Odds Supermodularity and The Luce Rule*

Serhat Dogan[†] Kemal Yildiz [‡]

July 15, 2020

Abstract

We present a characterization of the Luce rule in terms of *positivity* and a new choice axiom called *odds supermodularity* that strengthens the *regularity* axiom. This new characterization illuminates a connection that goes unnoticed, and sheds light on the behavioral underpinnings of the Luce rule and its extensions from a different perspective. We show that *odds supermodularity* per se characterizes a structured extension of the Luce rule that accommodates zero probability choices. We identify the random choice model characterized via a stochastic counterpart of Plott (1973)'s *path independence* axiom, which strengthens *odds supermodularity*.

JEL Classification Numbers: D01, D07, D09.

Keywords: Random choice, Luce rule, supermodularity, zero probability choices, stochastic path independence.

*We are grateful to Ariel Rubinstein, Faruk Gul, Wolfgang Pesendorfer, and Pietro Ortoleva, seminar participants at Bilkent University, and anonymous referees for their helpful comments. Kemal Yıldız gratefully acknowledges the support from the Scientific and Research Council of Turkey (TUBITAK) under grant number 1059B191601712.

[†]Bilkent University, Ankara, Turkey. Email: dserhat@bilkent.edu.tr

[‡]Bilkent University, Ankara, Turkey. Email: kemal.yildiz@bilkent.edu.tr

1 Introduction

Luce rule –the most widely used random choice model in economics– asserts that each alternative has a fixed positive weight, and is chosen from a choice set proportionately to its relative weight. We present a new characterization of the Luce rule in which the Luce's axioms are substituted by a strengthening of the *regularity* axiom. The classical Luce's axioms, namely *independence of irrelevant alternatives* (IIA) and the *choice axiom*, play a fundamental role in characterizing many choice models that are widely used in empirical work. Examples include axiomatic characterizations of the *generalized discounted logit* processes (Fudenberg & Strzalecki (2015)), the *multinomial logit* (Cerreia-Vioglio et al. (2017)), and the *mixed logit* model (Saito (2017)).[1] This new characterization illuminates a connection that goes unnoticed, and sheds light on the behavioral underpinnings of the Luce rule and its extensions from a different perspective. As a strengthening of our new axiom, we propose a stochastic counterpart of Plott (1973)'s *path independence* axiom, and identify the random choice model it characterizes.

To introduce our new axiom, consider an alternative x that is chosen from a choice set S with positive probability. The *odds against x in S*, denoted by $\mathscr{O}(x,S)$, is the ratio of the probability that x is not chosen from S to the probability that x is chosen from S, that is

$$\mathscr{O}(x,S) = \frac{1-\rho_x(S)}{\rho_x(S)}.$$

If x is not chosen from a choice set S with positive probability, then $\mathscr{O}(x,S) = \infty$. For each pair of choice sets S and T that contain only x in their intersection, *odds supermodularity* requires the sum of the odds against x in S and T be less than or equal to the odds against x in $S \cup T$. That is,

$$\mathscr{O}(x,S) + \mathscr{O}(x,T) \leq \mathscr{O}(x,S \cup T).^{2}$$

[1]For the characterizations of these models, our new axiom, *odds supermodularity*, can replace the Luce's IIA or the *choice axiom*.

[2]In Section 3, we argue that this is equivalent to require that $\mathscr{O}(x,\cdot)$ is supermodular for each alternative x, i.e. for each pair of choice sets S and T that contain x, $\mathscr{O}(x,S) + \mathscr{O}(x,T) \leq \mathscr{O}(x,S \cup$

Odds supermodularity can be interpreted as a strengthening of the *regularity* axiom, which requires that if new alternatives are added to a choice set, then the choice probability of an existing alternative should not increase, equivalently, odds against an existing alternative should not decrease. *Odds supermodularity* strengthens *regularity* by requiring that the odds against an existing alternative increases at least additively as new alternatives are added to the choice set.

In Theorem 1, we show that *the preference-oriented Luce rules* (POLRs) are the only random choice functions that satisfy *odds supermodularity*. An POLR has two primitives: a *preference relation*[3] $\succsim$ that allows for indifferences, and a *weight function* v that assigns a positive real number to each alternative. From each choice set S, an agent first shortlists the $\succsim$-best alternatives in S, then chooses each shortlisted alternative with a probability that equals the alternative's relative weight in the shortlist. It follows as a corollary that Luce rules are the only random choice functions that satisfy *odds supermodularity* and *positivity*, which requires each alternative be chosen from each choice set with positive probability. After presenting our results, we compare and contrast *odds supermodularity* to Luce's original axioms.

In Theorem 1, we also provide another way to view POLRs. A random choice function is in the *closure of the Luce rules* if it is obtainable as the limit of a sequence of the Luce rules. This extension facilitates the use of Luce rule in applications. In this vein, Apesteguia & Ballester (2020) formulate and use the *closure of the Luce rules* to accommodate the Luce model into their framework to analyze a novel measure of goodness of fit. We show that POLRs are the only random choice functions that are in the *closure of the Luce rules.*[4]

The Luce rule satisfies the requirement of *odds supermodularity* as an equality. Since, in Theorem 1, we show that the converse is also true, it follows that *whenever odds supermodularity* is satisfied, it must be satisfied as an equality. Therefore, *odds*

$T) + \mathscr{O}(x, S \cap T)$. As we also discuss in Section 3, this is equivalent to require for each alternative x, $\mathscr{O}(x, \cdot)$ has *increasing marginal returns*.

[3]A preference relation is a complete and transitive binary relation on the finite alternative set.

[4]We are grateful to an anonymous referee for suggesting this connection.

(super)modularity can be thought as a stochastic counterpart of Plott (1973)'s *path independence* axiom, which requires that if the choice set is "split up" into smaller sets, and if the choices from the smaller sets are collected and a choice is made from the collection, the final result should be the same as the choice from the original choice set. Similarly, *odds (super)modularity* requires the odds likelihoods be independent of how the choice process is divided into parts.

In the second part, we propose a direct counterpart of *path independence* that strengthens *odds supermodularity*. Our *stochastic path independence* requires for each pair of choice sets S and T such that $S \cap T = \{x\}$,

$$\rho_x(S \cup T) = \rho_x(S) \cdot \rho_x(T).$$

In the deterministic choice setup, choice functions that satisfy *path independence* are the rational choice functions. This raises the question: Which random choice functions satisfy its stochastic counterpart? In Theorem 2, we show that the answer of this question is a specific POLR, in which each alternative is indifferent with at most one other alternative according to the preference relation that is to be maximized. It follows from this result that an agent who satisfies SPI can choose at most two alternatives from each choice set with positive probability. This clarifies that SPI is a rather restrictive condition in case of having many alternatives.

Since POLR allows an alternative to 'dominate' another, if both alternatives are available, then the dominated one is never chosen. However, the widely adopted formulation of the Luce rule discards this possibility,[5] which leads to documented problems due to 'zero probability choices' in estimating market outcomes. The common empirical method to deal with zero probability choices is to drop such sample points, or replace them with small numbers. However, in the international trade literature, Helpman et al. (2008) point out that most of the bias in the traditional estimates is due to the omission of the *extensive margin*, which refers to the exporters

[5]In the original formulation by Luce (1959), it is allowed that some alternatives are never chosen from any choice set. However, it is not allowed that an alternative may be chosen in a choice set but not in another, which is our focus in here. For more details see the recent paper by Horan (2018), who also extends the Luce's *connected domain* approach to accommodate zero choice probabilities.

that never trade with each other. As for the empirical literature in which Luce rule and its variants are widely used, problems emanating from ignoring zero probability choices have recently received attention. Both Gandhi et al. (2013) and Hortacsu & Joo (2016) argue that when samples with zero market shares are dropped or replaced by small shares, price coefficient estimates are either biased upward or the direction of the bias becomes unpredictable. Hortacsu and Joo demonstrate that ignoring zero choice probabilities can even result in upward-sloping demand curves.[6] As we discuss in Section 3, POLR turns out to be closely related to the empirical model developed by Hortacsu & Joo (2016) to accommodate zero market shares consistently.

