A Canon of Probabilistic Rationality*

Simone Cerreia-Vioglio[a], Per Olov Lindberg[b], Fabio Maccheroni[a]

Massimo Marinacci[a] and Aldo Rustichini[c]

[a]Università Bocconi and Igier, [b]The Royal Institute of Technology

[c]University of Minnesota

March 1, 2021

Abstract

We prove that a random choice rule satisfies Luce's Choice Axiom if and only if its support is a choice correspondence that satisfies the Weak Axiom of Revealed Preference, thus it consists of alternatives that are optimal according to some preference, and random choice then occurs according to a tie breaking among such alternatives that satisfies Renyi's Conditioning Axiom.

Our result shows that the Choice Axiom is, in a precise formal sense, a probabilistic version of the Weak Axiom. It thus supports Luce's view of his own axiom as a "canon of probabilistic rationality."

*This paper combines and supersedes two independent works: Paper 6 (Lindberg, 2012b) of the thesis of Lindberg (2012a) and a working paper of Cerreia-Vioglio Maccheroni, Marinacci, and Rustichini (2016). We thank Sean Horan, Marco Pavan (the editor), an anonymous associate editor, and two anonymous referees for very helpful comments, as well as the ERC (grants SDDM-TEA and INDIMACRO) and a PRIN grant (2017CY2NCA) for financial support. Aldo Rustichini thanks the US Army for financial support, contract W911NF2010242.

1 Introduction

In 1977, twenty years after proposing it, Duncan Luce commented as follows about his celebrated Choice Axiom:[1]

> "Perhaps the greatest strength of the choice axiom, and one reason it continues to be used, is as a *canon of probabilistic rationality.* It is a natural probabilistic formulation of K. J. Arrow's famed principle of the *independence of irrelevant alternatives*, and as such it is a possible underpinning for rational, probabilistic theories of social behavior."

This claim already appears in his 1957 and 1959 works that popularized the axiom and the resulting stochastic choice model.[2] The conceptual proximity of Arrow's principle, typically identified with the set-theoretic version of the Weak Axiom of Revealed Preference (WARP),[3] and Luce's Choice Axiom is indeed often invoked. As well-known, the former plays a key role in deterministic choice theory, the latter in stochastic choice theory.

Yet, the formal relation between these two *independence of irrelevant alternatives (IIA)* assumptions has remained elusive so far.[4] For instance, in analyzing several different IIA axioms Ray (1973) writes:

> "Obviously IIA (Luce) falls in a different category altogether [relative to IIA (Arrow)], being concerned with probabilistic choices."

This note provides the missing link by showing that a random choice rule satisfies Luce's Choice Axiom if and only if:

1. its support, the set of *alternatives that can be chosen*, is a rational choice correspondence à la Arrow (1948, 1959), so it consists of alternatives that are optimal according to some preference;

[1]Luce (1977, p. 229), emphasis added.
[2]See Luce (1957, p. 6) and Luce (1959, p. 9).
[3]Arrow himself put forth this version of Samuelson's WARP in his 1948 and 1959 works.
[4]See the discussion of Peters and Wakker (1991, p. 1789) and Wakker (2010, p. 373).

2. tie-breaking among the optimal alternatives is consistent in the sense of conditional probability à la Renyi (1955, 1956).

In this way, our analysis formally supports Luce's "canonical rationality" claim for his Choice Axiom via a lexicographic composition of deterministic rationality (WARP) and stochastic consistency (Renyi's Conditioning Axiom).

2 Preliminaries

2.1 Random choice rules

Let $\mathcal{A}$ be the collection of all non-empty finite subsets of a universal set X of possible alternatives. The elements of $\mathcal{A}$ are called *choice sets* and denoted by A, B and C.

A map $\Gamma : \mathcal{A} \to \mathcal{A}$ such that $\Gamma(A) \subseteq A$ for all choice sets A is called *choice correspondence*. It is *rational* when

$$B \subseteq A \text{ and } \Gamma(A) \cap B \neq \varnothing \Longrightarrow \Gamma(B) = \Gamma(A) \cap B \tag{WARP}$$

This is the set-theoretic form of WARP considered by Arrow (1948, 1959). Its IIA nature is best seen when Γ is a function:

$$B \subseteq A \text{ and } \Gamma(A) \in B \Longrightarrow \Gamma(B) = \Gamma(A)$$

In words, adding suboptimal alternatives is irrelevant for choice behavior.

We denote by $\Delta(X)$ the set of all finitely supported probability measures on X and, for each $A \subseteq X$, by $\Delta(A)$ the subset of $\Delta(X)$ consisting of the measures assigning mass 1 to A.

Definition 1 *A* random choice rule *is a function*

$$\begin{aligned} p: \ \mathcal{A} &\to \Delta(X) \\ A &\mapsto p_A \end{aligned}$$

such that $p_A \in \Delta(A)$ for all $A \in \mathcal{A}$.

Given any alternative $a \in A$, we interpret $p_A(\{a\})$, also denoted by $p(a, A)$, as the probability that an agent chooses a when the set of available alternatives is A.

More generally, if B is a subset of A, we denote by $p_A(B)$ or $p(B, A)$ the probability that the selected element lies in B.[5] This probability can be viewed as the frequency with which an element in B is chosen. In particular, the set of *alternatives that can be chosen* from A is the *support* of p_A, given by

$$\operatorname{supp} p_A = \{a \in X : p(a, A) > 0\}$$

The condition $p_A(A) = 1$ guarantees that it is a non-empty subset of A, so that the *support correspondence*

$$\begin{aligned} \operatorname{supp} p : \ \mathcal{A} &\to \mathcal{A} \\ A &\mapsto \operatorname{supp} p_A \end{aligned}$$

is a choice correspondence.

Finally, the standard way of comparing the probabilities of choices in two different sets B and C are the *odds* in favor of B over C, that is,

$$r_A(B, C) = \frac{p_A(B)}{p_A(C)} = \frac{\#\text{ of times an element in } B \text{ is chosen}}{\#\text{ of times an element in } C \text{ is chosen}}$$

for all $B, C \subseteq A$. As usual, given any b and c in X, we set $p(b, c) = p(b, \{b, c\})$ and

$$r(b, c) = \frac{p(b, c)}{p(c, b)}$$

2.2 Luce's model

The classical assumptions of Luce (1959) on p are:

Positivity $p(a, b) > 0$ *for all* $a, b \in X$.

Choice Axiom $p(a, A) = p(a, B)\, p(B, A)$ *for all* $B \subseteq A$ *in* $\mathcal{A}$ *and all* $a \in B$.

The latter axiom says that the probability of choosing an alternative a from the choice set A is the probability of first selecting B from A, then choosing a from B (provided a belongs to B). As observed by Luce, formally this assumption corresponds to the fact that $\{p_A : A \in \mathcal{A}\}$ is a conditional probability system in the

[5] Formally, $x \mapsto p(x, A)$ for all $x \in X$ is the discrete density of p_A, but with an abuse of notation $p_A(\cdot)$ is identified with $p(\cdot, A)$; we also write $p_A(a)$ instead of $p_A(\{a\})$.

sense of Renyi (1955, 1956).[6] Remarkably, Luce's Choice Axiom is also equivalent to:

Odds Independence

$$\frac{p(a,b)}{p(b,a)} = \frac{p(a,A)}{p(b,A)} \tag{OI}$$

for all $A \in \mathcal{A}$ and all $a, b \in A$ such that $p(a,A)/p(b,A)$ is well defined.[7]

This axiom says that the odds for a against b are independent of the other available alternatives.[8]

Theorem 1 (Luce) *A random choice rule $p : \mathcal{A} \to \Delta(X)$ satisfies Positivity and the Choice Axiom if and only if there exists $\alpha : X \to \mathbb{R}$ such that*

$$p(a,A) = \frac{e^{\alpha(a)}}{\sum_{b \in A} e^{\alpha(b)}} \tag{LM}$$

for all $A \in \mathcal{A}$ and all $a \in A$.

