A Characterization of the Luce Choice Rule for an Arbitrary Collection of Menus*

Carlos Alós-Ferrer†
Lancaster University

Maximilian Mihm‡
New York University Abu Dhabi

Forthcoming, *Journal of Economic Theory.*
This is the author-generated version accepted for publication.
October 2024

Abstract

The Luce Choice Rule (or, equivalently, the multinomial logit model) is extensively used in economics and other fields. Classical characterizations rest on Luce's Choice Axiom, when all choice sets are available, and Luce's Product Rule in the case of binary choice. Yet, actual datasets typically consist neither of all choice sets nor all binary choice sets. We provide a characterization for the general case, allowing also for zero choice probabilities. Building upon this characterization, we derive implications for experimental design in terms of three criteria: falsification, identification, and prediction.

JEL Classification: D01 · D90

Keywords: Luce Choice Rule · Multinomial Logit Model · Stochastic Choice · Censored Luce Choice Rule

1 Introduction

The Luce Choice Rule is one of the most prominent models of stochastic choice. With an appropriate transformation, it is equivalent to the multinomial logit

*We thank Yves Breitmoser, Federico Echenique, Fabio Maccheroni, three anonymous reviewers, and seminar participants at Boston University and Brown University for helpful comments.

†Lancaster University Management School. Lancaster, LA1 4YX, United Kingdom. Email: c.alosferrer@lancaster.ac.uk

‡New York University Abu Dhabi, Division of Social Sciences. Email: max.mihm@nyu.edu.

model, which is widely used to model empirical choice data (McFadden, 1974, 2001). Every logistic regression on discrete choice data implicitly assumes that utilities are translated into choices through the Luce Choice Rule (Moffatt, 2015). Applications extend to game theory, where Luce choice underlies the concept of (logit) Quantal Response Equilibrium (McKelvey and Palfrey, 1995, 1998; Goeree et al., 2005). Under the name "softmax function," logit choice is routinely used in neuroscience and cognitive psychology to fit models to data. The standard drift-diffusion model of Ratcliff (1978) (see also Ratcliff and Rouder, 1998; Fudenberg et al., 2018) and the rational inattention model of Sims (2003), applied to discrete choice, generate logit choice probabilities (Matejka and McKay, 2015).

In this paper, we provide a characterization of the Luce Choice Rule that applies when choice data is available for an arbitrary collection of menus (i.e., choice sets) and derive implications of the characterization for experimental design. Given the widespread applications of the Luce Choice Rule, it is surprising that both of the well-known existing characterizations do not apply to most experiments or surveys, where it is often infeasible to collect data for the requisite menus. The first characterization in Luce (1959) rests on the celebrated Luce Choice Axiom and requires data for all possible menus. The second characterization in Luce and Suppes (1965) rests on the Luce Product Rule and covers the case of binary choice (all menus of two alternatives). Neither characterization generalizes to a dataset with an arbitrary collection of menus. For instance, the two characterizations are not nested: the Luce Choice Axiom is vacuous for binary choice, while the Luce Product Rule is not sufficient in a framework with all menus.

To illustrate, consider a stochastic choice dataset consisting of a collection of menus $\mathcal{M}$ (the choice sets for which data is available) and the corresponding frequencies $P(a, M)$ with which any given alternative $a \in M$ is chosen from menu $M \in \mathcal{M}$. The Luce Choice Axiom states that the probability that alternative a will be chosen from menu M should be equal to the probability that a is chosen from any subset $M' \subseteq M$ such that $a \in M'$ times the total probability that alternatives from M' are chosen from M; that is, the following condition should be satisfied.

$$P(a, M) = P(a, M') \sum_{a' \in M'} P(a', M).$$

The Luce Product Rule considers only binary menus and states that, for any three alternatives,

$$P(a_1, M_{12})P(a_2, M_{23})P(a_3, M_{13}) = P(a_3, M_{23})P(a_2, M_{12})P(a_1, M_{13}),$$

where M_{ij} is shorthand for the binary menu $\{a_i, a_j\}$; that is, the product of choice frequencies forming an intransitive choice cycle should be equal in the clockwise and counterclockwise directions.

Now suppose that there are four alternatives $\{a, b, c, d\}$ but the dataset contains choice frequencies only for the following set of menus:

$$\mathcal{M} = \Big\{\{a, b, c\}, \{a, d\}, \{b, d\}, \{c, d\}\Big\}.$$

The condition for the Luce Choice Axiom is vacuously satisfied for the three binary menus and, for the menu $\{a, b, c\}$, the condition cannot be verified because the dataset does not include the menus $\{a, b\}$, $\{a, c\}$ or $\{b, c\}$. Moreover, there are no choice cycles over binary menus in this dataset and the Luce Product Rule is therefore vacuous. Thus, neither the Luce Choice Axiom nor the Luce Product Rule provide any testable implications of the Luce Choice Rule. However, there are testable implications. For instance, it is easily verified that observing both $P(a, \{a, d\}) > P(b, \{b, d\})$ and $P(a, \{a, b, c\}) < P(b, \{a, b, c\})$ is not consistent with the Luce Choice Rule.

We introduce an axiom—the *General Product Rule* (GPR)—that characterizes the Luce Choice Rule for an arbitrary collection of menus. First, we say that an array of alternatives and menus $[a_1, \ldots, a_{n+1}; M_1, \ldots, M_n]$ form an *overlapping sequence* if data for all menus $M_1, \ldots, M_n$ is available and, for all $i = 1, \ldots, n$, the alternative a_i and its successor a_{i+1} are different but are both available in the menu M_i; that is, the alternative represent choices along a chain of overlapping menus. The alternatives on such an overlapping sequence are said to be connected; which we show is an equivalence relation. We then say that an overlapping sequence is an overlapping cycle if $a_1 = a_{n+1}$; that is, when the alternatives form an intransitive choice cycle. Similar to the Luce Product Rule, the GPR requires that the product of choice frequencies along an intransitive cycle should be equal in the clockwise and counterclockwise direction, but requires this to hold only for overlapping cycles (of any length) in the dataset.

