Statistics*, 85(4):909–922, 2003.

A. Rubinstein and Y. Salant. Two comments on the principle of revealed preference. *Unpublished paper*, 2006.

J. S. Sagi. Anchored preference relations. *Journal of Economic Theory*, 130(1):283–295, 2006.

Y. Salant and A. Rubinstein. (A, f): Choice with frames. *Review of Economic Studies*, 75 (4):1287–1296, 2008.

W. Samuelson and R. Zeckhauser. Status quo bias in decision making. *Journal of Risk and Uncertainty*, 1(1):7–59, 1988.

R. Sugden. Reference-dependent subjective expected utility. *Journal of Economic Theory*, 111(2):172–191, 2003.

P. Suppes and R. D. Luce. Preference, utility, and subjective probability. *Handbook of Mathematical Psychology*, 3:249–410, 1965.

J. Sydnor. (Over) insuring modest risks. *American Economic Journal: Applied Economics*, 2(4):177–199, 2010.

R. H. Thaler and S. Benartzi. Save more tomorrow$^{\mathrm{TM}}$: Using behavioral economics to increase employee saving. *Journal of Political Economy*, 112(S1):S164–S187, 2004.

L. Thurstone. A law of comparative judgement. *Pschological Review*, 34:273–286, 1927.

G. Tserenjigmid. Choosing with the worst in mind: A reference-dependent model. *Journal of Economic Behavior & Organization*, 157:631–652, 2019.

A. Tversky. Intransitivity of preferences. *Psychological Review*, 76(1):31–48, 1969.

A. Tversky and D. Kahneman. Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics*, 106(4):1039–1061, 1991.

P. P. Wakker. *Prospect Theory: For Risk and Ambiguity.* Cambridge University Press, 2010.

A Proof of Theorem 1

We first define reference-dependent preferences.

Definition 1. *For any x and $y \neq z$, let yP_xz if and only if there exists $S \supseteq \{y, z\}$ such that at least one of the following is satisfied:*

(i) $p(y|S) = 1$,

(ii) $p(y|S) > 0$ *and* $x = y$,

(iii) $p(y|S) > p(y|S \setminus x)$.

Claim 1. *If $p(\cdot|\cdot)$ satisfies Axioms 1-4, then P_x is transitive for all $x \in X$.*

Proof. Let yP_xzP_xt. We will show that yP_xt. There are nine cases (3 by 3) to consider. Each case is named according to the corresponding conditions in the definition.

(i)-(i): There exist $S \supseteq \{y, z\}$ and $T \supseteq \{z, t\}$ such that $p(y|S) = 1$ and $p(z|T) = 1$. By Axiom 2, $p(y|\{y, z\}) = 1$ and $p(z|\{z, t\}) = 1$. By Axiom 1, $p(y|\{y, z, t\}) = 1$, and yP_xt follows.

(i)-(ii): There exist $S \supseteq \{y, z\}$ and $T \supseteq \{z, t\}$ such that $p(y|S) = 1$ and $p(z|T) > 0$, where $x = z$. Axiom 2 implies $p(y|\{y, z\}) = 1$. Axiom 1 implies $p(z|\{z, t\}) > 0$ and $p(z|\{y, z, t\}) = 0$. Axiom 2 implies that either $p(y|\{y, z, t\}) = 1$ or $p(y|\{y, z, t\}) > p(y|\{y, t\})$. In both cases yP_zt follows, and since $x = z$, yP_xt follows.

(i)-(iii): There exist $S \supseteq \{y, z\}$ and $T \supseteq \{z, t\}$ such that $p(y|S) = 1$ and $p(z|T) > p(z|T \setminus x)$. Axiom 2 implies $p(y|\{y, z\}) = 1$. Hence, Axiom 1 implies $p(z|\{x, y, z, t\}) = 0$. Since $p(z|T) > p(z|T \setminus x)$, Axiom 1 implies $p(z|\{x, z\}) > 0$ and $p(z|\{z, t\}) > 0$. Then, Axiom 2 implies $p(z|\{x, z, t\}) > 0$. Moreover, since $p(z|T) > p(z|T \setminus x)$, Axiom 3 implies $p(x|T) = 0$, and hence Axioms 2 and 4 imply $p(x|\{x, z\}) = 0$. Then, Axiom 1 implies $p(x|\{x, y, z, t\}) = 0$. Note that, by Axioms 1 and 2, $p(t|\{x, y, z, t\}) < 1$ as $p(z|\{x, z, t\}) > 0$. In addition, since $p(z|T) > p(z|T \setminus x)$, Axiom 4 implies $p(t|\{x, y, z, t\}) \leq p(t|\{y, z, t\})$. Therefore, by Axiom 2, either $p(y|\{x, y, z, t\}) = 1$ or $p(y|\{x, y, z, t\}) > p(y|\{y, z, t\})$. In both cases yP_xt follows.

(ii)-(i): There exist $S \supseteq \{y, z\}$ and $T \supseteq \{z, t\}$ such that $p(y|S) > 0$ and $p(z|T) = 1$, where $x = y$. Axiom 1 implies $p(y|\{y, z\}) > 0$. Axiom 2 implies $p(z|\{z, t\}) = 1$. Therefore, Axiom 1 implies $p(t|\{y, z, t\}) = 0$. Since $p(y|\{y, z\}) > 0$ and $p(t|\{y, z, t\}) = 0$, by Axiom 2, we cannot have $p(y|\{y, z, t\}) = 0$. Hence, $p(y|\{y, z, t\}) > 0$ which implies yP_yt, and since $x = y$, yP_xt follows.

(ii)-(ii): There exist $S \supseteq \{y, z\}$ and $T \supseteq \{z, t\}$ such that $p(y|S) > 0$ and $p(z|T) > 0$, where $x = y$ and $x = z$. Since the definition of P_x requires that $y \neq z$ whenever yP_xz, this case is not possible.