This fundamental result in random choice theory also shows that, under the Choice Axiom, Positivity is equivalent to the stronger assumption that p_A has full support for all choice sets A.

Full Support $\operatorname{supp} p_A = A$ *for all $A \in \mathcal{A}$.*

From a choice-theoretic perspective, this axiom is unduly restrictive and may permit the choice of "dominated" actions. This note shows what happens when removing from the Luce analysis this extra baggage.

Finally, when X is a separable metric space we may introduce a continuity axiom.

---

[6] See Lemma 2 of Luce (1959) and Lemma 6 in the appendix. For bibliographic accuracy, we remark that here we consider the Choice Axiom in the form stated by Luce (1957) as Axiom 1. Under Positivity, this version coincides with Axiom 1 of the 1959 book, and is the version later analyzed by Luce himself in the retrospective of 1977.

[7] That is, different from 0/0. See Lemma 3 of Luce (1959) when Positivity holds and Lemma 6 in the appendix for the general case.

[8] For this reason, also this axiom often goes under the IIA name. To avoid confusion, we use a less popular label.

Continuity *Given any* $x, y \in X$, *if* $\{x_n\}_{n\in\mathbb{N}}$ *converges to* x, *then*

$$p(x_n, y) > 0 \text{ for all } n \in \mathbb{N} \implies p(x, y) > 0$$
$$p(y, x_n) > 0 \text{ for all } n \in \mathbb{N} \implies p(y, x) > 0$$

This axiom has a natural interpretation: if, eventually, x_n may be always chosen (rejected) over y, and x_n converges to x, then x can be chosen (rejected) over y. Continuity is automatically satisfied under Full Support as well as when X is countable and endowed with the discrete metric.

3 Main result

The next result generalizes Luce's Theorem 1 by getting rid of the Full Support assumption.

Theorem 2 *The following conditions are equivalent for a random choice rule* $p : \mathcal{A} \to \Delta(X)$*:*

(i) p *satisfies the Choice Axiom;*

(ii) there exist a function $\alpha : X \to \mathbb{R}$ *and a rational choice correspondence* $\Gamma : \mathcal{A} \to \mathcal{A}$ *such that*

$$p(a, A) = \begin{cases} \dfrac{e^{\alpha(a)}}{\sum_{b\in\Gamma(A)} e^{\alpha(b)}} & \text{if } a \in \Gamma(A) \\ 0 & \text{else} \end{cases} \tag{CA}$$

for all $A \in \mathcal{A}$ *and all* $a \in A$.

In this case, Γ *is unique and given by* $\Gamma(A) = \operatorname{supp} p_A$ *for all* $A \in \mathcal{A}$.

Since Γ is a rational choice correspondence, the relation $\succ$ defined by

$$a \succ b \iff a \neq b \text{ and } \Gamma(\{a, b\}) = \{a\} \iff b \notin \Gamma(\{a, b\})$$

is a strict preference (see Kreps, 1988) and the corresponding weak preference

$$b \succsim a \iff a \not\succ b \iff b \in \Gamma(\{a, b\}) \iff p(b, a) > 0$$

is such that $\Gamma(A) = \{a \in A : a \succsim b \text{ for all } b \in A\}$.

When X is countable, $\succsim$ is automatically represented by a utility function u and so we have

$$\Gamma(A) = \arg\max_{a \in A} u(a)$$

In general, some additional conditions are needed, as next we show.

Proposition 3 *If X is a separable metric space, then the random choice rule p in Theorem 2 satisfies Continuity if and only if there exists a continuous $u : X \to \mathbb{R}$ such that $\Gamma(A) = \arg\max_{a \in A} u(a)$ for all $A \in \mathcal{A}$.*

A two-stage decision process appears in formula (CA): first rational selection from the choice set A via maximization of preference $\succsim$ (or utility u), then Lucean tie-breaking to choose among the optimal alternatives.

While the optimization structure of the first stage is clear, more can be said about the tie-breaking structure of the second stage in that Theorem 2 describes only its functional form. To this end, recall that a random choice rule p is based on a *Random Preference Model* if there is a (measurable) collection

$$\{\succ_\omega\}_{\omega \in (\Omega, \mathcal{F}, \Pr)}$$

of strict preferences such that, for all $a \in A \in \mathcal{A}$,

$$p(a, A) = \Pr(\omega \in \Omega : a \succ_\omega b \quad \forall b \in A \setminus \{a\})$$

In particular, a Random Preference Model is *Lucean* if $p(\cdot, A)$ has the Luce form (LM).

A piece of terminology: the *lexicographic composition* of two binary relations $\succ$ and $\succ'$ is the binary relation $\succ \circ \succ'$ defined by

$$a \succ \circ \succ' b \iff a \succ b \text{ or } a \sim b \text{ and } a \succ' b$$

For instance, $>_1 \circ >_2$ is the usual lexicographic preference on the Cartesian plane.[9]

We can now state the announced characterization.

[9] Here $>_i$ is defined by $(a_1, a_2) >_i (b_1, b_2) \iff a_i > b_i$.

Proposition 4 *The following conditions are equivalent for a random choice rule* $p : \mathcal{A} \to \Delta(X)$*:*

(i) p *satisfies the Choice Axiom;*

(ii) $\operatorname{supp} p : \mathcal{A} \to \mathcal{A}$ *is a rational choice correspondence and*

$$p_B(a) = \frac{p_A(a)}{p_A(B)} \tag{COND}$$

for all $B \subseteq A \in \mathcal{A}$ *and all* $a \in B \cap \operatorname{supp} p_A$.

(iii) there exist a strict preference $\succ$ *on* X *and a Lucean Random Preference Model* $\{\succ_\omega\}_{\omega \in (\Omega, \mathcal{F}, \Pr)}$ *such that* p *is based on the lexicographic Random Preference Model*

$$\{\succ \circ \succ_\omega\}_{\omega \in (\Omega, \mathcal{F}, \Pr)}$$

This result presents two "deconstructions" of the Choice Axiom that both shed light on the second tie-breaking stage in (CA).

Specifically, to interpret (ii) observe that WARP says that, if a can be chosen from A (i.e., $p_A(a) > 0$) and belongs to $B \subseteq A$, then it can be chosen also from B. But, this axiom is silent about the relation between the frequencies of choice in the two sets A and B. Formula (COND) requires them to be related by the Conditioning Axiom of Renyi (1955, 1956), a classical probabilistic consistency condition. In particular, (COND) per se is weaker than Luce's Choice Axiom, which imposes $p_A(a) = p_B(a)\, p_A(B)$ for all $a \in B \subseteq A$, not just for the elements a in B that can be chosen from A.