2 The model

Given a finite alternative set X, any nonempty subset S is called a **choice set**. Let Ω denote the collection of all choice sets. A **random choice function** (RCF) p assigns each choice set $S \in \Omega$ a probability measure over S. We denote by $\rho_x(S)$ the probability that alternative x is chosen from choice set S. A preference relation $\succsim$ is a complete and transitive binary relation on X. We denote the strict part of $\succsim$ with $\succ$, and the indifference part of $\succsim$ with $\sim$. For each $S \in \Omega$, $max(S, \succsim)$ is the set of $\succsim$-best alternatives in S. Next, we define the preference-oriented Luce rule.

Definition 1 An RCF ρ is a **preference-oriented Luce rule** (POLR) if there is a preference relation $\succsim$ on X and a weight function $v : X \rightarrow \mathbb{R}_{++}$ such that for each choice set $S \in \Omega$,

$$\rho_x(S) = \begin{cases} \dfrac{v(x)}{\displaystyle\sum_{y \in max(S,\succsim)} v(y)} & \text{if } x \in max(S, \succsim), \\ 0 & \text{otherwise.} \end{cases}$$

In words, an agent first shortlists the $\succsim$-best alternatives from a choice set S, then chooses each shortlisted alternative with a probability that equals the alterna-

[6] Hortacsu & Joo (2016) illustrate this by using Dominick's supermarket cola sales scanner data, which covers 100 chain stores in the Chicago area for 400 weeks, from September 1989 to May 1997.

tive's relative weight in the shortlist. If no alternative is strictly better than another, that is if $\succ$ is empty, then we obtain the classical Luce rule.

To see an example of POLR, consider a set of alternatives with m-many binary attributes. That is, each alternative $x \in \{0,1\}^m$. An alternative x has attribute $i \in \{1,\ldots,m\}$, if $x_i = 1$. Now, consider a society consisting of rational agents who seek to maximize strict preferences. Assume that for each agent in the society, an alternative x *dominates* another alternative y, if x has more attributes than y. However, agents may be heterogeneous in terms of how they rank the alternatives that have equal number of attributes. The aggregate choice behavior of the society is a POLR, where the Luce weight of each alternative x with k-many attributes is the proportion of the agents who prefer x to all other alternatives with k-many attributes. Another natural example of *dominance* in this context might be coordinate-wise dominance. But, since coordinate-wise dominance is not *complete*, the resulting RCF is not a POLR.

In our analysis, we will provide another way of viewing the POLRs. An RCF ρ is in the **closure of the Luce rules**, if there is a sequence of Luce rules $\{\rho^n\}_{n\geq 1}$ that converges to ρ. That is, for each choice set S and alternative $x \in S$, $lim_{n\to\infty}\, \rho^n_x(S) = \rho_x(S)$. It follows from our Theorem 1 that an RCF is in the *closure of the Luce rules* if and only if it is a POLR.

3 Results

Let ρ be an RCF. For each $S \in \Omega$ and $x \in S$, the **odds against** x **in** S, denoted by $\mathcal{O}(x,S)$, is the ratio of the probability that x is not chosen from S to the probability that x is chosen from S, i.e. $\mathcal{O}(x,S) = \frac{1-\rho_x(S)}{\rho_x(S)}$ if $\rho_x(S) > 0$, and $\mathcal{O}(x,S) = \infty$ if $\rho_x(S) = 0$. Next, we introduce our new axiom.

Odds supermodularity: For each $S, T \in \Omega$ and $x \in X$ such that $S \cap T = \{x\}$,

$$\mathcal{O}(x,S) + \mathcal{O}(x,T) \leq \mathcal{O}(x, S \cup T).$$

As we discuss in the introduction, *odds supermodularity* strengthens the *regular-*

ity axiom by requiring that the odds against an existing alternative increases at least additively as new alternatives are added to the choice set. We can easily see that *odds supermodularity* is equivalent to require for each $x \in X$, $\mathcal{O}(x,\cdot)$ is supermodular, i.e. for each $S,T \in \Omega$ such that $x \in S \cap T$, $\mathcal{O}(x,S)+\mathcal{O}(x,T) \leq \mathcal{O}(x,S\cup T)+\mathcal{O}(x,S\cap T)$.[7] It is well known in the literature that supermodularity of a function $\mathcal{O}(x,\cdot)$ is equivalent to require the odds against x, $\mathcal{O}(x,\cdot)$, has **increasing marginal returns**, that is for each $S,T \in \Omega$ such that $x \in S \subset T$, and $y \in S \setminus \{x\}$, $\mathcal{O}(x,S) - \mathcal{O}(x,S\setminus\{y\}) \leq \mathcal{O}(x,T) - \mathcal{O}(x,T\setminus\{y\})$.[8]

In Theorem 1, we show that *odds supermodularity* characterizes preference-oriented Luce rules. Our new characterization of the Luce rule follows as a direct corollary to Theorem 1. The Luce rule satisfies the requirement of *odds supermodularity* as an equality. That is, a Luce rule ρ satisfies the following stronger condition.

Odds modularity: For each $S,T \in \Omega$ and $x \in X$ such that $S \cap T = \{x\}$,

$$\mathcal{O}(x,S)+\mathcal{O}(x,T) = \mathcal{O}(x,S\cup T).$$

In Theorem 1, we additionally show that an RCF satisfies *odds supermodularity* if and only if it satisfies *odds modularity*. It follows that there is no choice rule that satisfies *odds supermodularity* with a strict inequality for some pair of choice sets containing a single common alternative. A related question is if the POLR is the only RCF that satisfies *odds submodularity*, that is for each $S,T \in \Omega$ and $x \in X$ such that $S \cap T = \{x\}$, $\mathcal{O}(x,S)+\mathcal{O}(x,T) \geq \mathcal{O}(x,S\cup T)$. To see that this is not true, let $X = \{x,y,z\}$, and suppose $\rho_x(x,y) = \rho_y(y,z) = \rho_z(x,z) = 2/3$ and $\rho_x(X) = \rho_y(X) = \rho_z(X) = 1/3$. Clearly, ρ is not a POLR, but for each distinct $x,y,z \in X$, $\mathcal{O}(x,X) = 2 < 1/2+2 = \mathcal{O}_x(x,y)+\mathcal{O}_x(x,z)$. Next we present Theorem 1 and its proof.

[7]To see this, first let $T' = (T\setminus S)\cup\{x\}$. It follows from *odds supermodularity* that $\mathcal{O}(x,S)+\mathcal{O}(x,T') \leq \mathcal{O}(x,S\cup T)$ and $\mathcal{O}(x,T')+\mathcal{O}(x,S\cap T) = \mathcal{O}(x,T)$. If we subtract the second equality from the first inequality side by side, then we obtain that $\mathcal{O}(x,\cdot)$ is supermodular.