(ii)-(iii): There exist $S \supseteq \{y, z\}$ and $T \supseteq \{z, t\}$ such that $p(y|S) > 0$ and $p(z|T) > p(z|T \setminus x)$, where $x = y$. Axiom 1 implies $p(y|\{y, z\}) > 0$. Axiom 3 implies $p(x|T) = 0$. Axioms 2 and 4 imply $p(x|\{x, z\}) = 0$. This contradicts $p(y|\{y, z\}) > 0$ as $x = y$. Hence, this case is not possible.

(iii)-(i): There exist $S \supseteq \{y, z\}$ and $T \supseteq \{z, t\}$ such that $p(y|S) > p(y|S \setminus x)$ and $p(z|T) = 1$. Axiom 1 implies $p(y|\{x, y\}) > 0$ and $p(y|\{y, z\}) > 0$, and hence Axiom 2 implies $p(y|\{x, y, z\}) > 0$. Since $p(z|T) = 1$, Axiom 2 implies $p(z|\{z, t\}) = 1$, and hence Axiom 1 implies $p(t|\{x, y, z, t\}) = 0$. Now since $p(y|S) > p(y|S \setminus x)$, Axiom 3 implies $p(x|S) = 0$, and hence Axioms 2 and 4 imply $p(x|\{x, y\}) = 0$. Therefore, Axiom 1 implies $p(x|\{x, y, z, t\}) = 0$. Note that, by Axioms 1 and 2, $p(z|\{x, y, z, t\}) < 1$ as $p(y|\{x, y, z\}) > 0$. Moreover, by Axiom 4, $p(z|\{x, y, z, t\}) \leq p(z|\{y, z, t\})$ as $p(y|S) > p(y|S \setminus x)$. Since $p(x|\{x, y, z, t\}) = 0$, by Axiom 2, either $p(y|\{x, y, z, t\}) = 1$ or $p(y|\{x, y, z, t\}) > p(y|\{y, z, t\})$. In both cases $yP_x t$ follows.

(iii)-(ii): There exist $S \supseteq \{y, z\}$ and $T \supseteq \{z, t\}$ such that $p(y|S) > p(y|S \setminus x)$ and $p(z|T) > 0$, where $x = z$. Since $x = z$ and $p(y|S) > p(y|S \setminus x)$, Axiom 3 implies $p(z|S) = 0$, and hence Axioms 2 and 4 imply $p(z|\{y, z\}) = 0$. Now, Axiom 1 implies $p(z|\{y, z, t\}) = 0$ and $p(z|\{z, t\}) > 0$. Hence, by Axiom 2, $p(t|\{y, z, t\}) < 1$. Moreover, by Axiom 4, $p(t|\{y, z, t\}) \leq p(t|\{y, t\})$. Therefore, by Axiom 2, either $p(y|\{y, z, t\}) = 1$ or $p(y|\{y, z, t\}) > p(y|\{y, t\}) > 0$. In both cases $yP_z t$, and hence $yP_x t$, follows.

(iii)-(iii): There exist $S \supseteq \{y, z\}$ and $T \supseteq \{z, t\}$ such that $p(y|S) > p(y|S \setminus x)$ and $p(z|T) > p(z|T \setminus x)$. By Axiom 3, $p(x|S) = p(x|T) = 0$. Axioms 2 and 4 imply $p(x|\{x, y\}) = p(x|\{x, z\}) = 0$. Then, Axiom 1 implies $p(x|\{x, y, z, t\}) = 0$. Since $p(z|T) > p(z|T \setminus x)$, by Axiom 1, $p(z|\{x, z\}) > 0$ and $p(z|\{z, t\}) > 0$, and hence, by Axiom 2, $p(z|\{x, z, t\}) > 0$. Therefore, by Axioms 1 and 2, $p(t|\{x, y, z, t\}) < 1$. Moreover, by Axiom 4, $p(t|\{x, y, z, t\}) \leq p(t|\{y, z, t\})$. By the same argument, $p(y|S) > p(y|S \setminus x)$ implies that $p(z|\{x, y, z, t\}) < 1$ and $p(z|\{x, y, z, t\}) \leq p(z|\{y, z, t\})$. Hence, by Axiom 2, either $p(y|\{x, y, z, t\}) = 1$ or $p(y|\{x, y, z, t\}) > p(y|\{y, z, t\})$. In both cases $yP_x t$ follows.

We have now shown that in all possible cases $yP_x t$ follows. Hence, P_x is transitive. $\square$

Now let $\succ_x$ be an arbitrary completion of P_x. The next claim shows that $\{\succ_x\}_{x \in X}$ satisfies the SQB property.

Claim 2. *If $y \succ_x z$, then $y \succ_y z$.*

Proof. Suppose $z \succ_y y$. Then it must be that there exists no $S \supseteq \{y, z\}$ with $p(y|S) > 0$. In particular, $p(z|\{y, z\}) = 1$. But then, by definition, $zP_x y$ which contradicts $y \succ_x z$. Hence, $y \succ_y z$ must be true. $\square$

Next, we let q and $\mathcal{T}$ be defined as in the main text. The next claim shows that, under Axiom 5, q is generated by the Luce rule.

Claim 3. *Suppose $p(\cdot|\cdot)$ satisfies Axioms 1-4. If q satisfies Axiom 5, then there exist weights $\{w_x\}_{x \in X}$ such that for any $S \in \mathcal{T}$ and $x \in S$,*

$$q(x|S) = \frac{w_x}{\sum_{y \in S} w_y}.$$

Proof. First, we construct a partition $\mathcal{P}$ of the set of alternatives as follows. Let x and y belong to the same partition element $P \in \mathcal{P}$ if there exist $S_1, \ldots, S_{N-1} \in \mathcal{T}$ and $\{x_1, \ldots, x_N\}$

such that $x_1 = x$, $x_N = y$, and $\{x_i, x_{i+1}\} \subseteq S_i$ for $i \in \{1, \dots, N-1\}$. We let $P(x)$ denote the partition element corresponding to the alternative x.