To interpret (iii), note that the first-stage preference $\succ$ determines the support of p, while the second stage Random Preference Model $\{\succ_\omega\}_{\omega \in (\Omega, \mathcal{F}, \Pr)}$ is the formal description of the Lucean tie-breaking among optimizers that we previously discussed.

Finally, (iii) also says that, when X is countable, random choice rules that satisfy the Choice Axiom are random utility models (RUM), something not obvious from the definition.[10] This opens the way to the study of general compositions of strict

[10] See Section A.1 below for an independent RUM representation.

preferences and random utility models. The object of current research, a such study goes beyond the scope of this note.

4 Remarks

1. By considering a random choice rule p_A to describe the frequency with which elements are chosen from A, we make the standard interpretation of the choice correspondence $\Gamma(A) = \operatorname{supp} p_A$ as the the set of *alternatives that can be chosen from* A (cf. Sen, 1993) operational and formally meaningful. Here, "can be chosen" means chosen with positive frequency.

2. The second stage of randomization, disciplined by α, can be interpreted in the spirit of Salant and Rubinstein (2008) as capturing observable information which is irrelevant in the rational assessment of the alternatives, but nonetheless affects choice and may reveal how previous experiences and mental associations affect the selection from the optimal $\Gamma(A)$.

3. The distinct roles of u and α become clear once our result is related to the random utility representation of the Luce model. In fact, u corresponds to the *systematic* component of the agent *utility*, and α to the *alternative-specific bias* in the Multinomial Logit Model.[11] Specifically, Theorem 2 shows that a random choice rule p has the form (CA) if and only if, given any $A \in \mathcal{A}$ and any $a \in A$,

$$p_A(a) = \lim_{\lambda \to 0} \Pr\left(\omega \in \Omega : u(a) + \lambda\epsilon_a(\omega) > u(b) + \lambda\epsilon_b(\omega) \quad \forall b \in A \setminus \{a\}\right)$$

where u is a utility function that rationalizes Γ, $\{\epsilon_x\}_{x \in X}$ is a collection of independent errors with type I extreme value distribution, specific mean $\alpha(a)$, common variance $\pi^2/6$, and λ is the noise level. In fact,

$$\begin{aligned}
& \lim_{\lambda \to 0} \Pr\left(\omega \in \Omega : u(a) + \lambda\epsilon_a(\omega) > u(b) + \lambda\epsilon_b(\omega) \quad \forall b \in A \setminus \{a\}\right) \\
= \; & \lim_{\lambda \to 0} \Pr\left(\omega : \frac{u(a)}{\lambda} + \alpha(a) + [\epsilon_a(\omega) - \alpha(a)] > \frac{u(b)}{\lambda} + \alpha(b) + [\epsilon_b(\omega) - \alpha(b)] \ \forall b \neq a\right) \\
= \; & \lim_{\lambda \to 0} \frac{e^{\frac{u(a)}{\lambda} + \alpha(a)}}{\sum_{b \in A} e^{\frac{u(b)}{\lambda} + \alpha(b)}} = \frac{e^{\alpha(a)}}{\sum_{b \in \arg\max_A u} e^{\alpha(b)}} \delta_a(\arg\max_A u) = p_A(a)
\end{aligned}$$

[11] See the seminal McFadden (1973) as well as Ben Akiva and Lerman (1985) and Train (2009) for textbook treatments.

Our analysis thus shows that, when noise vanishes, optimal choice is governed by u and tie-breaking among optimal alternatives is stochastically driven by alternative-specific biases captured by α.

4. A similar interpretation arises when adopting the perspective of Matejka and McKay (2015) on the Multinomial Logit Model as the outcome of an optimal information acquisition problem. In this case, u is the true (initially unknown) payoff of alternatives, α captures a prior belief on payoffs held before engaging in experimentation, and λ is the cost of one unit of information.

   Here our analysis shows that, when the cost of information vanishes, optimal alternatives are selected without error, and prior beliefs only govern the tie-breaking among such alternatives.

5 Related literature

The study of the relations between axiomatic decision theory and stochastic choice has been recently an active field of research. Horan (2020) and Ok and Tserenjigmid (2020) are the most recent works that we are aware of. The former also provides an insightful review of the state of the art. The latter expands on the main conceptual topic of this note: the relation between deterministic and probabilistic "rationality."

Horan (2020) axiomatically unifies Luce (1956, 1959) in a random choice model of imperfect discrimination of the form

$$p(a, A) = \begin{cases} \dfrac{e^{\alpha(a)}}{\sum_{b \in \Gamma(A)} e^{\alpha(b)}} & \text{if } a \in \Gamma(A) \\ 0 & \text{else} \end{cases} \tag{GLM}$$

where Γ is a utility correspondence based on α. Specifically, in Horan, Γ describes the degree of imperfection in the discrimination of the α-values of alternatives; on the contrary, in this note α and Γ are independent, with the former tie-breaking the optimizers identified by the latter.

Horan also compares and provides alternative axiomatizations of several "General Luce Models" (the name is of Echenique and Saito, 2019) of the form (GLM), which correspond to different specifications of the properties of Γ: Ahumada and Ulku (2019), Dogan and Yildiz (2019), Echenique and Saito (2019), and McCausland (2009).

In particular, Dogan and Yildiz (2019) and Horan (2020) provide alternative characterizations of (CA): the former based on supermodularity of odds, the latter on the product rule and a transitivity condition of Fishburn (1978). These results —together with our characterizations of (CA) through the Choice Axiom alone, or WARP and conditioning— provide a full perspective on "rational choice" followed by "rational tie-breaking."

Like us, Ok and Tserenjigmid (2020) regard the support of a random choice rule as a deterministic choice correspondence, and they analyze its rationality properties for several different random choice rules. Following Fishburn (1978), they also consider the entire family of deterministic choice correspondences that lie between the support of p and its subset consisting of the alternatives that are chosen with highest frequency (rather than with positive frequency).

A Proofs and related analysis

A.1 Independent RUM representations

At the end of Section 3, we observed how Proposition 4.(iii) shows that, when X is countable, random choice rules that satisfy the Choice Axiom are random utility models. Here we expand on this topic by providing an explicit independent random utility representation for the random choice rule (CA) of Theorem 2, which holds whenever Γ is the "$\arg\max$" of a utility function $u : X \to \mathbb{R}$ with discrete range. Note that, while this requires $u(X)$ to be countable, no assumption is made on the cardinality of X.

Proposition 5 *Let $u, \alpha : X \to \mathbb{R}$ and consider the random choice rule $p : \mathcal{A} \to$*

$\Delta(X)$ *defined by*

$$p(a,A)=\begin{cases}\dfrac{e^{\alpha(a)}}{\sum_{b\in\arg\max_{z\in A}u(z)}e^{\alpha(b)}} & \text{if } a\in\arg\max_{z\in A}u(z)\\ 0 & \text{else}\end{cases}$$

for all $A\in\mathcal{A}$ and all $a\in A$. If $u(X)$ is a discrete subset of $\mathbb{R}$, then there exists a collection $\{U_x\}_{x\in X}$ of independent random variables such that

$$p(a,A)=\Pr(\omega\in\Omega:U_a(\omega)>U_b(\omega)\quad\forall b\in A\setminus\{a\})$$

for all $A\in\mathcal{A}$ and all $a\in A$.