[8]The proof of the general result can be found, for example, in Moulin (1991).

Theorem 1 *For an RCF ρ, the following are equivalent:*

i. ρ *satisfies odds supermodularity.*

ii. ρ *satisfies odds modularity.*

iii. ρ *is a preference-oriented Luce rule.*

iv. ρ *is in the closure of the Luce rules.*

Proof. First, we present several observations that are critical to prove the result. For any given RCF ρ, define the binary relation $\succsim_\rho$ on X such that for each $x, y \in X$, $x \succsim_\rho y$ if and only if $\rho_x(\{x, y\}) > 0$. It follows that $x \succ_\rho y$ if and only if $\rho_x(\{x, y\}) = 1$. In showing that i implies iii, we use the $\succsim_\rho$ relation as the one that is to be maximized. In Lemma 3, we show that if an RCF ρ satisfies *odds supermodularity*, then $\succsim_\rho$ is complete and transitive. To prove this result, first, we prove the following two lemmas. Lemma 1 shows that if ρ satisfies *odds supermodularity*, then ρ satisfies *regularity*, i.e. for each $S, T \in \Omega$ such that $S \subset T$, and $x \in S$, $\rho_x(T) \leq \rho_x(S)$.

Lemma 1 *If an RCF ρ satisfies odds supermodularity, then ρ satisfies regularity.*

Proof. Let $T \in \Omega$ and $x \in T$. First, suppose $\rho_x(T) > 0$. Since for each $S \subset T$ such that $x \in S$, $\mathcal{O}(x, S) + \mathcal{O}(x, (T \setminus S) \cup \{x\}) \leq \mathcal{O}(x, T)$, we have $\mathcal{O}(x, S) \leq \mathcal{O}(x, T)$. It follows that $\rho_x(T) \leq \rho_x(S)$. If $\rho_x(T) = 0$, then the conclusion follows directly. ■

Lemma 2 *Let ρ be an RCF that satisfies odds supermodularity. For each $S \in \Omega$ and $x, y \in S$, if $\rho_x(S) > 0$ and $\rho_y(S) = 0$, then $\rho_x(\{x, y\}) = 1$.*

Proof. For each $S \in \Omega$ and $x, y \in S$, since ρ satisfies *odds supermodularity*, we have $\mathcal{O}(x, \{x, y\}) + \mathcal{O}(x, S \setminus \{y\}) \leq \mathcal{O}(x, S)$. Suppose $\rho_x(S) > 0$ and $\rho_y(S) = 0$, we show that $\rho_x(\{x, y\}) = 1$. We first show that $\rho_x(S \setminus \{y\}) = \rho_x(S)$. To see this, note that since ρ satisfies *regularity* by Lemma 1, for each $z \in S \setminus \{y\}$, $\rho_z(S \setminus \{y\}) \geq \rho_z(S)$. Suppose $\rho_x(S \setminus \{y\}) > \rho_x(S)$. Since the choice probabilities sum to one and $\rho_y(S) = 0$, there exists $z \in S \setminus \{x, y\}$ with $\rho_z(S \setminus \{y\}) < \rho_z(S)$. This is a contradiction. Now, since

$\rho_x(S \setminus \{y\}) = \rho_x(S)$, we have $\mathcal{O}(x, S \setminus \{y\}) = \mathcal{O}(x, S)$. Therefore, $\mathcal{O}(x, \{x, y\}) = 0$, which implies $\rho_x(\{x, y\}) = 1$. ■

Lemma 3 *If an RCF ρ satisfies odds supermodularity, then $\succsim_\rho$ is complete and transitive.*

Proof. Since either $\rho_x(\{x, y\}) \geq 0$ or $\rho_y(\{x, y\}) \geq 0$, $\succsim_\rho$ is complete. To see that $\succsim_\rho$ is transitive, by contradiction suppose there exist $x, y, z \in X$ such that $x \succsim_\rho y \succsim_\rho z$ and $z \succ_\rho x$, which means $\rho_x(\{x, z\}) = 0$. Now, consider $S = \{x, y, z\}$, since ρ satisfies *regularity* and $\rho_x(\{x, z\}) = 0$, $\rho_x(S) = 0$. Since $\rho_x(S) = 0$ and $x \succsim_\rho y$, it follows from the contrapositive of Lemma 2 that $\rho_y(S) = 0$. Therefore, $\rho_z(S) = 1$. But, since ρ satisfies *regularity*, we must have $\rho_z(\{y, z\}) = 1$, which contradicts $y \succsim_\rho z$. ■

Lemma 4 *Let ρ be an RCF that satisfies odds supermodularity. For each $S \in \Omega$ and $x \in S$, $\rho_x(S) > 0$ if and only if $x \in \max(S, \succsim_\rho)$.*

Proof. If part: For each $S \in \Omega$ and $x \in S$, $x \in \max(S, \succsim_\rho)$ implies for each $y \in S \setminus \{x\}$, $\rho_x(\{x, y\}) > 0$. Since $\rho_y(S) > 0$ for some $y \in S$, it follows from Lemma 2 that $\rho_x(S) > 0$. Only if part: Suppose $\rho_x(S) > 0$. We show that $x \in \max(S, \succsim_\rho)$. By contradiction, suppose there exists $y \in S$ such that $y \succ_\rho x$. It follows that $\rho_x(\{x, y\}) = 0$. Since ρ satisfies *regularity*, this implies that $\rho_x(S) = 0$, which contradicts $\rho_x(S) > 0$. ■

Next we present two critical results for the construction of the Luce weights.

Lemma 5 *For each $i, j \in \{1, 2, \ldots, n\}$, if $a_{ij} > 0$, $a_{ij}a_{ji} = 1$, and $c_j = \sum_{i=1}^{n} a_{ij}$, then*

i. $\sum_{j=1}^{n} \frac{1}{c_j} \leq 1$.

ii. $\sum_{j=1}^{n} \frac{1}{c_j} = 1$ *if and only if for each $i, j \in \{1, 2, \ldots, n\}$, there exist $\alpha_i, \alpha_j \in \mathbb{R}_+$ such that $a_{ij} = \frac{\alpha_i}{\alpha_j}$.*[9]

Proof. Let $c = \sum_{j=1}^{n} \frac{1}{c_j}$ First, we show that $c \leq 1$. To see this for each $k \in \{1, 2, \ldots, n\}$,

[9] We thank Selim Bahadir for his contribution in extending the result for more than three positive numbers.

let $x_k = \frac{1}{c_j}$. Next, let $j \in \{1, 2, \ldots, n\}$ be fixed. For each $k \in \{1, 2, \ldots, n\}$, let $u_k = x_k$ and $v_k = a_{kj}$. Now, Titu's Lemma[10] implies that

$$\sum_{k=1}^{n} \frac{u_k^2}{v_k} \geq \frac{\left(\sum_{k=1}^{n} u_k\right)^2}{\sum_{k=1}^{n} v_k} \tag{1}$$

It follows from (2) that for each $j \in \{1, 2, \ldots, n\}$,

$$\sum_{i=1}^{n} \frac{x_i^2}{a_{ij}} \geq \frac{\left(\sum_{i=1}^{n} x_i\right)^2}{\sum_{i=1}^{n} a_{ij}} \tag{2}$$

Since for each $i, j \in \{1, 2, \ldots, n\}$, $a_{ij}a_{ji} = 1$, we can replace a_{ij} with $\frac{1}{a_{ji}}$. If we also substitute $\frac{1}{c_i}$ instead of x_i, we obtain that for each $j \in \{1, 2, \ldots, n\}$,

$$\sum_{i=1}^{n} \frac{a_{ji}}{c_i^2} \geq c^2 \frac{1}{c_j} \tag{3}$$

Now, if we sum both sides of the inequality in (3) over j, then we obtain that

$$\sum_{j=1}^{n} \sum_{i=1}^{n} \frac{a_{ji}}{c_i^2} \geq c^2 \sum_{j=1}^{n} \frac{1}{c_j} = c^3 \tag{4}$$

Note that we also have

$$\sum_{j=1}^{n} \sum_{i=1}^{n} \frac{a_{ji}}{c_i^2} = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{a_{ji}}{c_i^2} = \sum_{i=1}^{n} \left(\frac{1}{c_i^2} \sum_{j=1}^{n} a_{ji} \right) = \sum_{i=1}^{n} \left(\frac{1}{c_i^2} c_i \right) = \sum_{i=1}^{n} \frac{1}{c_i} = c \tag{5}$$

But, now, (4) and (5) implies $c \geq c^3$. Since $c \geq 0$, it follows that $c \leq 1$.