Next, we construct the weights as follows. Pick an arbitrary element $x \in X$ and let $w_x = 1$. Choose $y \in P(x)$ and suppose $S_1, \dots, S_{N-1} \in \mathcal{T}$ and $\{x_1, \dots, x_N\}$ are such that $x_1 = x$, $x_N = y$, and $\{x_i, x_{i+1}\} \subseteq S_i$ for $i \in \{1, \dots, N-1\}$. Then, we let

$$w_y = \prod_{i=1}^{N-1} \frac{q(x_{i+1}|S_i)}{q(x_i|S_i)}.$$

To see that w_y is well-defined, let $S'_1, \dots, S'_{K-1} \in \mathcal{T}$ and $\{x'_1, \dots, x'_K\}$ be such that $x'_1 = x$, $x'_K = y$, and $\{x_j, x_{j+1}\} \subseteq S'_j$ for $j \in \{1, \dots, K\}$. Then, by Axiom 5, we get

$$1 = \frac{q(x_2|S_1)}{q(x_1|S_1)} \cdots \frac{q(x_{i+1}|S_i)}{q(x_i|S_i)} \cdots \frac{q(x_N|S_{N-1})}{q(x_{N-1}|S_{N-1})} \frac{q(x'_{K-1}|S'_{K-1})}{q(x'_K|S'_{K-1})} \cdots \frac{q(x'_j|S'_j)}{q(x'_{j+1}|S'_j)} \cdots \frac{q(x'_1|S'_1)}{q(x'_2|S'_1)},$$

since $x_N = x'_K$, $x_1 = x'_1 = x$, and $q(x|\{x\}) = 1$. Hence,

$$w_y = \prod_{i=1}^{N-1} \frac{q(x_{i+1}|S_i)}{q(x_i|S_i)} = \prod_{j=1}^{K-1} \frac{q(x'_{j+1}|S'_j)}{q(x'_j|S'_j)}.$$

If $X \setminus P(x)$ is empty, then we are done. Otherwise, pick an alternative $y \notin P(x)$ and repeat the procedure above until we are done.

Now let $S \in \mathcal{T}$ be given. If S is a singleton, the claim follows trivially. Hence, let $y, z \in S$ be given. Then, we know that $S \subseteq P(x)$ for some $x \in X$. Let $S_1, \dots, S_{N-1}$ and $\{x_1, \dots, x_N\}$ be such that $x_1 = x$, $x_N = y$, and $\{x_i, x_{i+1}\} \subseteq S_i$ for $i \leq N-1$, and let $S'_1, \dots, S'_{K-1}$ and $\{x'_1, \dots, x'_K\}$ be such that $x'_1 = x$, $x'_K = z$, and $\{x'_j, x'_{j+1}\} \subseteq S'_j$ for $j \leq K-1$. Now, by Axiom 5,

$$\frac{q(y|S)}{q(z|S)} = \underbrace{\frac{q(x_N|S_{N-1})}{q(x_{N-1}|S_{N-1})} \cdots \frac{q(x_{i+1}|S_i)}{q(x_i|S_i)} \cdots \frac{q(x_2|S_1)}{q(x_1|S_1)}}_{w_y} \underbrace{\frac{q(x'_1|S'_1)}{q(x'_2|S'_1)} \cdots \frac{q(x'_{j+1}|S'_j)}{q(x'_j|S'_j)} \cdots \frac{q(x'_{K-1}|S'_{K-1})}{q(x'_K|S'_{K-1})}}_{1/w_z},$$

since $x_N = y$, $x'_K = z$, and $x_1 = x'_1 = x$. Hence, we get that for any $S \in \mathcal{T}$ and any $y, z \in S$,

$$\frac{q(y|S)}{q(z|S)} = \frac{w_y}{w_z}.$$

It then follows that

$$q(y|S) = \frac{w_y}{\sum_{t \in S} w_t},$$

as desired. $\square$

We next show that Axioms 1-5 guarantee that the characterization theorem holds for all sets with at most three alternatives.

Claim 4. *Suppose $p(\cdot|\cdot)$ satisfies Axioms 1-5, and let $\{\succ_x\}_{x\in X}$ and $\{w_x\}_{x\in X}$ be defined as in the previous claims. Then, for any S with $|S|\leq 3$ and $x\in S$,*

$$p(x|S)=\sum_{y\in S}\frac{w_y}{\sum_{z\in S}w_z}\mathbb{1}(x=\arg\max(\succ_y,S)).$$

Proof. Let S with $|S|\leq 3$ and $x\in S$ be given. There are a few cases to consider.

Case 1: $p(x|S)=1$ for some $x\in S$. Then, by our definition of $\succ_y$, $x\succ_y z$ for any $y,z\in S$. Since x is $\succ_y$-maximal element in S for all $y\in S$, the characterization follows.

Case 2: $p(x|S)>0$ for all $x\in S$. B our definition of $\succ_x$, each x is $\succ_x$-best element in S. Moreover, by our definition of $q(\cdot|\cdot)$, we have that $p(x|S)=q(x|S)$. Since by Claim 3

$$q(x|S)=\frac{w_x}{\sum_{y\in S}w_y},$$

the characterization follows.

Case 3: $S=\{x,y,z\}$, $p(x|S)>0$, $p(y|S)>0$, and $p(z|S)=0$. Note that by our definition of $\succ_x$ and $\succ_y$, x and y are $\succ_x$ and $\succ_y$ maximal elements in S, respectively. Moreover, by Axioms 2 and 4, exactly one of $p(x|S)>p(x|S\setminus z)$ and $p(y|S)>p(y|S\setminus z)$ must hold. Without loss, suppose $p(y|S)>p(y|S\setminus z)$. Then, we must have that y is $\succ_z$-maximal element in S, which shows that the representation holds for z. In addition, by our definition of $q(\cdot|\cdot)$, we have $p(x|S)=q(x|S)$. Since by Claim 3

$$q(x|S)=\frac{w_x}{w_x+w_y+w_z},$$

the representation follows for x. Since $p(y|S)=1-p(x|S)$, the representation also follows for y. This concludes the proof of Claim 4. $\square$

The next claim shows that if we assume Axioms 1-6, the characterization holds for all sets where all alternatives are chosen with positive probability.