Proof Let $\{V_x\}_{x\in X}$ be a collection of independent random variables such that

$$\frac{e^{\alpha(a)}}{\sum_{b\in A}e^{\alpha(b)}}=\Pr(\omega\in\Omega:V_a(\omega)>V_b(\omega)\quad\forall b\in A\setminus\{a\})$$

for all $A\in\mathcal{A}$ and all $a\in A$, and assume that $-1<V_x(\omega)<1$ for all $x\in X$ and all $\omega\in\Omega$.[12] Since $u(X)$ is discrete, for each $x\in X$ there exists a constant $r_x>0$ which only depends on $u(x)$ such that

$$u(x)>u(y)\implies u(x)-r_x>u(y)+r_y$$

Define $U_x=u(x)+r_xV_x$ and note that $\{U_x\}_{x\in X}$ is a collection of independent random variables too.

Now arbitrarily choose $A\in\mathcal{A}$ and set $B=\arg\max_{z\in A}u(z)$ and $C=A\setminus B$. Two cases have to be considered.

If $a\in B$, then

$$p(a,A)=\frac{e^{\alpha(a)}}{\sum_{b\in B}e^{\alpha(b)}}=\Pr(\omega\in\Omega:V_a(\omega)>V_b(\omega)\quad\forall b\in B\setminus\{a\})$$

since $u(b)=u(a)$ for all $b\in B$, then $r_b=r_a$ for all $b\in B$, thus

$$\begin{aligned}p(a,A)&=\Pr(\omega\in\Omega:u(a)+r_aV_a(\omega)>u(b)+r_bV_b(\omega)\quad\forall b\in B\setminus\{a\})\\&=\Pr(\omega\in\Omega:U_a(\omega)>U_b(\omega)\quad\forall b\in B\setminus\{a\})\end{aligned}$$

[12] This is without loss of generality because one can always take the representation of McFadden (1973) and apply an arctangent transformation.

But, for all $c \in C = A \setminus B$ and all $\omega \in \Omega$,

$$U_c(\omega) = u(c) + r_c V_c(\omega) < u(c) + r_c$$

and $u(a) > u(c)$ implies $u(a) - r_a > u(c) + r_c$, hence

$$U_a(\omega) = u(a) + r_a V_a(\omega) > u(a) - r_a > u(c) + r_c > U_c(\omega)$$

Thus, $U_a(\omega) > U_c(\omega)$ for all $c \in C$ and all $\omega \in \Omega$, so

$$\begin{aligned} p(a, A) &= \Pr\left(\omega \in \Omega : U_a(\omega) > U_b(\omega) \quad \forall b \in B \setminus \{a\} \text{ and } \forall b \in C\right) \\ &= \Pr\left(\omega \in \Omega : U_a(\omega) > U_b(\omega) \quad \forall b \in A \setminus \{a\}\right) \end{aligned}$$

If instead $c \in C$, then taking $a \in B$ as above, $U_a(\omega) > U_c(\omega)$ for all $\omega \in \Omega$, then

$$\begin{aligned} 0 &= \Pr\left(\omega \in \Omega : U_c(\omega) > U_a(\omega)\right) \\ &\geq \Pr\left(\omega \in \Omega : U_c(\omega) > U_b(\omega) \quad \forall b \in A \setminus \{c\}\right) \end{aligned}$$

whence

$$p(c, A) = 0 = \Pr\left(\omega \in \Omega : U_c(\omega) > U_b(\omega) \quad \forall b \in A \setminus \{c\}\right)$$

as wanted. ■

A.2 Proofs

A preference on X can be given in either strict form, $\succ$, or weak form, $\succsim$.

- In the first case, $\succ$ is required to be asymmetric and negatively transitive, and $\succsim$ is defined by

$$a \succsim b \text{ if and only if } \neg(b \succ a) \tag{1}$$

- In the second case, $\succsim$ is required to be complete and transitive, and $\succ$ is defined by

$$b \succ a \text{ if and only if } \neg(a \succsim b) \tag{2}$$

These approaches are well known to be interchangeable,[13] and for this reason we call *weak order* both $\succ$ and $\succsim$ with the understanding that they are related by the equivalent (1) or (2).

[13] See Kreps (1988, p. 11).

Lemma 6 *Let $p : \mathcal{A} \to \Delta(X)$ be a random choice rule. The following conditions are equivalent:*

(i) p is such that, $p_A(C) = p_B(C)\, p_A(B)$ for all $C \subseteq B \subseteq A$ in $\mathcal{A}$;

(ii) p satisfies the Choice Axiom;

(iii) p is such that $p(b, B)\, p(a, A) = p(a, B)\, p(b, A)$ for all $B \subseteq A$ in $\mathcal{A}$ and all $a, b \in B$;

(iv) p satisfies Odds Independence;

(v) p is such that $p(Y \cap B, A) = p(Y, B)\, p(B, A)$ for all $B \subseteq A$ in $\mathcal{A}$ and all $Y \subseteq X$.

Moreover, in this case, p satisfies Positivity if and only if it satisfies Full Support.

Proof *(i) implies (ii).* Choose as C the singleton a appearing in the statement of the axiom.

(ii) implies (iii). Given any $B \subseteq A$ in $\mathcal{A}$ and any $a, b \in B$, by the Choice Axiom, $p(a, A) = p(a, B)\, p(B, A)$, but then $p(b, B)\, p(a, A) = p(a, B)\, p(b, B)\, p(B, A) = p(a, B)\, p(b, A)$ where the second equality follows from another application of the Choice Axiom.

(iii) implies (iv). Let $A \in \mathcal{A}$ and arbitrarily choose $a, b \in A$ such that $p(a, A) / p(b, A) \neq 0/0$. By (iii),

$$p(b, a)\, p(a, A) = p(b, \{a, b\})\, p(a, A) = p(a, \{a, b\})\, p(b, A) = p(a, b)\, p(b, A)$$

three cases have to be considered:

- $p(b, a) \neq 0$ and $p(b, A) \neq 0$, then $p(a, A) / p(b, A) = p(a, b) / p(b, a)$;
- $p(b, a) = 0$, then $p(a, b)\, p(b, A) = 0$, but $p(a, b) \neq 0$ (because $p(a, b) / p(b, a) \neq 0/0$), thus $p(b, A) = 0$ and $p(a, A) \neq 0$ (because $p(a, A) / p(b, A) \neq 0/0$); therefore
$$\frac{p(a, b)}{p(b, a)} = \infty = \frac{p(a, A)}{p(b, A)}$$

- $p(b,A)=0$, then $p(b,a)\,p(a,A)=0$, but $p(a,A)\neq 0$ (because $p(a,A)/p(b,A)\neq 0/0$), thus $p(b,a)=0$ and $p(a,b)\neq 0$ (because $p(a,b)/p(b,a)\neq 0/0$); therefore

$$\frac{p(a,A)}{p(b,A)}=\infty=\frac{p(a,b)}{p(b,a)}$$

(iv) implies (iii). Given any $B\subseteq A$ in $\mathcal{A}$ and any $a,b\in B$:

- If $p(a,A)/p(b,A)\neq 0/0$ and $p(a,B)/p(b,B)\neq 0/0$, then by (OI)

$$\frac{p(a,A)}{p(b,A)}=\frac{p(a,b)}{p(b,a)}=\frac{p(a,B)}{p(b,B)}$$

  - If $p(b,A)\neq 0$, then $p(b,B)\neq 0$ and $p(b,B)\,p(a,A)=p(a,B)\,p(b,A)$.
  - Else $p(b,A)=0$, then $p(b,B)=0$ and again $p(b,B)\,p(a,A)=p(a,B)\,p(b,A)$.

