Odds Supermodularity and The Luce Rule*

Serhat Dogan† Kemal Yildiz ‡

January 1, 2019

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 also show that *odds supermodularity* per se characterizes a structured extension of the Luce rule that accommodates zero probability choices.

JEL Classification Numbers: D01, D07, D09.

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

*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. An earlier version of this paper is circulated as "A New Characterization of the Luce Rule". 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 *mixed logit* model (Saito (2017)), the *multinomial logit* (Cerreia-Vioglio et al. (2017)), and the *generalized discounted logit* processes (Fudenberg & Strzalecki (2015)).[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.

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 $\mathcal{O}(x, S)$, is the ratio of the probability that x is not chosen from S to the probability that x is chosen from S, that is

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

If x is not chosen from a choice set S with positive probability, then $\mathcal{O}(x, S) = \infty$. 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,

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

[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 $\mathcal{O}(x, \cdot)$ is supermodular for each alternative x, that is for each pair of choice sets S and T that contain x, $\mathcal{O}(x, S) + \mathcal{O}(x, T) \leq \mathcal{O}(x, S \cup T) + \mathcal{O}(x, S \cap T)$. As we also discuss in Section 3, this is equivalent to require for each alternative x, $\mathcal{O}(x, \cdot)$ has *increasing marginal returns*.

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. We show that a random choice function is a Luce rule if and only if it satisfies *odds supermodularity* and *positivity*, which requires each alternative be chosen from each choice set with positive probability.[3] After presenting our results, we compare and contrast *odds supermodularity* to Luce's original axioms.

To argue that *odds supermodularity* can substitute the Luce's *choice axiom*, one should show that both the *choice axiom* and *odds supermodularity* lead to the same choice behavior in the absence of *positivity*. To see this, we characterize random choice functions that satisfy *odds supermodularity*, but possibly fails to satisfy *positivity*. We show that *the preference-oriented Luce rules* (POLRs) are the only random choice functions that satisfy *odds supermodularity*. In a recent study, Cerreia-Vioglio et al. (2017) show that *the choice axiom* also characterizes POLR. It follows that, *odds supermodularity* can substitute the *choice axiom* for characterizing extensions of the Luce rule that are widely used in empirical work

A POLR has two primitives: a *preference relation*[4] $\succsim$ 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. 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

[3]One can easily verify that the Luce rule satisfies the requirement of *odds supermodularity* as an equality. Since we show that the converse is also true, it follows that whenever a random choice function satisfies *odds supermodularity* it satisfies in the form of an equality.

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

[5]In the original formulation by Luce (1959), it is allowed that some alternatives are never chosen

to documented problems due to 'zero probability choices' in estimating market outcomes. POLR offers an extension of the Luce rule that allows zero probability choices, yet retain the simplicity of the Luce rule.

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 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. In what follows, we first introduce and characterize the POLR. As a corollary of this result, we obtain the new characterization of the Luce rule. We conclude by relating our work to the existing literature.

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 prob-

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.

[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.

ability 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 $v : X \to \mathbb{R}_{++}$ such that for each choice set $S \in \Omega$,

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

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 alternative'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 $x_i = 1$. 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 are 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.

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 *regularity* axiom. In the proof, we show that the Luce rule satisfies the requirement of *odds supermodularity* as an equality. Since we show that the converse is also true, it follows that whenever an RCF satisfies *odds supermodularity*, it satisfies as an equality. Then, 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]

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 our characterization of preference-oriented Luce rule and its proof. Our new characterization of the Luce rule follows as a direct corollary.

[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 *An RCF ρ is a preference-oriented Luce rule if and only if ρ satisfies odds supermodularity.*

Proof. 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 the sufficiency part of the proof, we show that the $\succsim_\rho$ relation is the relation 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 $x, y, z \in \mathbb{R}_{++}$, $\frac{x}{1+x+xy} + \frac{y}{1+y+yz} + \frac{z}{1+z+zx} \leq 1$. Moreover, equality holds if and only if $xyz = 1$.*

Proof. If we substitute $\frac{1}{1+x+xy} + \frac{x}{1+x+xy} + \frac{xy}{1+x+xy}$ instead of 1, then we get the following inequality

$$\frac{x}{1+x+xy} + \frac{y}{1+y+yz} + \frac{z}{1+z+zx} \leq \frac{1}{1+x+xy} + \frac{x}{1+x+xy} + \frac{xy}{1+x+xy}.$$

By rearranging the terms we get

$$\left[\frac{y}{1+y+yz} - \frac{xy}{1+x+xy}\right] + \left[\frac{z}{1+z+zx} - \frac{1}{1+x+xy}\right] \leq 0.$$

By equating denominators we get

$$\frac{y + xy + xy^2 - xy - xy^2 - xy^2 z}{(1+y+yz)(1+x+xy)} + \frac{z + zx + xyz - 1 - z - zx}{(1+z+zx)(1+x+xy)} \leq 0.$$

Then,

$$\frac{y(1-xyz)}{(1+y+yz)(1+x+xy)} - \frac{1-xyz}{(1+z+zx)(1+x+xy)} \leq 0.$$

Finally we obtain

$$\frac{1-xyz}{1+x+xy}\left(\frac{xyz-1}{(1+y+yz)(1+z+zx)}\right)=\frac{-(1-xyz)^2}{(1+x+xy)(1+y+yz)(1+z+zx)}\leq 0.$$

It is clear that this inequality holds, and holds as an equality if and only if $xyz=1$. ■

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 $\frac{\mathcal{O}(x,\{x,z\})}{\mathcal{O}(x,\{x,y\})}=\frac{\rho_z(\{y,z\})}{\rho_y(\{y,z\})}$.*

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, consider the sum of the three rightmost terms. It follows from 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 Lemma 5, we know that this equality holds if and only if $\mathcal{O}(x,\{x,y\})\cdot\mathcal{O}(y,\{y,z\})\cdot\mathcal{O}(z,\{x,z\})=1$. 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.