Claim 5. *Suppose $p(\cdot|\cdot)$ satisfies Axioms 1-6, and let $\{\succ_x\}_{x\in X}$ and $\{w_x\}_{x\in X}$ be defined as in the previous claims. Then, for any S such that $p(y|S)>0$ for all $y\in S$, and any $x\in S$,*

$$p(x|S)=\sum_{y\in S}\frac{w_y}{\sum_{z\in S}w_z}\mathbb{1}(x=\arg\max(\succ_y,S)).$$

Proof. Let S be such that $p(y|S)>0$ for all $y\in S$ and let $x\in S$. By Axiom 1, $p(x|\{x,y\})>0$ for any $y\in S$. In addition, by our definition of $\succ_y$, each $y\in S$ is $\succ_y$-maximal element in S. Hence, by Claim 4, for any $y\in S$,

$$\frac{p(x|\{x,y\})}{p(y|\{x,y\})}=\frac{w_x}{w_y}.$$

Now let $z\in S$ be given. By Axiom 3, we get $p(z|S)\leq p(z|S\setminus t)$ for all $t\in S\setminus z$. Hence, by

Axiom 6,

$$\frac{p(x|S)}{p(y|S)} = \frac{p(x|S \setminus z)}{p(y|S \setminus z)}.$$

for any $y \in S \setminus z$. Now note that if $\{x, y\} \subseteq T \subseteq S$, then Axioms 1 and 2 imply that $p(t|T) > 0$ for all $t \in T$. Hence, by repeatedly applying the above reasoning, we get that for any $y \in S \setminus x$,

$$\frac{p(x|S)}{p(y|S)} = \frac{p(x|\{x,y\})}{p(y|\{x,y\})} = \frac{w_x}{w_y}.$$

Hence,

$$\frac{1 - p(x|S)}{p(x|S)} = \frac{\sum_{y \in S \setminus x} w_y}{w_x} \quad \Rightarrow \quad p(x|S) = \frac{w_x}{\sum_{y \in S} w_y},$$

as desired. Since x was arbitrary, this concludes the proof of the claim. □

For any S and $x \in S$, let $A(x|S)$ denote the set of alternatives in S for which x is the maximal element in S:

$$A(x|S) = \{y \in S| \, x = \arg\max(\succ_y, S)\}.$$

Hence, the representation we want to prove can alternatively be stated as

$$p(x|S) = \frac{\sum_{a \in A(x|S)} w_a}{\sum_{b \in S} w_b}.$$

The last claim shows that Axioms 1-8 are sufficient for the representation.

Claim 6. *Suppose $p(\cdot|\cdot)$ satisfies Axioms 1-8, and let $\{\succ_x\}_{x \in X}$ and $\{w_x\}_{x \in X}$ be defined as in the previous claims. Then, for any S and $x \in S$,*

$$p(x|S) = \frac{\sum_{a \in A(x|S)} w_a}{\sum_{y \in S} w_b}.$$

Proof. Note that we have already proven the result for S with $|S| \leq 3$. We will extend the result for all S by induction. Suppose the characterization holds for all S with $|S| \leq n$, where $n \geq 3$, and let $S \ni x$ with $|S| = n + 1$ be given. If $p(x|S) = 1$ for some $x \in S$, then the same argument used in Claim 4 can still be used to show the characterization. Hence, we can assume that $p(x|S) = 1$ for no $x \in S$. In addition, if $p(x|S) > 0$ for all $x \in S$, then Claim 5 guarantees that the representation holds. Hence, suppose there exists $z \in S$ such that $p(z|S) = 0$. There are two cases to consider.

Case 1: $p(z|S) = 0$ and $p(x|S) > 0$ for all $x \in S \setminus z$. By definition of $\succ_z$, we should have that $z \in A(x|S)$ for some $x \in S \setminus z$, and the representation holds for z. In addition, every $x \in S \setminus z$ is $\succ_x$-maximal in S. Hence, assume $A(x|S) = \{x, z\}$ for some $x \in S \setminus z$ and $A(y|S) = \{y\}$ for $y \in S \setminus \{x, z\}$. Since $|S| \geq 4$, there exist at least two alternatives $y, y' \in S \setminus x$ such that $p(y|S)p(y'|S) > 0$. By Axioms 3 and 4, $p(y|S) \leq p(y|S \setminus t)$ for all $t \in S \setminus y$ and

$p(y'|S) \leq p(y'|S \setminus t)$ for all $t \in S \setminus y'$. By Axiom 6, for any $t \in S \setminus \{x, y, z\}$,

$$\frac{p(t|S)}{p(x|S)} = \frac{p(t|S \setminus y)}{p(x|S \setminus y)}.$$

In addition, by Axiom 6,

$$\frac{p(y|S)}{p(x|S)} = \frac{p(y|S \setminus y')}{p(x|S \setminus y')}.$$

By induction argument, we have

$$\frac{p(t|S \setminus y)}{p(x|S \setminus y)} = \frac{w_t}{w_x + w_z} \quad \text{and} \quad \frac{p(y|S \setminus y')}{p(x|S \setminus y')} = \frac{w_y}{w_x + w_z}.$$

Combining the previous two lines, we get that for any $t \in S \setminus \{x, z\}$,

$$\frac{1 - p(x|S)}{p(x|S)} = \frac{\sum_{t \in S \setminus \{x,z\}} w_t}{w_x + w_z},$$

which implies

$$p(x|S) = \frac{w_x + w_z}{\sum_{y \in S} w_y} \quad \text{and} \quad p(t|S) = \frac{w_t}{\sum_{y \in S} w_y},$$

as desired.

Case 2: $p(z|S) = p(t|S) = 0$ for $z \neq t \in S$. There are two subcases to consider.