- Else, either $p(a,A)/p(b,A)=0/0$ or $p(a,B)/p(b,B)=0/0$, and in both cases

$$p(b,B)\,p(a,A)=p(a,B)\,p(b,A)$$

(iii) implies (v). Given any $B\subseteq A$ in $\mathcal{A}$ and any $Y\subseteq X$, since $p(B,B)=1$, it follows $p(Y,B)=p(Y\cap B,B)$. Therefore

$$\begin{aligned}
p(Y\cap B,A)&=\sum_{y\in Y\cap B}p(y,A)=\sum_{y\in Y\cap B}\left(\sum_{x\in B}p(x,B)\right)p(y,A)=\sum_{y\in Y\cap B}\left(\sum_{x\in B}p(x,B)\,p(y,A)\right)\\
[\text{by (iii)}]&=\sum_{y\in Y\cap B}\left(\sum_{x\in B}p(y,B)\,p(x,A)\right)=\sum_{y\in Y\cap B}p(y,B)\left(\sum_{x\in B}p(x,A)\right)\\
&=\sum_{y\in Y\cap B}p(y,B)\,p(B,A)=p(Y\cap B,B)\,p(B,A)=p(Y,B)\,p(B,A)
\end{aligned}$$

(v) implies (i). Take $Y=C$.

Finally, let p satisfy the Choice Axiom. Assume – *per contra* – Positivity holds and $p(a,A)=0$ for some $A\in\mathcal{A}$ and some $a\in A$. Then $A\neq\{a\}$ and, for all $b\in A\setminus\{a\}$, the Choice Axiom implies $0=p(a,A)=p(a,\{a,b\})\,p(\{a,b\},A)=p(a,b)\,(p(a,A)+p(b,A))=p(a,b)\,p(b,A)$ whence $p(b,A)=0$ (because $p(a,b)\neq 0$), contradicting $p(A,A)=1$. Therefore Positivity implies Full Support. The converse is trivial. ■

If $p:\mathcal{A}\to\Delta(X)$ is a random choice rule, denote by $\sigma_p(A)$ the support of p_A, for all $A\in\mathcal{A}$.

Lemma 7 *If $p : \mathcal{A} \to \Delta(X)$ is a random choice rule that satisfies the Choice Axiom, then $\sigma_p : \mathcal{A} \to \mathcal{A}$ is a rational choice correspondence.*

Proof Clearly, $\varnothing \neq \sigma_p(A) \subseteq A$ for all $A \in \mathcal{A}$, then $\sigma_p : \mathcal{A} \to \mathcal{A}$ is a choice correspondence. Let $A, B \in \mathcal{A}$ be such that $B \subseteq A$ and assume that $\sigma_p(A) \cap B \neq \varnothing$.

We want to show that $\sigma_p(A) \cap B = \sigma_p(B)$. Since p satisfies the Choice Axiom, if $a \in \sigma_p(A) \cap B$, then $0 < p(a, A) = p(a, B)\, p(B, A)$. It follows that $p(a, B) > 0$, that is, $a \in \sigma_p(B)$. Thus, $\sigma_p(A) \cap B \subseteq \sigma_p(B)$. As to the converse inclusion, let $a \in \sigma_p(B)$, that is, $p(a, B) > 0$. By contradiction, assume that $a \notin \sigma_p(A) \cap B$. Since $a \in B$, it must be the case that $a \notin \sigma_p(A)$, that is, $p(a, A) = 0$. Since p satisfies the Choice Axiom, we then have $0 = p(a, A) = p(a, B)\, p(B, A)$. Since $p(a, B) > 0$, it must be the case that $p(B, A) = 0$, that is, $\sigma_p(A) \cap B = \varnothing$. This contradicts $\sigma_p(A) \cap B \neq \varnothing$; therefore, a belongs to $\sigma_p(A) \cap B$. Thus, $\sigma_p(B) \subseteq \sigma_p(A) \cap B$. ■

Lemma 8 *The following conditions are equivalent for a function $p : \mathcal{A} \to \Delta(X)$:*

(i) p is a random choice rule that satisfies the Choice Axiom;

(ii) p is a random choice rule such that σ_p is a rational choice correspondence, and

$$p_B(a) = \frac{p_H(a)}{p_H(B)} \tag{3}$$

for all $B \subseteq H \in \mathcal{A}$ and all $a \in \sigma_p(H) \cap B$;

(iii) there exist a function $v : X \to (0, \infty)$ and a rational choice correspondence $\Gamma : \mathcal{A} \to \mathcal{A}$ such that, for all $x \in X$ and $A \in \mathcal{A}$

$$p(x, A) = \begin{cases} \dfrac{v(x)}{\sum_{b \in \Gamma(A)} v(b)} & \text{if } x \in \Gamma(A) \\ 0 & \text{else} \end{cases} \tag{4}$$

In this case, Γ is unique and coincides with σ_p.

Proof *(iii) implies (i).* Let p be given by (4) with Γ a rational choice correspondence and $v : X \to (0, \infty)$. It is easy to check that p is a well defined random choice rule, that the support correspondence $\operatorname{supp} p$ coincides with Γ, and that

$$p(Y, A) = \sum_{y \in Y \cap \Gamma(A)} \frac{v(y)}{\sum_{d \in \Gamma(A)} v(d)}$$

for all $Y \subseteq X$ and all $A \in \mathcal{A}$.

Let $A, B \in \mathcal{A}$ be such that $B \subseteq A$ and $a \in B$. We have two cases:

- If $\Gamma(A) \cap B \neq \varnothing$, since Γ satisfies WARP, $\Gamma(A) \cap B = \Gamma(B)$.

  - If $a \in \Gamma(B)$, then $a \in \Gamma(A)$ and $p(a, B) = v(a) / \sum_{b \in \Gamma(B)} v(b)$, it follows that

$$p(a, A) = \frac{v(a)}{\sum_{d \in \Gamma(A)} v(d)} = \frac{v(a)}{\sum_{b \in \Gamma(B)} v(b)} \frac{\sum_{b \in \Gamma(A) \cap B} v(b)}{\sum_{d \in \Gamma(A)} v(d)} = p(a, B) p(B, A)$$

  - Else $a \notin \Gamma(B)$, and since $a \in B$, it must be the case that $a \notin \Gamma(A)$, so $p(a, A) = 0 = p(a, B) = p(a, B) p(B, A)$.

- Else $\Gamma(A) \cap B = \varnothing$. It follows that $a \notin \Gamma(A)$ and $p(B, A) = 0 = p(a, A)$; again, we have $p(a, A) = p(a, B) p(B, A)$.

These cases prove that p satisfies the Choice Axiom.