To see that (ii) holds, Titu's Lemma tells that (2) holds as an equality if and only if there exists $\alpha \in \mathbb{R}_+$ such that for each $k \in \{1, 2, \ldots, n\}$, $u_k = \alpha v_k$. It follows that $c = 1$ if and only if for each $j \in \{1, 2, \ldots, n\}$ there exists $\alpha_j \in \mathbb{R}_+$ such that for each $i \in \{1, 2, \ldots, n\}$, $x_i = \alpha_j a_{ij}$. That is for each $i, j \in \{1, 2, \ldots, n\}$, $a_{ij} = \frac{x_i}{\alpha_j}$. ■

Lemma 6 *Let ρ be an RCF that satisfies odds supermodularity. For each distinct $x, y, z \in X$, if $\rho_x(x, y) \in (0, 1)$, $\rho_x(x, z) \in (0, 1)$, and $\rho_y(y, z) \in (0, 1)$, then $\frac{\mathcal{O}(x,\{x,z\})}{\mathcal{O}(x,\{x,y\})} = \frac{\rho_z(\{y,z\})}{\rho_y(\{y,z\})}$.*

[10] Titu's Lemma is a direct consequence of the Cauchy–Bunyakovsky–Schwarz inequality.

Proof. First, note that for each $S, T \in \Omega$ with $S \cap T = \{x\}$, we have $\rho_x(S \cup T) = \dfrac{1}{1 + \mathcal{O}(x, S \cup T)}$. Second, note that *odds supermodularity* implies $\rho_x(S \cup T) \leq \frac{1}{1+\mathcal{O}(x,S)+\mathcal{O}(x,T)}$. Now, consider the choice set $\{x, y, z\}$. It follows from these two observations that

$$\rho_x(x,y,z) \leq \frac{1}{\mathcal{O}(x,\{x,z\}) + 1 + \mathcal{O}(x,\{x,y\})} = \frac{\mathcal{O}(z,\{x,z\})}{1 + \mathcal{O}(z,\{x,z\}) + \mathcal{O}(z,\{x,z\})\mathcal{O}(x,\{x,y\})},$$
$$\rho_y(x,y,z) \leq \frac{1}{\mathcal{O}(y,\{x,y\}) + 1 + \mathcal{O}(y,\{y,z\})} = \frac{\mathcal{O}(x,\{x,y\})}{1 + \mathcal{O}(x,\{x,y\}) + \mathcal{O}(x,\{x,y\})\mathcal{O}(y,\{y,z\})},$$
$$\rho_z(x,y,z) \leq \frac{1}{\mathcal{O}(z,\{y,z\}) + 1 + \mathcal{O}(z,\{x,z\})} = \frac{\mathcal{O}(y,\{y,z\})}{1 + \mathcal{O}(y,\{y,z\}) + \mathcal{O}(y,\{y,z\})\mathcal{O}(z,\{x,z\})}.$$

Note that to obtain the rightmost terms, we multiply both the numerator and the denominator of the middle terms with $\mathcal{O}(z, \{x, z\})$, $\mathcal{O}(x, \{x, y\})$, and $\mathcal{O}(y, \{y, z\})$ respectively.

Now, for each $i, j \in \{x, y, z\}$, define $a_{ij} = \mathcal{O}(i, \{i, j\})$. Since for each $i, j \in \{x, y, z\}$, $\rho_i(i, j) \in (0, 1)$, we have $a_{ij} > 0$, and by definition of $\mathcal{O}(i, \{i, j\})$, we have $a_{ij}a_{ji} = 1$. Now, consider the sum of the three middle terms. These terms correspond to $\frac{1}{c_x}, \frac{1}{c_y}, \frac{1}{c_z}$ in the statement of Lemma 5. Thus, it follows from part i. of Lemma 5 that this sum is less than or equal to 1. On the other hand, since the sum of the three leftmost terms equals 1, the sum of the rightmost terms must be equal to 1. By part ii. of Lemma 5, we know that if this equality holds, then for each $i, j \in \{x, y, z\}$ there exist $\alpha_i, \alpha_j \in \mathbb{R}_+$ such that $\mathcal{O}(i, \{i, j\}) = \frac{\alpha_i}{\alpha_j}$. It follows that $\mathcal{O}(x, \{x, y\}) \cdot \mathcal{O}(y, \{y, z\}) \cdot \mathcal{O}(z, \{x, z\}) = 1$. If we substitute $\frac{1}{\mathcal{O}(x,\{x,z\})}$ instead of $\mathcal{O}(z, \{x, z\})$, then we get $\mathcal{O}(y, \{y, z\}) = \frac{\mathcal{O}(x,\{x,z\})}{\mathcal{O}(x,\{x,y\})}$. Since $\mathcal{O}(y, \{y, z\}) = \frac{\rho_z(\{y,z\})}{\rho_y(\{y,z\})}$, we get the desired conclusion. ∎

Now, we are ready to prove Theorem 1.

$(i \implies iii)$: Suppose ρ is an RCF that satisfies *odds supermodularity*. We construct a weak order $\succsim$ and a weight function $v : X \to \mathbb{R}_{++}$ that recovers the choices of ρ. We choose $\succsim$ as the $\succsim_\rho$ that is defined above. It follows from Lemma 3 that $\succsim_\rho$ is complete and transitive. It follows from Lemma 4 that for each $S \in \Omega$ and $x \in S$, $\rho_x(S) > 0$ if and only if $x \in \max(S, \succsim_\rho)$.

Next, we construct the weight function $v : X \to \mathbb{R}_{++}$. First for each $x \in X$, define $x_\sim = \{z \in X : x \sim_\rho z\}$. Next, we specify the weights for each of these equivalence classes. For each $x \in X$, if $x_\sim = \{x\}$, i.e., there is no $z \in X \setminus \{x\}$ with $z \sim_\rho x$, then let $v(x) = 1$. If there is a single $z \in X \setminus \{x\}$ with $z \sim_\rho x$, then let $v(x) = \rho_x(\{x, z\})$ and $v(z) = \rho_z(\{x, z\})$. If $|x_\sim| > 2$, then define $v(x) = 1$ for some fixed $x \in x_\sim$, and for each $z \in x_\sim \setminus \{x\}$, define $v(z) = \mathcal{O}(x, \{x, z\})$. Thus we complete the construction of the weight function v.