First, suppose $z, t \in A(x|S)$ for some $x \in S$. Since we assumed that $p(x|S) \neq 1$, Axiom 2 implies $p(x|S) > p(x|S \setminus z)$ and $p(x|S) > p(x|S \setminus t)$. By Axiom 8, for any $y \in S \setminus x$ with $p(y|S) > 0$, we have

$$\frac{p(x|S)}{p(y|S)} = \frac{p(x|S \setminus z)}{p(y|S \setminus z)} + \frac{p(x|S \setminus t)}{p(y|S \setminus t)} - \frac{p(x|S \setminus \{z, t\})}{p(y|S \setminus \{z, t\})}.$$

Since $A(y|S) = A(y|S \setminus z) = A(y|S \setminus t) = A(y|S \setminus \{z, t\})$, by induction argument, we get

$$\frac{p(x|S)}{p(y|S)} = \frac{\sum_{a \in A(x|S) \setminus z} w_a}{\sum_{b \in A(y|S)} w_b} + \frac{\sum_{a \in A(x|S) \setminus t} w_a}{\sum_{b \in A(y|S)} w_b} - \frac{\sum_{a \in A(x|S) \setminus \{z,t\}} w_a}{\sum_{b \in A(y|S)} w_b} = \frac{\sum_{a \in A(x|S)} w_a}{\sum_{b \in A(y|S)} w_b}.$$

Since this is true for any $y \in S \setminus x$ with $p(y, S) > 0$, we get

$$p(x|S) = \frac{\sum_{a \in A(x|S)} w_a}{\sum_{b \in S} w_b} \quad \text{and} \quad p(y|S) = \frac{\sum_{a \in A(y|S)} w_a}{\sum_{b \in S} w_b}.$$

Next, suppose $z \in A(x|S)$ and $t \in A(y'|S)$ for some $y' \neq x$. Hence, we should have $p(x|S) > p(x|S \setminus z)$ and $p(x|S) \leq p(x|S \setminus t)$. By Axiom 7, for any $y \neq x$ such that $p(y|S) > 0$, we have

$$\frac{p(x|S)}{p(y|S)} = \frac{p(x|S \setminus z)}{p(y|S \setminus z)} \frac{p(x|S \setminus t)}{p(y|S \setminus t)} \frac{p(y|S \setminus \{z, t\})}{p(x|S \setminus \{z, t\})}.$$

Note that for $y \neq x, y'$, we have $A(y|S) = A(y|S \setminus z) = A(y|S \setminus t) = A(y|S \setminus \{z,t\})$. Similarly, we have $A(x|S) \setminus z = A(x|S \setminus t) \setminus z = A(x|S \setminus z) = A(x|S \setminus \{z,t\})$. Hence, for $y \neq x, y'$, the induction hypothesis implies

$$\frac{p(x|S)}{p(y|S)} = \frac{\sum_{a\in A(x|S)\setminus z} w_a}{\sum_{b\in A(y|S)} w_b} \frac{\sum_{a\in A(x|S)} w_a}{\sum_{b\in A(y|S)} w_b} \frac{\sum_{b\in A(y|S)} w_b}{\sum_{a\in A(x|S)\setminus z} w_a} = \frac{\sum_{a\in A(x|S)} w_a}{\sum_{b\in A(y|S)} w_b}.$$

In addition, since $A(y'|S) \setminus t = A(y'|S \setminus z) \setminus t = A(y'|S \setminus t) = A(y'|S \setminus \{z,t\})$,

$$\frac{p(x|S)}{p(y'|S)} = \frac{\sum_{a\in A(x|S)\setminus z} w_a}{\sum_{b\in A(y'|S)} w_b} \frac{\sum_{a\in A(x|S)} w_a}{\sum_{b\in A(y'|S)\setminus t} w_b} \frac{\sum_{b\in A(y'|S)\setminus t} w_b}{\sum_{a\in A(x|S)\setminus z} w_a} = \frac{\sum_{a\in A(x|S)} w_a}{\sum_{b\in A(y'|S)} w_b}.$$

Combining the previous two lines, we get

$$p(x|S) = \frac{\sum_{a\in A(x|S)} w_a}{\sum_{b\in S} w_b} \quad \text{and} \quad p(y|S) = \frac{\sum_{a\in A(y|S)} w_a}{\sum_{b\in S} w_b},$$

as desired. This concludes the proof of the theorem.

□

B Proof of Theorem 2

Suppose p has a GAR representation $(\{\succ_x\}_{x\in X}, \rho)$. We first show that it satisfies Axiom 9.

Let $X = \{x_1, \ldots, x_N\}$ so that $|X| = N$. We first introduce a binary relation $\gg$ on $\mathcal{X}$. For any non-empty $S \subseteq X$, let $k(S)$ denote the smallest integer in $\{1, \ldots, N\}$ such that $x_{k(S)} \in S$. Let $S \gg S'$ if $|S| < |S'|$ or $|S| = |S'|$ and $k(S \setminus S') > k(S' \setminus S)$. Note that $\gg$ is irreflexive, and since $S \setminus S' \neq \emptyset$, $S' \setminus S \neq \emptyset$, and $(S \setminus S') \cap (S' \setminus S) = \emptyset$ for any $S \neq S'$ with $|S| = |S'|$, it is weakly connected. In addition, suppose $|S| = |S'| = |S''|$, $k(S \setminus S') > k(S' \setminus S)$, and $k(S' \setminus S'') > k(S'' \setminus S')$. Then, we must have that either $k(S' \setminus S) > k(S'' \setminus S')$ or $k(S'' \setminus S') > k(S' \setminus S)$ (equality is not possible since $(S' \setminus S) \cap (S'' \setminus S') = \emptyset$). Since S and S' have the same elements from $\{x_1, \ldots, x_{k(S'\setminus S)-1}\}$ with $x_{k(S'\setminus S)} \in S' \setminus S$, and S' and S'' have the same elements from $\{x_1, \ldots, x_{k(S''\setminus S')-1}\}$ with $x_{k(S''\setminus S')} \in S'' \setminus S'$, in both cases above $k(S \setminus S'') > k(S'' \setminus S)$ follows. Hence, $\gg$ is transitive. Let $m : \mathcal{X} \to \{0, 1, 2, \ldots, 2^N - 2\}$ be a numeric representation of $\gg$ that satisfies $m(S) > m(S')$ if and only if $S \gg S'$. Hence, $m(X) = 0$, $m(\{x_N\}) = 2^N - 2$, and etc.