(i) implies (ii). Let $p : \mathcal{A} \to \Delta(X)$ be a random choice rule that satisfies the Choice Axiom. Then, by Lemma 7, $\sigma_p : \mathcal{A} \to \mathcal{A}$ is a rational choice correspondence. Moreover, if $B \subseteq H$ and all $a \in \sigma_p(H) \cap B$, then

$$p(a, H) = p(a, B) p(B, H)$$

but $p(B, H) \geq p(a, H) > 0$ because $a \in B$ and $a \in \sigma_p(H)$, and (3) follows.

(ii) implies (iii). Let $p : \mathcal{A} \to \Delta(X)$ be a random choice rule such that σ_p is a rational choice correspondence, and that satisfies (3). Since, σ_p is a rational choice correspondence, then the relation

$$a \succsim b \iff a \in \sigma_p(\{a, b\}) \iff p(a, b) > 0$$

is a weak order on X; and its symmetric part $\sim$ is an equivalence relation such that

$$a \sim b \iff p(a, b) > 0 \text{ and } p(b, a) > 0 \iff r(a, b) \in (0, \infty)$$

Moreover, by Theorem 3 of Arrow (1959), it follows that

$$\sigma_p(A) = \{a \in A : a \succsim b \quad \forall b \in A\} \qquad \forall A \in \mathcal{A} \tag{5}$$

in particular, all elements of $\sigma_p(A)$ are equivalent with respect to $\sim$, and

$$\sigma_p(S) = S \tag{6}$$

for all $S \in \mathcal{A}$ consisting of equivalent elements.

Let $\{X_i : i \in I\}$ be the family of all equivalence classes of $\sim$ in X. Choose $a_i \in X_i$ for all $i \in I$. For each $x \in X$, there exists one and only one $i = i_x$ such that $x \in X_i$, set

$$v(x) = r(x, a_i) \tag{7}$$

Since $x \sim a_i$, then $r(x, a_i) \in (0, \infty)$; and so $v : X \to (0, \infty)$ is well defined. Consider any $x \sim y$ in X and any $S \in \mathcal{A}$ consisting of equivalent elements and containing x and y. Notice that, by (6), $\sigma_p(S) = S$, hence $x \in \sigma_p(S) \cap \{x, y\}$, then by (3) with $H = S$ and $B = \{x, y\}$,

$$p(x, y) = \frac{p_S(x)}{p_S(\{x, y\})}$$

therefore

$$0 < p(x, S) = p(x, y)\, p(\{x, y\}, S)$$

and analogously

$$0 < p(y, S) = p(y, x)\, p(\{x, y\}, S)$$

yielding that

$$p(x, y)\, p(y, x)\, p(x, S)\, p(y, S) > 0 \text{ and } \frac{p(x, S)}{p(y, S)} = \frac{p(x, y)}{p(y, x)} = r(x, y) \tag{8}$$

We are ready to conclude our proof, that is, to show that (4) holds with $\Gamma = \sigma_p$. Let $a \in X$ and $A \in \mathcal{A}$. If $a \notin \sigma_p(A)$, then $p(a, A) = 0$ because $\sigma_p(A)$ is the support of p_A. Else, $a \in \sigma_p(A)$, and, by (5), all the elements in $\sigma_p(A)$ are equivalent with respect to $\sim$ and therefore they are equivalent to some a_i with $i \in I$. It follows that $\sigma_p(A) \cup \{a_i\} \in \mathcal{A}$ and it is such that $\sigma_p(A) \cup \{a_i\} \subseteq X_i$. By (6), we have that $\sigma_p(\sigma_p(A) \cup \{a_i\}) = \sigma_p(A) \cup \{a_i\}$, that is, $p(x, \sigma_p(A) \cup \{a_i\}) > 0$ for all $x \in \sigma_p(A) \cup \{a_i\}$ and $p(\sigma_p(A), \sigma_p(A) \cup \{a_i\}) > 0$. By (3) with $H = A$ and $B = \sigma_p(A)$, since $a \in \sigma_p(A) \cap B$, it follows

$$p(a, \sigma_p(A)) = \frac{p(a, A)}{p(\sigma_p(A), A)}$$

Since $p(\sigma_p(A), A) = 1$, then

$$p(a, A) = p(a, \sigma_p(A))$$

By (3) again, with $H = \sigma_p(A) \cup \{a_i\}$ and $B = \sigma_p(A)$, since $a \in \sigma_p(\sigma_p(A) \cup \{a_i\}) \cap \sigma_p(A)$, then

$$p(a, \sigma_p(A)) = \frac{p(a, \sigma_p(A) \cup \{a_i\})}{p(\sigma_p(A), \sigma_p(A) \cup \{a_i\})} = \frac{\frac{p(a,\sigma_p(A)\cup\{a_i\})}{p(a_i,\sigma_p(A)\cup\{a_i\})}}{\frac{p(\sigma_p(A),\sigma_p(A)\cup\{a_i\})}{p(a_i,\sigma_p(A)\cup\{a_i\})}}$$

applying (8) to the pairs $(x, y) = (a, a_i)$ and $(x, y) = (b, a_i)$, with $b \in \sigma_p(A)$, in $S = \sigma_p(A) \cup \{a_i\} \subseteq X_i$, we can conclude that

$$\frac{\frac{p(a,\sigma_p(A)\cup\{a_i\})}{p(a_i,\sigma_p(A)\cup\{a_i\})}}{\frac{p(\sigma_p(A),\sigma_p(A)\cup\{a_i\})}{p(a_i,\sigma_p(A)\cup\{a_i\})}} = \frac{\frac{p(a,\sigma_p(A)\cup\{a_i\})}{p(a_i,\sigma_p(A)\cup\{a_i\})}}{\sum_{b\in\sigma_p(A)} \frac{p(b,\sigma_p(A)\cup\{a_i\})}{p(a_i,\sigma_p(A)\cup\{a_i\})}} = \frac{r(a, a_i)}{\sum_{b\in\sigma_p(A)} r(b, a_i)} = \frac{v(a)}{\sum_{b\in\sigma_p(A)} v(b)}$$

as wanted.

As for the uniqueness part, we already observed that (iii) implies $\Gamma = \sigma_p$. ■

Theorem 2 immediately follows.

Proof of Proposition 3 In Theorem 2, Γ is a rational choice correspondence and the corresponding weak order is

$$a \succsim b \iff a \in \Gamma(\{a, b\}) \iff p(a, b) > 0$$

thus Continuity can be rewritten as: Given any $x, y \in X$, if $\{x_n\}_{n\in\mathbb{N}}$ converges to x, then

$$\begin{aligned} x_n \ \succsim \ y \text{ for all } n \in \mathbb{N} &\implies x \succsim y \\ y \ \succsim \ x_n \text{ for all } n \in \mathbb{N} &\implies y \succsim x \end{aligned}$$

This concludes the proof, because on a separable metric space, a weak order admits a continuous utility if and only if its upper and lower level sets are closed (see, e.g., Kreps, 1988, p. 27). ■

The set $\mathcal{W}$ of all weak orders on X is endowed with the σ-algebra $\mathfrak{W}$ generated by the sets of the form

$$W_{ab} = \{\succ : \ a \succ b\} \qquad \forall a, b \in X$$

Given $\succ$ and $\succ'$ in $\mathcal{W}$, the lexicographic composition $\succ \circ \succ'$ of $\succ$ and $\succ'$ is routinely seen to be a weak order too (see, e.g., Fishburn, 1974).