Now, let $\widehat{\rho}$ be the preference-oriented Luce rule defined by $\succsim_\rho$ and v. First, we show that for each $x, y \in X$ we have $\rho_x(\{x, y\}) = \widehat{\rho}_x(\{x, y\})$. To see this, note that for each $x, y \in X$, if $\rho_x(\{x, y\}) = 1$ ($\rho_x(\{x, y\}) = 0$), then $x \succ_\rho y$ ($y \succ_\rho x$). Therefore, $\widehat{\rho}_x(\{x, y\}) = 1$ ($\widehat{\rho}_x(\{x, y\}) = 0$). Next we show that if $x \sim_\rho y$, then $\frac{v(x)}{v(y)} = \frac{\rho_x(\{x,y\})}{\rho_y(\{x,y\})}$. Once we show this, then it directly follows that $\rho_x(\{x, y\}) = \widehat{\rho}_x(\{x, y\})$. To see this, let $x^* \in x_\sim$ such that $v(x^*) = 1$ and for each $z \in x_\sim$, $v(z) = \mathcal{O}(x^*, \{x^*, z\})$. If $x = x^*$ or $y = x^*$, then it directly follows from the definition of odds likelihood that $\frac{v(x)}{v(y)} = \frac{\rho_x(\{x,y\})}{\rho_y(\{x,y\})}$. If $x^* \notin \{x, y\}$, then it follows from Lemma 6 that $\frac{\mathcal{O}(x^*,\{x^*,x\})}{\mathcal{O}(x^*,\{x^*,y\})} = \frac{\rho_x(\{x,y\})}{\rho_y(\{x,y\})}$. Since $v(x) = \mathcal{O}(x^*, \{x^*, x\})$ and $v(y) = \mathcal{O}(x^*, \{x^*, y\})$, we obtain the desired conclusion.

Next, consider each $S \in \Omega$ and $x \in S$, note that by applying *odds supermodularity* recursively, we get $\mathcal{O}(x, S) \geq \sum_{y \in S \setminus \{x\}} \mathcal{O}(x, \{x, y\})$. Next, we argue that this inequality should be satisfied as an equality. By contradiction, suppose there exist $S \in \Omega$ and $x^* \in S$ such that $\mathcal{O}(x^*, S) > \sum_{y \in S \setminus \{x^*\}} \mathcal{O}(x^*, \{x^*, y\})$. It follows from the previous step of the current proof that there exist Luce weights $\{v(x)\}_{x \in X}$ such that for each $x, y \in X$, $\mathcal{O}(x, \{x, y\}) = v(y)/v(x)$. Next, multiply these weights with $1/\Sigma_{x \in S} v(x)$. Let v' denote the new Luce weights. Now, note that for each $x, y \in X$, $\mathcal{O}(x, \{x, y\}) = v'(y)/v'(x)$. Moreover, since $\Sigma_{x \in S} v'(x) = 1$, for each $x \in S$, $\sum_{y \in S \setminus \{x\}} \mathcal{O}(x, \{x, y\}) = \frac{1 - v'(x)}{v'(x)}$. On the other hand, by definition, $\mathcal{O}(x, S) = (1 - \rho_x(S))/\rho_x(S)$. It follows from our initial supposition that for each $x \in S$, $\rho_x(S) \geq v'(x)$ and $\rho_{x^*}(S) > v'(x^*)$. But, this implies $\sum_{x \in S} \rho_x(S) > \sum_{x \in S} v'(x) = 1$. Thus, we get a contradiction. Finally, since for each $S \in \Omega$ and $x \in S$, $\mathcal{O}(x, S) = \sum_{y \in S \setminus \{x\}} \mathcal{O}(x, \{x, y\})$, choices from pairs of alternatives uniquely specify ρ. Since for each pair of alternatives ρ and $\widehat{\rho}$ assign the same choice probabilities, we get $\rho = \widehat{\rho}$. It follows that ρ is a preference-oriented

Luce rule.

$(iii \implies iv)$: Let ρ be a preference-oriented Luce rule represented by $\succsim$ and v. Suppose that $\succsim$ has K indifference classes $I_1 \dots, I_K$ such that for each $k, k' \in \{1 \dots, K\}$, if $k > k'$ then $x \succ y$ for each $x \in I_k$ and $y \in I_{k'}$. For each $n > 0$, define the weight vector v^n such that $v^n(x) = k^n v(x)$ for each $x \in I_k$, and let ρ^n be the Luce rule with the weight vector v^n. Next, we show that ρ^n converges to ρ.

To see this, let x and y be a pair of distinct alternatives. If $x \sim y$, then $x, y \in I_k$ for some $k \in \{1 \dots, K\}$. Then, for each $n > 0$, we have $v^n(x) = k^n v(x)$ and $v^n(y) = k^n v(y)$, which implies that $lim_{n\to\infty} \frac{v^n(y)}{v^n(x)} = \frac{v(y)}{v(x)}$. If $x \succ y$, then $x \in I_k$ and $y \in I_{k'}$ for some $k, k' \in \{1 \dots, K\}$ with $k > k'$. It follows that $v^n(x) = k^n v(x)$ and $v^n(y) = k'^n v(y)$. Therefore, for each $n > 0$, $\frac{v^n(y)}{v^n(x)} = (\frac{k'}{k})^n \frac{v(y)}{v(x)}$. Since $k > k'$, we get $lim_{n\to\infty} \frac{v^n(y)}{v^n(x)} = 0$. Thus, we conclude that ρ^n converges to ρ.

$(iv \implies ii)$: *Step 1*: First we show that if ρ is a Luce rule, then ρ satisfies *odds modularity*. To see this, Let ρ be a Luce rule and $S, T \in \Omega$ be such that $S \cap T = \{x\}$. Since $\rho^(x, S \cup T) > 0$ we have $\mathcal{O}^(x, S \cup T) < \infty$. Now, since ρ is a Luce rule, we have

$$\rho_x(S) = \frac{v(x)}{\sum\limits_{y\in S} v(y)}\ , \ \rho_x(T) = \frac{v(x)}{\sum\limits_{y\in T} v(y)}, \text{ and } \rho_x(S \cup T) = \frac{v(x)}{\sum\limits_{y\in S\cup T} v(y)}$$

It follows that

$$\mathcal{O}^n(x, S) = \frac{\sum\limits_{y\in S\setminus\{x\}} v(y)}{v(x)}\ , \quad \mathcal{O}^n(x, T) = \frac{\sum\limits_{y\in T\setminus\{x\}} v(y)}{v(x)}\ , \text{ and } \quad \mathcal{O}^n(x, S\cup T) = \frac{\sum\limits_{y\in (S\cup T)\setminus\{x\}} v(y)}{v(x)}$$

Since $S \cap T = \{x\}$, we obtain that $\mathcal{O}(x, S) + \mathcal{O}(x, T) = \mathcal{O}(x, S \cup T)$, thus ρ satisfies *odds modularity*.

Step 2: Let $\{\rho^n\}_{n>0}$ be a sequence of Luce rules that converges to an RCF ρ. Then, by Step 1, each ρ^n satisfies *odds modularity*. Now we show that ρ satisfies *odds modularity*. Suppose $\rho_x(S) > 0$ and $\rho_x(T) > 0$. Then, since $t \to \frac{t}{1-t}$ is a continuous mapping at each $t > 0$, we get $lim_{n\to\infty} \mathcal{O}^n(x, S) = \mathcal{O}(x, S)$ and $lim_{n\to\infty} \mathcal{O}^n(x, T) = \mathcal{O}(x, T)$. Thus, we get $\mathcal{O}(x, S) + \mathcal{O}(x, T) = \mathcal{O}(x, S \cup T)$ and ρ satisfies *odds modularity*.