We next define the vector of choice probabilities. For any $S \subseteq X$, let $p(\cdot|S)$ denote the vector of choice probabilities for the choice set S, where the probability corresponding to alternative x_i is placed higher in the vector (alternatively, has a lower row number) than alternative x_j if $i < j$. Note that if $x_i \notin S$, then $p(x_i|S) = 0$. Each $p(\cdot|S)$ is an $N \times 1$ vector. Let $\mathbf{p}$ denote the vector of choice probabilities $[p(\cdot, S)]_{S\subseteq X}$ stacked as follows: if $m(S) < m(S')$, then the $p(\cdot|S)$ is placed higher in the vector $\mathbf{p}$ than $p(\cdot|S')$. Hence, the rows between $m(S)N + 1$ and $(m(S) + 1)N$ in the vector $\mathbf{p}$ correspond to $p(\cdot|S)$. Note that $\mathbf{p}$ is a $(2^N - 1)N \times 1$ vector.

We next define the vector of reference probabilities. Let $\rho(\cdot|S)$ denote the $N \times 1$ vector of reference probabilities, where $\rho(x_i|S) = 0$ if $x_i \notin S$. We stack the reference probabilities $\rho(\cdot|\cdot)$ in the vector $\boldsymbol{\rho}$ in the same order as $\mathbf{p}$.

Next, we encode the preferences in a matrix as follows. For any $S \subseteq X$, let $A(S)$ denote an $N \times N$ matrix of zeros and ones such that

$$[A(S)]_{ij} = 1 \text{ if and only if } x_i = \arg\max(\succ_{x_j}, S) \text{ and } x_j \in S.$$

Note that if $p(\cdot|\cdot)$ has a GAR representation, then for any nonempty $S \subseteq X$,

$$A(S)\rho(\cdot|S) = p(\cdot|S).$$

We stack the matrices $A(S)$ in a matrix $\mathbf{A}$ as follows. For a matrix $\mathbf{A}$ and integers m, n, k, l such that $n \geq m$ and $l \geq k$, let $\mathbf{A}[m : n, k : l]$ denote the $(n - m + 1) \times (l - k + 1)$ matrix consisting of elements $\mathbf{A}_{ij}$ such that $i \in \{n, \dots, m\}$ and $j \in \{k, \dots, l\}$. Then, $\mathbf{A}$ is an $(2^N - 1)N \times (2^N - 1)N$ matrix given by

$$[\mathbf{A}_{ij}] = A(S) \quad \text{where} \quad i, j \in \{m(S)N + 1, \dots, (m(S) + 1)N\} \text{ for some } S \in \mathcal{X}$$

and $\mathbf{A}_{ij} = 0$ otherwise. By construction, if $p(\cdot|\cdot)$ has a GAR representation $(\{\succ_x\}_{x \in X}, \rho)$, then

$$\mathbf{A}\boldsymbol{\rho} = \mathbf{p}.$$

Our next step is to incorporate strict positivity and regularity constraints for reference probabilities in a matrix. It will be convenient to introduce the following function: for any $S \in \mathcal{X}$, let $n_S : \{1, \dots, N\} \to \{0, 1, \dots, N\}$ be given by

$$n_S(i) = \begin{cases} |\{x_j \in S | j \leq i\}| & \text{if } x_i \in S, \\ 0 & \text{if } x_i \notin S. \end{cases}$$

Now, let $I(X)$ denote the $N \times N$ identity matrix. For any non-empty $S \subseteq X$, let $I(S)$ denote $|S| \times N$ matrix where the rows in $I(X)$ which correspond to elements in $X \setminus S$ are eliminated. That is,

$$[I(S)]_{ij} = 1 \quad \text{if and only if} \quad n_S(j) = i.$$

We stack the matrices $I(S)$ in a $(\sum_{S \in \mathcal{X}} |S|) \times (2^N - 1)N$ matrix $\mathbf{B_1}$ as follows:

$$\begin{aligned} [\mathbf{B_1}]_{ij} = 1 \text{ if and only if } & j \in \{m(S)N + 1, \dots, (m(S) + 1)N\} \text{ for some } S \in \mathcal{X} \\ & \text{and } i = n_S(j - m(S)N) + \sum_{S' \in \mathcal{X} : m(S') < m(S)} |S'| \end{aligned}$$

The matrix $\mathbf{B_1}$ encodes the requirement that $\rho(x_i|S) > 0$ for any $x_i \in S \subseteq X$.

We will use the matrix $\mathbf{B_2}$ to encode the regularity requirement: $\rho(x_i|S) < \rho(x_i|S \setminus x_j)$ for any $x_i \in S \subseteq X$ and $x_j \in S \setminus x_i$. To this end, for any $x_i \in S \subseteq X$, let $B_2(x_i, S)$ denote the $|S \setminus x_i| \times (2^N - 1)N$ matrix where

$$[B_2(x_i, S)]_{kl} = \begin{cases} -1 & \text{if } l = m(S)N + i, \\ 1 & \text{if } k = n_{S\setminus x_i}(j), l = m(S \setminus x_j)N + i \text{ for } x_j \in S \setminus x_i, \\ 0 & \text{otherwise.} \end{cases}$$

Stack the matrices $B_2(x_i, S)$ in the matrix $B_2(S)$ where the matrix corresponding to the element $x_i \in S$ is ranked higher in the matrix than element $x_j \in S$ if $i < j$. Lastly, create a matrix $\mathbf{B_2}$ consisting of matrices $B_2(S)$, where $B_2(S)$ is placed higher in $\mathbf{B}_2$ than $B_2(S')$ if $m(S) < m(S')$.