Lemma 9 *For each $\succ$ in $\mathcal{W}$, the map*

$$\begin{array}{rccc} f = f_{\succ}: & \mathcal{W} & \to & \mathcal{W} \\ & \succ' & \mapsto & \succ \circ \succ' \end{array}$$

is measurable with respect to $\mathfrak{W}$.

Proof Arbitrarily choose $a, b \in X$, and study

$$\begin{aligned} f^{-1}(W_{ab}) &= f^{-1}(\{\succ'' : a \succ'' b\}) = \{\succ' : f(\succ') \in \{\succ'' : a \succ'' b\}\} \\ &= \{\succ' : a\ f(\succ')\ b\} = \{\succ' : a \succ \circ \succ' b\} \end{aligned}$$

- if $a \prec b$, then there is no $\succ'$ in $\mathcal{W}$ such that $a \succ \circ \succ' b$, that is,

  $$\{\succ' : a \succ \circ \succ' b\} = \varnothing$$

  which is measurable (because $\varnothing \in \mathfrak{W}$),

- else if $a \succ b$, then $a \succ \circ \succ' b$ for all $\succ'$ in $\mathcal{W}$, that is,

  $$\{\succ' : a \succ \circ \succ' b\} = \mathcal{W}$$

  which is measurable (because $\mathcal{W} \in \mathfrak{W}$),

- else, it must be the case that $a \sim b$ and $a \succ \circ \succ' b$ if and only if $a \succ' b$, that is,

  $$\{\succ' : a \succ \circ \succ' b\} = \{\succ' : a \succ' b\} = W_{ab}$$

  which is measurable (because $W_{ab} \in \mathfrak{W}$).

Therefore f is measurable since the counterimage of a class of generators of $\mathfrak{W}$ is contained in $\mathfrak{W}$. ■

A *Random Preference Model* is a measurable function

$$\begin{array}{rccc} P: & (\Omega, \mathcal{F}, \Pr) & \to & \mathcal{W} \\ & \omega & \mapsto & P(\omega) \end{array}$$

It is common practice to write $\succ_\omega$ instead of $P(\omega)$. The *Random Selector* p based on the RPM P is given by

$$p(a, A) = \Pr(\omega \in \Omega : a \succ_\omega b \quad \forall b \in A \setminus \{a\}) \qquad \forall a \in A \in \mathcal{A}$$

The latter is well defined because

$$\begin{aligned} \{\omega \in \Omega : a \succ_\omega b \quad \forall b \in A \setminus \{a\}\} &= \{\omega \in \Omega : P(\omega) \in W_{ab} \quad \forall b \in A \setminus \{a\}\} \\ &= \left\{\omega \in \Omega : P(\omega) \in \bigcap\nolimits_{b \in A \setminus \{a\}} W_{ab}\right\} \\ &= P^{-1}\left(\bigcap\nolimits_{b \in A \setminus \{a\}} W_{ab}\right) \in \mathcal{F} \end{aligned}$$

since P is measurable. Moreover, depending on P, the RS p might not define a random choice rule. For instance, if P is constantly equal to the trivial weak order according to which all alternatives are indifferent, then $p(a, A) = 0$ for all $a \in A \in \mathcal{A}$ such that $|A| \geq 2$.

The proof of Proposition 4 hinges on the study of the composition of the functions $f_\succ$ and P.

First, such a composition defines a random preference model, because

$$\begin{aligned} f_\succ \circ P : \quad (\Omega, \mathcal{F}, \Pr) &\to \mathcal{W} \\ \omega &\mapsto f_\succ(P(\omega)) = \ \succ \circ \succ_\omega \end{aligned}$$

—being a composition of measurable functions, it is measurable.

Second, the random selector based on the random preference model $f_\succ \circ P$ is a lexicographic version of P, that first selects the maximizers of $\succ$, then breaks the ties according to P.

In order to state these results formally, we denote by $\Gamma = \Gamma_\succ$ the rational choice correspondence induced by $\succ$.[14]

Lemma 10 *Let $\succ$ be a weak order, $P = \{\succ_\omega\}_{\omega \in \Omega}$ be a RPM, and p be the RS based on P. Then $f_\succ \circ P = \{\succ \circ \succ_\omega\}_{\omega \in \Omega}$ is a RPM and the RS based on it is given by*

$$p_\succ(a, A) = \begin{cases} p(a, \Gamma(A)) & \text{if } a \in \Gamma(A) \\ 0 & \text{else} \end{cases} \tag{9}$$

for all $a \in A \in \mathcal{A}$.

[14] $\Gamma_\succ(A) = \{a \in A : a \succsim b \text{ for all } b \in A\}$ also recall that $a \succsim b$ if and only if $a \not\prec b$.

Proof We already observed that $f_{\succ} \circ P = \{\succ \circ \succ_{\omega}\}_{\omega \in \Omega}$ is a RPM. By definition of random selector based on a RPM

$$p_{\succ}(a, A) = \Pr(\omega \in \Omega : a \succ \circ \succ_{\omega} b \quad \forall b \in A \setminus \{a\})$$

We have to verify that this formula coincides with (9) for all $a \in A \in \mathcal{A}$.

For each $A \in \mathcal{A}$ and each $a \in \Gamma(A)$, set

$$\begin{aligned} J_A(a) &= \{\omega \in \Omega : a \succ_{\omega} c \text{ for all } c \in \Gamma(A) \setminus \{a\}\} = J \\ K_A(a) &= \{\omega \in \Omega : a \succ \circ \succ_{\omega} b \text{ for all } b \in A \setminus \{a\}\} = K \end{aligned}$$

Next we check that $J = K$.

If $\omega \in J$, then $a \succ_{\omega} c$ for all $c \in \Gamma(A) \setminus \{a\}$; take any $b \in A \setminus \{a\}$,

- if b is such that $b \notin \Gamma(A)$, then, $a \succ b$ and hence $a \succ \circ \succ_{\omega} b$,

- else $b \in \Gamma(A)$, then $a \sim b$ and $a \succ_{\omega} b$, again $a \succ \circ \succ_{\omega} b$,

then $a \succ \circ \succ_{\omega} b$ for all $b \in A \setminus \{a\}$, thus $\omega \in K$.

Conversely, if $\omega \in K$, then $a \succ \circ \succ_{\omega} b$ for all $b \in A \setminus \{a\}$. Thus, for all $b \in \Gamma(A) \setminus \{a\}$, since relation $a \sim b$, it must be the case that $a \succ_{\omega} b$. Therefore ω is such that $a \succ_{\omega} b$ for all $b \in \Gamma(A) \setminus \{a\}$, and $\omega \in J$.

Summing up, for all $A \in \mathcal{A}$ and $a \in \Gamma(A)$,

$$p(a, \Gamma(A)) = \Pr J_A(a) = \Pr K_A(a) = p_{\succ}(a, A)$$

and the first line of (9) is true.

Let $A \in \mathcal{A}$ and $a \notin \Gamma(A)$, then there exists $\bar{b} \in A \setminus \{a\}$ such that $a \prec \bar{b}$, and for no ω it holds $a \succ \circ \succ_{\omega} \bar{b}$, that is,

$$K_A(a) = \{\omega \in \Omega : a \succ \circ \succ_{\omega} b \text{ for all } b \in A \setminus \{a\}\} = \varnothing$$

therefore $p_{\succ}(a, A) = \Pr K_A(a) = 0$, and the second line of (9) is true too. ■

Proof of Proposition 4 The equivalence between points (i) and (ii) corresponds with the equivalence between the points with the same name of Lemma 8.