Next, suppose without loss of generality that $\rho_x(S) = 0$. It follows that $lim_{n\to\infty}\ \rho_x^n(S) = 0$. Since each ρ^n is a Luce rule, ρ^n satisfies *regularity*. Therefore, we have $\rho_x^n(S \cup T) \leq \rho_x^n(S)$, which implies that $lim_{n\to\infty}\ \rho_x^n(S \cup T) = 0$. Thus, we get $\rho_x(S \cup T) = 0$, then since the associated odds ratio is ∞, we conclude that ρ satisfies *odds modularity*.

$(ii \implies i)$: This directly follows from the definition of *odds modularity*. ■

In the representation, the preference relation is identified uniquely, and the identified weight function is unique up to multiplication by a positive scalar for each equivalence class of the preference relation. In their empirical model, Hortacsu & Joo (2016) consider a consumer who assigns each product a (Luce) weight in a specific way. In the vein of Helpman et al. (2008), they model the consumer's choice as a two-stage procedure, in which the consumer first eliminates the products with weights lower than a given threshold, and then decides how much to consume of each remaining product according to its relative weight. The Luce weights are used both to determine which products are consumed with positive probability and their market share. For each POLR, we can rescale the weight function and create threshold values for the choice sets so that the given POLR fits into Hortacsu & Joo (2016)'s empirical model.[11]

The main difference between a preference-oriented Luce rule and the classic Luce rule is *positivity*, which requires each alternative be chosen from each choice set with positive probability, i.e. for each $S \in \Omega$ and $x \in S$, $\rho_x(S) > 0$. As a corollary to Theorem 1, we obtain a new characterization of the Luce rule in terms of *positivity* and *odds modularity*.

Corollary 1 *An RCF ρ is a Luce rule if and only if ρ satisfies positivity and odds modularity.*

[11] In Hortacsu & Joo (2016)'s formulation, the threshold value is fixed, yet the weight of an alternative has a choice set dependent component. In mapping a POLR to a threshold model, we keep the weights fixed, but allow the threshold values depend on the choice sets such that as new alternatives are added to a choice set the threshold value does not decrease.

Luce (1959) characterizes his model in terms of *positivity* and IIA, which says that the ratio of choice probabilities between two alternatives does not depend on other alternatives in the choice set, i.e. for each $S \in \Omega$ and $x, y \in S$, $\rho_x(\{x,y\})/\rho_y(\{x,y\}) = \rho_x(S)/\rho_y(S)$. Luce also shows that under *positivity*, IIA is equivalent to the *choice axiom*, which requires that the probability of choosing an alternative x from a choice set S equals the probability of choosing any subset T of S, which contains x, multiplied by the probability of choosing x from T, i.e. for each $S, T \in \Omega$ with $T \subset S$ and $x \in T$, $\rho_x(S) = \Sigma_{y \in T} \rho_y(S) \cdot \rho_x(T)$.

Both *odds modularity* and Luce's axioms are formulated in terms of choice probability ratios. To obtain an additive form of IIA, one needs a logarithmic transformation, whereas *odds modularity* directly renders an additive formulation. The behavioral message carried by *odds modularity* is substantially different from that of IIA and the *choice axiom*. In that, *odds modularity* strengthens *regularity*, whereas both IIA and the *choice axiom* indicate that agents chooses as if taking conditional probabilities with respect to an underlying subjective probability measure.[12]

Note that IIA directly implies *odds modularity*. To see this, suppose that $S = \{x,y,z\}$. IIA requires $\rho(y,\{x,y\})/\rho(x,\{x,y\}) = \rho(y,S)/\rho(x,S)$ and $\rho(z,\{x,z\})/\rho(x,\{x,z\}) = \rho(z,S)/\rho(x,S)$. If we sum up side-by-side, then we obtain that $\mathcal{O}(x,\{x,y\}) + \mathcal{O}(x,\{x,z\}) = \mathcal{O}(x,\{x,y,z\})$. That is, *odds supermodularity* holds as an equality. This connection indicates that *odds modularity* can also be thought as a, weakened, *aggregate form of IIA*.[13] Conversely, to see that IIA follows from *odds modularity*, our Lemma 5 and Lemma 6 are critical. The observation in Lemma 5, which is proved by using Cauchy–Bunyakovsky–Schwarz inequality, might be of mathematical interest for its own sake.

To argue that *odds modularity* can substitute the Luce's *choice axiom*, one should show that both the *choice axiom* and *odds modularity* lead to the same choice behavior in the absence of *positivity*. In a recent study, Cerreia-Vioglio et al. (2017) show that *the choice axiom* also characterizes POLR. It follows that, *odds modularity* can

[12] For the precise connection see Lemma 2 of Luce (1959).

[13] We are grateful to an anonymous referee for suggesting this alternative interpretation.

substitute the *choice axiom* for characterizing extensions of the Luce rule that are widely used in empirical work.

There is a recent literature, which offers extensions of the Luce rule that accommodate zero choice probabilities. Echenique & Saito (2019) and Ahumada & Ülkü (2018) analyze a general Luce model that accommodates zero probability choices.[14] In this model, an agent first forms his consideration sets in any arbitrary way, then applies the Luce rule. Our POLR can be seen as a general Luce rule with a rather structured consideration set. A key difference between our work and these studies is our axiomatic characterization. In that, *odds modularity* plays the main role in our results, whereas *cyclical independence*, which is a multiplicative extension of IIA, plays the main role in theirs.

4 Stochastic path independence and POLR

For a given deterministic choice function c, *path independence* requires for each pair of disjoint choice sets S and T, $c(S \cup T) = c(c(S), c(T))$. First, we introduce our *stochastic path independence* condition. Then, we show that the random choice procedure characterized by this requirement is a specific POLR.[15]

Stochastic Path Independence (SPI): For each $S, T \in \Omega$, such that $S \cap T = \{x\}$,

$$\rho_x(S \cup T) = \rho_x(S) \cdot \rho_x(T).$$

Equivalently, an RCF ρ satisfies SPI if for each $S \in \Omega$, $x \in S$, and $y \notin S$, $\rho_x(S \cup \{y\}) = \rho_x(S) \cdot \rho_x(x, y)$. We show that the random choice procedure characterized by SPI is a specific form of POLR, in which the preference relation that is to be maximized is a near linear order. A weak preference relation $\succsim$ is **a near linear order** if each alternative is indifferent with at most one other alternative.

[14] As a simple exercise along these lines, Doğan & Yıldız (2020) show that a permissive class of RFCs using Luce weights are characterized by using a weakening of *odds supermodularity*.

[15] This section partially subsumes Section 2.3.2 of the PhD Thesis Yildiz (2013).