Now, let

$$\mathbf{B} = \begin{bmatrix} \mathbf{B_1} \\ \mathbf{B_2} \end{bmatrix}.$$

By construction, $p(\cdot|\cdot)$ has a GAR representation $(\{\succ_x\}_{x \in X}, \rho)$ if and only if

$$\mathbf{A}\boldsymbol{\rho} = \mathbf{p} \quad \text{and} \quad \mathbf{B}\boldsymbol{\rho} > 0. \tag{1}$$

Let

$$\mathbf{C} = \begin{bmatrix} \mathbf{A} \\ -\mathbf{A} \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} \mathbf{p} \\ -\mathbf{p} \end{bmatrix}.$$

Then, we can alternatively write equation 1 as

$$\mathbf{C}\boldsymbol{\rho} \leq \mathbf{b} \quad \text{and} \quad \mathbf{B}\boldsymbol{\rho} > 0. \tag{2}$$

By Motzkin transposition theorem [Motzkin, 1936], the system described by equation 2 has a solution if and only if for all vectors $\mathbf{y} \geq 0$ and $\mathbf{z} \geq 0$,

$$\mathbf{C}^T\mathbf{y} = \mathbf{B}^T\mathbf{z} \quad \Rightarrow \quad \mathbf{b}^T\mathbf{y} \geq 0 \tag{3}$$

with strict inequality if $\mathbf{z} \neq 0$. Note that $\mathbf{C}^T\mathbf{y}$ and $\mathbf{B}^T\mathbf{z}$ are $(2^N - 1)N \times 1$ vectors.

Note that each row in the matrix $\mathbf{B}$ is associated either with a strict positivity or a regularity constraint, and for each row i in $\mathbf{B}$ there is a corresponding $z_i \geq 0$ in the vector $\mathbf{z}$. For row i such that $[\mathbf{B}]_i$ is associated with the positivity constraint $\rho(x_k|S) > 0$, we use the notation $\lambda_{x_k}(S) = z_i$ to indicate that row i represents the strict positivity requirement for x_k in S. Similarly, for row j such that $[\mathbf{B}]_j$ is associated with the constraint $\rho(x_k|S) < \rho(x_k|S \setminus x_l)$, we use the notation $\lambda_{x_k x_l}(S) = z_j$. Let $\boldsymbol{\lambda} = \mathbf{z}$. Then, we can rewrite equation 3 as

$$\mathbf{C}^T\mathbf{y} = \mathbf{B}^T\boldsymbol{\lambda} \quad \Rightarrow \quad \mathbf{b}^T\mathbf{y} \geq 0 \tag{4}$$

with strict inequality if $\boldsymbol{\lambda} \neq 0$.

Now, let $\kappa = (2^N - 1)N$. We can express $\mathbf{b}^T\mathbf{y}$ as

$$\mathbf{b}^T\mathbf{y} = \sum_{S \in \mathcal{X}} \sum_{i \in \{1,\ldots,N\}} p(x_i|S)(y_{m(S)N+i} - y_{\kappa+m(S)N+i}).$$

For any $S \in \mathcal{X}$, let η_S denote the function $\eta_S : \{1, \ldots, N\} \to \{1, \ldots, N\}$ such that

$$\eta_S(i) = \{j \in \{1, \ldots, N\} |\, x_j = \arg\max(\succ_{x_i}, S)\}.$$

Then, for $i \in \{1, \ldots, N\}$ and $S \in \mathcal{X}$,

$$[\mathbf{C}^T \mathbf{y}]_{m(S)N+i} = \begin{cases} y_{m(S)N+\eta_S(i)} - y_{\kappa+m(S)N+\eta_S(i)} & \text{if } x_i \in S, \\ 0 & \text{otherwise.} \end{cases}$$

and

$$[\mathbf{B}^T \boldsymbol{\lambda}]_{m(S)N+i} = \begin{cases} \lambda_{x_i}(S) + \sum_{x_k \notin S} \lambda_{x_i x_k}(S \cup x_k) - \sum_{x_j \in S \setminus x_i} \lambda_{x_i x_j}(S) & \text{if } x_i \in S, \\ 0 & \text{otherwise.} \end{cases}$$

Hence, for $x_i \in S \subseteq X$, the constraint in equation 4 implies

$$y_{m(S)N+\eta_S(i)} - y_{\kappa+m(S)N+\eta_S(i)} = \lambda_{x_i}(S) + \sum_{x_k \notin S} \lambda_{x_i x_k}(S \cup x_k) - \sum_{x_j \in S \setminus x_i} \lambda_{x_i x_j}(S). \tag{5}$$

In addition, since $p(x_i|S) = 0$ if and only if $\eta_S(i) \neq i$, using the above equations, we get

$$\begin{aligned} \mathbf{b}^T \mathbf{y} &= \sum_{S \in \mathcal{X}} \sum_{i \in \{1,\ldots,N\}} p(x_i|S) \big[\lambda_{x_i}(S) + \sum_{x_k \notin S} \lambda_{x_i x_k}(S \cup x_k) - \sum_{x_j \in S \setminus x_i} \lambda_{x_i x_j}(S)\big] \\ &= \sum_{S \in \mathcal{X}} \sum_{x_i \in S} \lambda_{x_i}(S) p(x_i|S) + \sum_{S \in \mathcal{X}} \sum_{x_i \in S} \sum_{x_j \in S \setminus x_i} \lambda_{x_i x_j}(S)(p(x_i|S \setminus x_j) - p(x_i|S)) \end{aligned}$$

Lastly, given the above equation, the only implication of equation 5 is that for $x_i, x_l \in S$ such that $\eta_S(i) = \eta_S(l)$, we have

$$\lambda_{x_i}(S) + \sum_{x_k \notin S} \lambda_{x_i x_k}(S \cup x_k) - \sum_{x_j \in S \setminus x_i} \lambda_{x_i x_j}(S) = \lambda_{x_l}(S) + \sum_{x_k \notin S} \lambda_{x_l x_k}(S \cup x_k) - \sum_{x_j \in S \setminus x_l} \lambda_{x_l x_j}(S).$$

Since for $x_i \neq x_l$ in S, $\eta_S(i) = \eta_S(l) = i$ if and only if $p(x_i|S) = 1$ or $p(x_i|S) > p(x_i|S \setminus x_l)$, letting

$$\Gamma_\lambda(x_i, S) = \lambda_{x_i}(S) + \sum_{x_k \notin S} \lambda_{x_i x_k}(S \cup x_k) - \sum_{x_j \in S \setminus x_i} \lambda_{x_i x_j}(S)$$

and

$$V(\lambda, p) = \sum_{S \in \mathcal{X}} \sum_{x_i \in S} \lambda_{x_i}(S) p(x_i|S) + \sum_{S \in \mathcal{X}} \sum_{x_i \in S} \sum_{x_j \in S \setminus x_i} \lambda_{x_i x_j}(S)(p(x_i|S \setminus x_j) - p(x_i|S))$$

yields the axiom.