(i) implies (iii). By Theorem 2, there exist a function $\alpha : X \to \mathbb{R}$ and a rational choice correspondence $\Gamma : \mathcal{A} \to \mathcal{A}$ such that

$$p(a, A) = \begin{cases} \dfrac{e^{\alpha(a)}}{\sum_{b \in \Gamma(A)} e^{\alpha(b)}} & \text{if } a \in \Gamma(A) \\ 0 & \text{else} \end{cases}$$

for all $a \in A \in \mathcal{A}$. Denote by $\succ$ the weak order that corresponds to Γ.

As shown by McFadden (1973), the Lucean random choice rule

$$q(a, A) = \frac{e^{\alpha(a)}}{\sum_{b \in A} e^{\alpha(b)}} \qquad \forall a \in A \in \mathcal{A}$$

is based on a (Lucean) RPM $P = \{\succ_\omega\}_{\omega \in \Omega}$. By Lemma 10, it follows that $f_\succ \circ P = \{\succ \circ \succ_\omega\}_{\omega \in \Omega}$ is a RPM and the RS based on it is given by

$$q_\succ(a, A) = \begin{cases} q(a, \Gamma(A)) & \text{if } a \in \Gamma(A) \\ 0 & \text{else} \end{cases} = p(a, A) \qquad \forall a \in A \in \mathcal{A}$$

Therefore, there exist a Lucean Random Preference Model $\{\succ_\omega\}_{\omega \in (\Omega, \mathcal{F}, \Pr)}$ and a weak order $\succ$ on X such that p is based on $\{\succ \circ \succ_\omega\}_{\omega \in (\Omega, \mathcal{F}, \Pr)}$.

(iii) implies (i). If there exist a Lucean Random Preference Model $\{\succ_\omega\}_{\omega \in (\Omega, \mathcal{F}, \Pr)}$ and a weak order $\succ$ on X such that p is based on $\{\succ \circ \succ_\omega\}_{\omega \in (\Omega, \mathcal{F}, \Pr)}$; in particular, there exists $\alpha : X \to \mathbb{R}$ such that

$$\Pr(\omega \in \Omega : a \succ_\omega b \quad \forall b \in A \setminus \{a\}) = \frac{e^{\alpha(a)}}{\sum_{b \in A} e^{\alpha(b)}} \qquad \forall a \in A \in \mathcal{A}$$

Denoting

$$q(a, A) = \frac{e^{\alpha(a)}}{\sum_{b \in A} e^{\alpha(b)}} \qquad \forall a \in A \in \mathcal{A}$$

the RS based on $\{\succ_\omega\}_{\omega \in (\Omega, \mathcal{F}, \Pr)}$, by Lemma 10, the RS is based on $\{\succ \circ \succ_\omega\}_{\omega \in (\Omega, \mathcal{F}, \Pr)}$ is

$$\begin{aligned} q_\succ(a, A) &= \begin{cases} q(a, \Gamma(A)) & \text{if } a \in \Gamma(A) \\ 0 & \text{else} \end{cases} \\ &= \begin{cases} \dfrac{e^{\alpha(a)}}{\sum_{b \in \Gamma(A)} e^{\alpha(b)}} & \text{if } a \in \Gamma(A) \\ 0 & \text{else} \end{cases} \end{aligned}$$

But, by assumption (iii), $q_{\succ}$ coincides with p (p is based on $\{\succ \circ \succ_\omega\}_{\omega\in(\Omega,\mathcal{F},\Pr)}$), and Γ is a rational choice correspondence because $\succ$ is a weak order. Then Theorem 2 guarantees that p satisfies the Choice Axiom. ■

References

[1] Ahumada, A., and Ulku, L. (2018). Luce rule with limited consideration. *Mathematical Social Sciences*, 93, 52-56.

[2] Arrow, K. J. (1948). The possibility of a universal social welfare function. RAND Document P-41, 1948.

[3] Arrow, K. J. (1959). Rational choice functions and orderings, *Economica*, 26, 121-127, 1959.

[4] Ben-Akiva, M. E., and Lerman, S. R. (1985). *Discrete choice analysis: theory and application to travel demand*. MIT Press.

[5] Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., and Rustichini, A. (2016). Law of demand and forced choice. IGIER Working Paper 593. http://www.igier.unibocconi.it/files/593.pdf

[6] Dogan, S., and Yildiz, K. (2019). Odds supermodularity and the Luce rule. Manuscript.

[7] Echenique, F., and Saito, K. (2019). General Luce Model. *Economic Theory*, 68, 811-826.

[8] Fishburn, P. C. (1974). Lexicographic orders, utilities and decision rules: A survey. *Management science*, 20, 1442-1471.

[9] Fishburn, P. C. (1978). Choice probabilities and choice functions. *Journal of Mathematical Psychology*, 18, 205-219.

[10] Horan, S. (2020). Stochastic semi-orders. Manuscript.

[11] Kreps, D. M. (1988). *Notes on the theory of choice*. Westview.

[12] Lindberg, P. O. (2012a). *Contributions to probabilistic discrete choice*. Thesis, Department of Transport Science, Royal Institute of Technology. http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-95402

[13] Lindberg, P. O. (2012b). The Choice Axiom revisited. In Lindberg (2012a). http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-95399

[14] Luce, R. D. (1957). A theory of individual choice behavior. Bureau of Applied Social Research records, Columbia University Library, 1957.

[15] Luce, R. D. (1959). *Individual choice behavior: A theoretical analysis*. Wiley.

[16] Luce, R. D. (1977). The choice axiom after twenty years. *Journal of Mathematical Psychology*, 15, 215-233.

[17] Matejka, F., and McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105, 272-298.

[18] McCausland, W. J. (2009). Random consumer demand. *Economica*, 76, 89-107.

[19] McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In Zarembka, P. (Ed.). *Frontiers in econometrics* (pp. 105-142). Academic Press.

[20] Ok, E. A., and Tserenjigmid, G. (2020). Deterministic rationality of stochastic choice behavior. Manuscript.

[21] Peters, H., and Wakker, P. (1991). Independence of irrelevant alternatives and revealed group preferences. *Econometrica*, 59, 1787-1801.

[22] Ray, P. (1973). Independence of irrelevant alternatives. *Econometrica*, 41, 987-991.

[23] Renyi, A. (1955). On a new axiomatic theory of probability. *Acta Mathematica Academiae Scientiarum Hungaricae*, 6, 285-335.

[24] Renyi, A. (1956). On conditional probability spaces generated by a dimensionally ordered set of measures. *Theory of Probability and Its Applications*, 1, 55-64.

[25] Salant, Y., and Rubinstein, A. (2008). (A, f): choice with frames. *Review of Economic Studies*, 75, 1287-1296.

[26] Samuelson, P. A. (1938). A note on the pure theory of consumer's behaviour. *Economica*, 5, 61-71.

[27] Sen, A. (1993). Internal consistency of choice. *Econometrica*, 61, 495-521.

[28] Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge University Press.

[29] Wakker, P. P. (2010). *Prospect theory: For risk and ambiguity*. Cambridge University Press.