Theorem 2 *An RCF ρ satisfies stochastic path independence if and only if there exists a near linear order $\succsim$ on X and a weight function $v : X \to R_{++}$ such that for each choice set $S \in \Omega$,*

$$\rho_x(S) = \begin{cases} 1 & \text{if } x = max(S, \succsim), \\ \dfrac{v(x)}{v(x)+v(y)} & \text{if } max(S, \succsim) = \{x, y\}, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. To get a short and simple proof of this result, we use definitions and results from Yildiz (2016). In this paper the list rational random choice (LRRC) procedure is analyzed. This procedure has two primitives: A *list* and *a random binary relation* used to compare pairs of alternatives. The list $\mathbf{f}$ followed by the decision maker is deterministic and gives an ordering of the alternatives. The decision maker's binary comparisons are probabilistic. The probability that an alternative x is chosen from a choice set is equal to the probability that x is chosen from the set consisting of alternatives that precede x in the list, multiplied by the probability that x defeats all the alternatives that follow x in a binary comparison.[16] For each given list $\mathbf{f}$, random binary relation P, choice set S, and alternative $x \in S$, let $P_x^{\mathbf{f}}(S)$ denote the probability that x is chosen from S according to the LRRC procedure specified by $(P, \mathbf{f})$.

For a given RCF p and for each distinct $x, y \in X$, x is **revealed-to-follow** y, denoted by $x \ \mathbf{f}_p \ y$, if SPI is violated between x and y, i.e. there exists $S \in \Omega$ such that $x \in S$, $y \notin S$, and $p_x(S \cup \{y\}) \neq p_x(S) \cdot p_x(x, y)$. In Proposition 1, Yildiz (2016) shows that an RCF p is list rational if and only if $\mathbf{f}_p$ is acyclic. Moreover, the followed list $\mathbf{f}$ is identified unique up to the completions of the transitive closure of $\mathbf{f}_p$. We use these results to derive the following two lemmas that play a key role in proving the only if

[16]For example, suppose that you make a choice between any pair of alternatives x, y and z with equal probabilities. Now, you follow the list $z \ \mathbf{f} \ y \ \mathbf{f} \ x$ (i.e., x comes first, y follows x and comes second, and z comes last) to make a final choice as follows: You first compare x to y, x wins the comparison to y with 0.5 probability, if this is the case then you compare x to z and x wins this comparison with 0.5 probability and gets chosen. Assuming independent randomization, this happens with a total probability of 0.25, so x is chosen with 0.25 probability. Similarly, y is chosen with 0.25 probability. Since z defeats the winner of the previous round with 0.5 probability, whether it is x or y, z is chosen with 0.5 probability.

part of our Theorem 2.

Lemma 7 *Let p be an RCF that satisfies SPI. For each distinct $x, y, z \in X$, if $p_x(x, y) \in (0, 1)$ then $p_x(x, z) \in \{0, 1\}$.*

Proof. Since p satisfies SPI, there is no pair $x, y \in X$ with $x\ \mathbf{f}_p\ y$. It follows from Proposition 1 of Yildiz (2016) that for each list $\mathbf{f}$ and $S \in \Omega$, $P^{\mathbf{f}}(S) = p(S)$. Now, consider the lists $\mathbf{f}$ and $\mathbf{f}'$ such that $x\ \mathbf{f}\ y\ \mathbf{f}\ z$ (i.e., z comes first, y follows z and comes second, and x comes last) and $z\ \mathbf{f}'\ y\ \mathbf{f}'\ x$. By definition of LRRC, $P_x^{\mathbf{f}'}(\{x, y, z\}) = \rho_x(x, y) \cdot \rho_x(x, z)$ and $P_x^{\mathbf{f}}(\{x, y, z\}) \geq min\{\rho_x(x, y), \rho_x(x, z)\}$. Since $P_x^{\mathbf{f}'}(\{x, y, z\}) = P_x^{\mathbf{f}}(\{x, y, z\})$, we get $\rho_x(x, y) \cdot \rho_x(x, z) \geq min\{\rho_x(x, y), \rho_x(x, z)\}$. Therefore, if $\rho_x(x, y) \in (0, 1)$, then this is possible only if $\rho_x(x, z) \in \{0, 1\}$. ■

Lemma 8 *Let p be an RCF that satisfies SPI. For each distinct $x, y, z \in X$ such that $p_x(x, y) \in (0, 1)$, if $\rho_x(x, z) = t$ then $\rho_y(y, z) = t$, where $t \in \{0, 1\}$.*

Proof. If $p_x(x, y) \in (0, 1)$, then it follows from Lemma 7 that $\rho_x(x, z) \in \{0, 1\}$ and $\rho_y(y, z) \in \{0, 1\}$. Let $p_x(x, z) = t$. Consider the lists $\mathbf{f}$ and $\mathbf{f}'$ used in proving Lemma 7. We show that $P_x^{\mathbf{f}}(\{x, y, z\}) = P_x^{\mathbf{f}'}(\{x, y, z\})$ implies $\rho_y(y, z) = t$. Suppose $p_x(x, z) = 1$ but $\rho_y(y, z) = 0$, then we get $P_x^{\mathbf{f}}(\{x, y, z\}) = 1 \neq p_x(x, y) = P_x^{\mathbf{f}'}(\{x, y, z\})$. Suppose $p_x(x, z) = 0$ but $\rho_y(y, z) = 1$, then we get $P_x^{\mathbf{f}}(\{x, y, z\}) = p_x(x, y) \neq 0 = P_x^{\mathbf{f}'}(\{x, y, z\})$. ■

Now we are ready to prove Theorem 2.

Only if part: Let $x, y \in X$ and $S \in \Omega$. We show that $\rho_x(S \cup \{y\}) = \rho_x(x, y) \cdot \rho_x(S)$. (i) Suppose $p_x(S \cup \{y\}) > 0$. Now, for each $z \in S \cup \{y\}$, we have $x \succsim z$. If $x \sim y$ then for each $z \in S \setminus \{x, y\}$, $x \succ z$. It follows that $p_x(S) = 1$ and $p_x(x, y) \cdot \rho_x(S) = p_x(S \cup \{y\})$. If $x \succ y$ then $p_x(x, y) = 1$. Since $\succsim$-maximal elements in S are same as that of $S \setminus \{y\}$, we get $p_x(S) = p_x(S \cup \{y\})$, which implies $\rho_x(S \cup \{y\}) = \rho_x(x, y) \cdot \rho_x(S)$.

(ii) Suppose $p_x(S \cup \{y\}) = 0$. It follows that there exists $z \in S \cup \{y\}$ such that $z \succ x$. If $z = y$ then we get $p_x(x, y) = 0$. If $z \neq y$ then we get $p_x(S) = 0$. Thus, in both cases we get $\rho_x(x, y) \cdot \rho_x(S) = 0$.

If part: Suppose ρ satisfies SPI. We will show that there exists a near linear order $\succsim$ and weight function v that satisfy condition (2). For each $x, y \in X$, define $x \sim y$ if and only if $\rho_x(x, y) \in (0, 1)$, and $x \succ y$ if and only if $p(x, y) = 1$. Define $\succsim = \succ \cup \sim$. It follows from Lemma 7 that if $x \sim y$, then there is no other $z \in X$ such that $x \sim z$. It follows from Lemma 8 that $\succsim$ is transitive. Therefore, $\succsim$ is a near linear order.

Next, we show that for each $S \in \Omega$ and $x \in S$, $x = max(S, \succsim)$ if and only if $p_x(S) = 1$. To see this, first, suppose $x = max(S, \succsim)$. Let $S = \{x, x_1, \ldots, x_k\}$. Since ρ satisfies SPI, $p_x(S) = p_x(x, x_1) \cdot p_x(S \setminus \{x_1\})$. Since $x \succ x_1$, we have $p_x(x, x_1) = 1$. This implies $p_x(S) = p_x(S \setminus \{x_1\})$. By proceeding similarly, we get $p_x(S) = p_x(x, x_k) = 1$. To see the converse, suppose $p_x(S) = 1$. Since, by SPI, for each $y \in S$, $p_x(S) = p_x(x, y) \cdot p_x(S \setminus \{y\})$, it directly follows that $p_x(x, y) = 1$. By definition of $\succsim$, this means $x = max(S, \succsim)$.