To prove sufficiency, we first define preferences $\{\succ_x\}_{x \in X}$ as in the proof of Theorem 1. By Claim 1, Axioms 1-4 guarantee that preferences are well defined. We then define the vectors $\mathbf{p}$, $\boldsymbol{\rho}$, and matrices $\mathbf{A}$, $\mathbf{B}$ as in the proof of the necessity of Axiom 9. Here the vector $\boldsymbol{\rho}$ is unknown. Proving the representation is equivalent to showing that equation 1, and hence equation 2, holds for some $\boldsymbol{\rho}$. We can then use Motzkin transposition theorem to show that the representation holds if and only if equation 3, and hence equation 4 holds. Let $\eta_S(i)$ be defined as in the proof of the necessity. Notice that our definition of $\{\succ_x\}_{x\in X}$ and the axioms guarantee that (i) $\eta_S(i) \neq i$ if and only if $p(x_i|S) = 0$, and (ii) $\eta_S(i) = \eta_S(l) = i$ for $x_i \neq x_l$ in S if and only if $p(x_i|S) = 1$ or $p(x_i|S) > p(x_i|S \setminus x_l)$. The rest of the proof is the same as the proof of the necessity.

C Related Literature: An Extended Discussion

Our paper also contributes to the growing decision-theoretic literature on stochastic choice. Here we elaborate on the relationship between our model and several other models mention in the introduction.

Our paper is closely related to a few recent papers which generalize the Luce model (Ahumada and Ulku [2018], Echenique and Saito [2019], Horan [2021]) by relaxing the requirement that an alternative must be chosen from every choice problem once it is chosen from one choice problem. In these models, the DM first constructs a consideration set by eliminating "dominated alternatives". Within the DM's consideration set, choices are made according to the Luce rule. Hence, in these models IIA holds once we restrict attention to chosen alternatives. On the other hand, in RAR "dominated alternatives" influence the choice probability of chosen alternatives, and IIA does not necessarily hold even if we restrict attention to chosen alternatives. In addition, in contrast to these models, RAR requires that a regularity violation must occur when there is an alternative that is chosen with zero probability unless one of the alternatives is chosen with probability 1. The main difference is that while in RAR zero probabilities occur due to reference-dependent preferences, in these models they occur due to limited consideration. Hence, the ways in which RAR and these models address zero probabilities are distinct. All these models (including RAR) require that IIA is satisfied if all alternatives are always chosen with positive probability.

Kovach and Tserenjigmid [2019] study an extension of the Luce model, called Nested Stochastic Choice (NSC). In their model, the set of alternatives is partitioned into distinct categories. Each category X_i is endowed with a choice set dependent weight $v(X_i \cap S)$, and each alternative x is endowed with a weight $u(x)$. When faced with a choice problem, the decision maker first picks a category via the Luce rule with weights $v(X_i \cap S)$ and next picks an alternative within the category via the Luce rule with weights $u(x)$. Note that in their model IIA holds (i) for any two alternatives in the same category, and (ii) for any two alternatives that have degenerate (singleton) categories. Consider a choice rule p that has a RAR representation with preferences $z, t \succ_x x \succ_x y$ and any alternative other than x is the best when it is the reference point. We show that p is not NSC. First, since p is RAR, removing x from $\{x, y, z\}$, $\{x, y, t\}$, and $\{x, z, t\}$ will cause an IIA violation. Then, observation (i) implies that in any NSC representation y, z, and t must be in distinct categories. On the other hand, observation (ii) implies that x must be in the same category with one of y and z, one of y and t, and one of z and t. This is clearly impossible. Lastly, NSC allows $p(x|\{x, y\}) = 0$ but $p(x|S) > 0$ for some $S \supseteq \{x, y\}$ (a violation of Axiom 1). Hence, RAR and NSC are distinct in terms of observed choices as well as behavioral motivations.

Another strand of the literature focuses on the extensions or special cases of the random utility model (RUM). Apesteguia et al. [2017] studies a RUM where the support satisfies a single-crossing property. Relatedly, Filiz-Ozbay and Masatlioglu [2020] extends randomization over utilities to randomization over deterministic choice functions and study a random choice rule whose support satisfies a progressivity property.[10]

Another line of research studies random attention. Manzini and Mariotti [2014] analyze a model where attention is independent between alternatives and each alternative is considered with a fixed probability, called the attention parameter. Brady and Rehbeck [2016] consider a model where the attention set is formed á la Luce. Aguiar [2017] generalizes the Manzini and Mariotti [2014] model by relaxing the independence assumption.[11] Cattaneo et al. [2020] analyze a general non-parametric model of monotonic random attention.

Following the seminal work of Machina [1985], another important line of research analyzes stochastic choice as maximization of preferences over lotteries. Fudenberg et al. [2015] analyze perturbations of the expected utility model. This model satisfies regularity. Cerreia-Vioglio et al. [2019] analyze an alternative deliberate random choice model which allows regularity violations.

Our study is significantly different from the aforementioned papers on several dimensions. In particular, the stochastic choice behavior in our model is due to stochastic reference points and from the viewpoint of the decision maker, choices are deterministic. To the best of our knowledge, our paper is the first study of random endogenous reference formation.

[10] With the exception of the general model in Filiz-Ozbay and Masatlioglu [2020], which is irrefutable, none of these models are logically related to ours.

[11] Kovach and Suleymanov [2021] show that when there is a reference alternative the Manzini and Mariotti [2014] model is precisely the intersection of the Brady and Rehbeck [2016] and Aguiar [2017] models.