Only if part: Let ρ be a preference-oriented Luce rule represented by $\succsim$ and v. First, suppose $\mathcal{O}(x,S\cup T)=\infty$ for some $S,T\in\Omega$ such that $S\cap T=\{x\}$. It follows that $\rho_x(S\cup T)=0$. This implies there exists $y\in S\cup T$ such that $y\succ x$. It follows

that either $y \in S$ or $y \in T$. Therefore, either $\mathcal{O}(x, S) = \infty$ or $\mathcal{O}(x, T) = \infty$. Thus, we obtain $\mathcal{O}(x, S) + \mathcal{O}(x, T) = \mathcal{O}(x, S \cup T)$. Next, suppose $\mathcal{O}(x, S \cup T) < \infty$. It follows that $\rho_x(S \cup T) > 0$ and $x \in \max(S \cup T, \succsim)$. Let $S' = \max(S, \succsim) \setminus \{x\}$ and $T' = \max(T, \succsim) \setminus \{x\}$. Since $x \in max(S, \succsim)$ and $x \in max(T, \succsim)$, all the alternatives in S' and T' belong to the same indifference class of $\succsim$. Therefore, $\max(S \cup T, \succsim) = S' \cup T' \cup \{x\}$. Now, since ρ is a preference-oriented Luce rule, we have

$$\rho_x(S) = \frac{v(x)}{v(x) + \sum_{y \in S'} v(y)}, \ \rho_x(T) = \frac{v(x)}{v(x) + \sum_{y \in T'} v(y)}, \text{ and } \rho_x(S \cup T) = \frac{v(x)}{v(x) + \sum_{y \in S' \cup T'} v(y)}$$

It follows that

$$\mathcal{O}(x, S) = \frac{\sum_{y \in S'} v(y)}{v(x)}, \quad \mathcal{O}(x, T) = \frac{\sum_{y \in T'} v(y)}{v(x)}, \text{ and } \quad \mathcal{O}(x, S \cup T) = \frac{\sum_{y \in S' \cup T'} v(y)}{v(x)}$$

Since $S \cap T = \{x\}$, we have $S' \cap T' = \emptyset$. Thus, we obtain that $\mathcal{O}(x, S) + \mathcal{O}(x, T) = \mathcal{O}(x, S \cup T)$.

If part: 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. ■

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.[9]

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 supermodularity*.

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

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 supermodularity* 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 supermodularity* directly renders an additive formulation. The behavioral message carried by *odds supermodularity* is substantially different from that of IIA and the *choice axiom*. In that, *odds supermodularity* strengthens *regularity*, whereas both IIA and the *choice axiom* indicate that agents chooses as if taking conditional probabilities with respect an underlying subjective probability measure.[10]

[9] 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.

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

Note that IIA directly implies *odds supermodularity*. 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 supermodularity* can also be thought as a, weakened, *aggregate form of IIA*.[11] To see that IIA follows from *odds supermodularity*, our Lemma 5 and Lemma 6 are critical. The observation in Lemma 5 might be of mathematical interest for its own sake, and can be generalized to any collection of positive real numbers by using Cauchy–Bunyakovsky–Schwarz inequality.[12]

4 Relation to the literature and final comments

We obtain a new axiomatic characterization of the Luce rule in terms of *positivity* and *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*, without *positivity*, characterizes POLR,

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

[12] Along these lines, Selim Bahadir proved an extension of our result in Lemma 5 for more than three positive numbers.

which offers a structured 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.

There is a recent literature, which offers extensions of the Luce rule that accommodate zero choice probabilities. Cerreia-Vioglio et al. (2017) show that *the choice axiom* of Luce also characterizes POLR. We compare and contrast *odds supermodularity* and the *choice axiom* in the previous section. Among others, Yildiz (2013) analyzes a specific POLR, in which the preference to be maximized is a weak order such that an alternative can be indifferent with at most one other alternative. He shows that this model is characterized by a random counterpart of Plott (1973)'s *path independence*. POLR generalizes this model by allowing an agent to be indifferent among any number of alternatives. Echenique & Saito (2015) and Ahumada & Ülkü (2018) analyze a general Luce model that accommodates zero probability choices. 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 supermodularity* plays the main role in our results, whereas *cyclical independence*, which is a multiplicative extension of IIA, plays the main role in theirs.

References

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

Ahumada, A. & Ülkü, L. (2018), 'Luce rule with limited consideration', *Mathematical Social Sciences* **93**, 52–56. 14

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, 3, 14

Echenique, F. & Saito, K. (2015), 'General Luce model', *mimeo* . 14

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* . 4

Gul, F., Natenzon, P. & Pesendorfer, W. (2014), 'Random choice as behavioral optimization', *Econometrica* **82**(5), 1873–1912. 13

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, 11

Horan, S. (2018), 'Extended Luce rules', *mimeo* . 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. 4, 11, 12, 14

Luce, R. (1959), *Individual choice behavior: a theoretical analysis*, Wiley, New York. 3, 12, 13

Moulin, H. (1991), *Axioms of cooperative decision making*, number 15, Cambridge university press. 6

Plott, C. (1973), 'Path independence, rationality, and social choice', *Econometrica* **41**, 1075– 1091. 14

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

Yildiz, K. (2013), Essays in Microeconomic Theory, PhD thesis, New York University. 14