Now, we show that for each $S \in \Omega$, $\rho(S) = \rho(max(S, \succsim))$. If there exists $x \in S$ such that $x = max(S, \succsim)$, then the conclusion directly follows from the previous observation. Suppose there exist distinct $x, y \in S$ such that $\{x, y\} = max(S, \succsim)$. By SPI, $p_x(S) = p_x(x, y) \cdot p_x(S \setminus \{y\})$. Since $x = max(S \setminus \{y\}, \succsim)$, again by the previous observation, we get $p_x(S \setminus \{y\}) = 1$. Thus, $p_x(S) = p_x(x, y)$. Symmetric arguments hold with y, so it follows that $\rho(S) = \rho(max(S, \succsim))$. Finally, for each $x, y \in X$ such that $x \sim y$, define $v(x) = p_x(x, y)$ and $v(y) = p_y(x, y)$, for each $z \in X$ such that there is no $y \in X$ with $z \sim y$, define $v(z) = 1$. Thus, $\succsim$ and v yield the desired representation. ■

The direct translation of Plott (1973)'s *path independence* into random choice environment requires for each $S, T \in \Omega$, $p(x, S \cup T) = p(x, \rho^+(S) \cup \rho^+(T))$ where $\rho^+(S)$ and $\rho^+(T)$ are the set of alternatives chosen from S and T with positive probability. Our SPI is a rather demanding counterpart of path independence. In that, to formulate SPI, we depart from *odds modularity* and follow the previous work by Sattath & Tversky (1976) and Kalai & Megiddo (1980). The representation we obtain in Theorem 2 makes how restrictive is this condition visible with naked eye.

Kalai & Megiddo (1980) propose a similar path independence condition and show that any random choice rule which satisfies this condition can choose at most two alternatives from each choice set with positive probability. They restrict the alternative set to have a vector structure and employs this structure to formulate their path independence condition. In our setup, we formulate SPI condition for general alternative sets and provide a full characterization of the underlying procedure. Thus, from our characterization we learn that the alternatives chosen with positive probability can be rationalized as the best alternatives of a preference relation. In another early study, Sattath & Tversky (1976) focus on *multiplicative inequality* that requires the sum of the choice probabilities of an alternative x in any pair of choice sets S and T be less than or equal to the choice probability of x in $S \cup T$.[17] In their analysis, they show that independent random utility model and elimination by aspects of Tversky (1972) satisfy *multiplicative inequality*.

5 Final comments

We obtain a new axiomatic characterization of the Luce rule in terms of *odds supermodularity* that substitutes Luce's IIA and the *choice axiom*. To best of our knowledge this is the first alternative characterization of the Luce rule in its classical setting with finitely many alternatives. There are two recent papers that offer new axiomatizations of the Luce rule in more structured choice environments. Gul et al. (2014) offer a characterization of the Luce rule for the "rich" choice environment. An implication of *richness* is that the alternative set should be infinite. Ahn et al. (2018) characterize the counterpart of the Luce rule in "average choice" environment, in which the alternative set is also infinite and endowed with a vector space structure that allows for the existence of convex combinations and averages. Here, we retain the original setting of Luce (1959), and consider a finite set of alternatives. These different assumptions on the choice environments lead to substantially different characterizations that shed light on different aspects of the Luce rule.

We show that *odds supermodularity* characterizes POLR, which offers a struc-

[17] This is *odds supermodularity* written in terms of choice probabilities instead of odds ratios.

tured extension of the Luce rule that accommodates zero probability choices. Hortacsu & Joo (2016) develop an empirical choice model to accommodate zero choice probabilities. As we discuss in the previous section, by rescaling the Luce weights and choosing the thresholds properly, our POLR can also be represented in this form. Thus, besides its normative appeal, our characterization of POLR paves the way for identification of a model that has already received attention in the empirical literature.

We observe that if a random choice rule satisfies *odd supermodularity*, then it should satisfy *odds modularity*. To clarify this connection, we propose a *stochastic path independence* condition, and show that the random choice procedure characterized by this requirement is a specific POLR, in which each alternative is indifferent with at most one other alternative according to the preference relation that is to be maximized. Our analysis complements the previous work by Sattath & Tversky (1976) and Kalai & Megiddo (1980).

References

Ahn, D. S., Echenique, F. & Saito, K. (2018), 'On path independent stochastic choice', *Theoretical Economics* **13**(1), 61–85. 20

Ahumada, A. & Ülkü, L. (2018), 'Luce rule with limited consideration', *Mathematical Social Sciences* **93**, 52–56. 16

Apesteguia, J. & Ballester, M. A. (2020), 'Separating predicted randomness from residual behavior', *Journal of the European Economic Association* . 3

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., Rustichini, A. et al. (2017), Multinomial logit processes and preference discovery: inside and outside the black box, Technical report. 2, 15

Doğan, S. & Yıldız, K. (2020), 'Category oriented luce rule', *METU Studies in Development* **47**(1), 29–37.

Echenique, F. & Saito, K. (2019), 'General Luce model', *Economic Theory* **68**(4), 811–826. 16

Fudenberg, D. & Strzalecki, T. (2015), 'Dynamic logit with choice aversion', *Econometrica* **83**(2), 651–691. 2

Gandhi, A., Lu, Z. & Shi, X. (2013), 'Estimating demand for differentiated products with error in market shares: A moment inequalities approach', *Manuscript, University of Wisconsin-Madison* . 5

Gul, F., Natenzon, P. & Pesendorfer, W. (2014), 'Random choice as behavioral optimization', *Econometrica* **82**(5), 1873–1912. 20

Helpman, E., Melitz, M. & Rubinstein, Y. (2008), 'Estimating trade flows: Trading partners and trading volumes', *The Quarterly Journal of Economics* **123**(2), 441–487. 4, 14

Horan, S. (2018), 'Threshold Luce rules', *Manuscript, Université de Montréal. Département de sciences économiques* . 4

Hortacsu, A. & Joo, J. (2016), Semiparametric estimation of CES demand system with observed and unobserved product characteristics, *in* '2016 Meeting Papers', number 36, Society for Economic Dynamics. 5, 14, 21

Kalai, E. & Megiddo, A. (1980), 'Path independent choices,', *Econometrica* **48**, 781–784. 19, 21

Luce, R. (1959), *Individual choice behavior: A theoretical analysis*, Wiley, New York. 4, 14, 15, 20

Moulin, H. (1991), *Axioms of cooperative decision making*, number 15, Cambridge University Press. 7

Plott, C. (1973), 'Path independence, rationality, and social choice', *Econometrica* **41**, 1075– 1091. 1, 2, 4, 19

Saito, K. (2017), 'Axiomatizations of the mixed logit model', *mimeo* . 2

Sattath, S. & Tversky, A. (1976), 'Unite and conquer: A multiplicative inequality for choice probabilities', *Econometrica: Journal of the Econometric Society* pp. 79–89. 19, 20, 21

Tversky, A. (1972), 'Elimination by aspects: A theory of choice.', *Psychological review* **79**(4), 281. 20

Yildiz, K. (2013), Essays in Microeconomic Theory, PhD thesis, New York University. 16

Yildiz, K. (2016), 'List-rationalizable choice', *Theoretical Economics* **11**(2), 587–599. 17, 18