In contrast to the Luce Choice Axiom and the Luce Product Rule, the GPR therefore captures restrictions on both binary and non-binary menus, but does not require data from all menus or all binary menus to be available. For instance, the above example has many overlapping cycles. In particular, $[a, b, d, a; \{a, b, c\}, \{b, d\}, \{a, d\}]$ and $[d, b, a, d; \{b, d\}, \{a, b, c\}, \{a, d\}]$ are two overlapping cycles that form intransitive choice cycles in the clockwise and counterclockwise direction and the GPR requires that

$$P(a, \{a, b, c\})P(b, \{b, d\})P(d, \{a, d\}) = P(d, \{b, d\})P(b, \{a, b, c\})P(a, \{a, d\}),$$

which is violated when $P(a, \{a, d\}) > P(b, \{b, d\})$ and $P(a, \{a, b, c\}) < P(b, \{a, b, c\})$. Our main result shows that a stochastic choice dataset can be rationalized by the Luce Choice Rule if and only if the product of choice frequencies along *every* overlapping cycle in the dataset is equal in either direction of the cycle.

Additionally, in our characterization, we allow for *censored alternatives*; that is, we do not impose *Positivity* (an axiom in Luce, 1959; Luce and Suppes, 1965) but rather allow some alternatives in a menu to be chosen with probability zero. This is important in practice because, empirically, it cannot be guaranteed that all alternatives are chosen in each menu. Moreover, as we will show, there is a close connection between restrictions on menus and censoring of alternatives. The reason is that, if a menu is not available in a dataset, one could artificially add it while specifying that all alternatives in this menu but one are not chosen. On the censored experiment, the added menu is then not part of any overlapping cycle and therefore introduces no testable restrictions for the Luce Choice Rule.

A recent literature has considered the Luce Choice Rule with censored alternatives, providing characterizations of two-stage logit models where first a criterion is used to discard some alternatives (e.g., if the utility is below a certain threshold) and then logit choice is applied to the remaining alternatives. Existing models with censored alternatives include Ahumada and Ülkü (2018), Echenique and Saito (2019), Horan (2021), and Doğan and Yıldız (2021). All of these, however, consider a framework where all menus are available; for this particular case, our result boils down to previous theorems in Ahumada and Ülkü (2018) and Echenique and Saito (2019). Actually, because of the relationship between restrictions on menus and censoring of alternatives, one could also build an alternative proof of our main result starting from the theorems in Ahumada and Ülkü (2018) and Echenique and Saito (2019), although none of these provide the characterization when the menus are restricted (see Remark 2 in Section 3.1).

The extensive literature on logit choice has also considered many other generalizations and variants, but all require a rich class of menus. For instance, generalizations allowing for established behavioral anomalies (Gul et al., 2014; Faro, 2023), variants focusing on attention and perception (Echenique et al., 2018; Tserenjigmid, 2021; Kovach and Tserenjigmid, 2022b; Heydari, 2021), and characterizations pinning down or fixing the utility function (Ahn et al., 2018; Breitmoser, 2021). We discuss related literature further in Section 5.

Our results are of interest for the empirical discrete-choice literature since they clarify the testable conditions that characterize the Luce Choice Rule for general datasets. They are also of interest for experimental work where practical considerations (such as monetary or time constraints) often restrict the set of menus on which choice data is collected. We apply our characterization to consider three potential criteria that may factor into an experimental design, where the experiment consists of the menus chosen by the experimenter *before* stochastic choice data is collected from the experimental participants.

First, we consider experiments that allow for a *falsification* of the Luce Choice Rule. With a rich collection of menus (e.g., all menus or all binary menus), it is always possible to falsify the Luce Choice Rule. However, as we show, there are potential experiments on which *every* stochastic choice dataset is Luce rationalizable. For instance, if the menu $\{a, b, c\}$ is not included in the above example, there are no overlapping cycles, and the Luce Choice Rule is not falsifiable. An experimental researcher interested in providing evidence of Luce behavior may therefore need to include (or add) specific menus to ensure that the Luce Choice Rule has testable implications. We hence characterize when the choice data from a new experiment can be used to falsify the Luce Choice Rule on a previous experiment.

Second, we consider the *identification* of the utilities for the alternatives in the Luce Choice Rule. With a rich collection of menus (e.g., all menus or all binary menus), the utilities in the Luce Choice Rule are always unique up to a rescaling by a strictly positive constant. However, as we show, this identification does not extend: for an arbitrary collection of menus, utilities are identified up to rescaling *within* each equivalence class (generated by the overlapping sequences) but are completely independent *across* equivalence classes. As a result, there are experiments on which, for any Luce rationalizable dataset, it would not be possible to compare utilities across some alternatives. We hence characterize the experiments that allow for identification (up to rescaling) of all alternatives and also highlight some trade-offs between identification and falsification.

Finally, when practical considerations necessitate a restriction of the menus to be included for data collection, the ability to *predict* Luce choice behavior on other menus may also factor into an experimental design. We hence characterize when the stochastic choice data from an experiment can be used to uniquely predict choice behavior consistent with the Luce Choice Rule for out-of-sample menus.

The paper is structured as follows. Section 2 introduces the framework. Section 3 describes the GPR, states and proves the main characterization result, and discusses the relation to previous axioms. We also provide a tighter characterization that removes some redundancies. Section 4 discusses implications for experimental design. Section 5 reviews the related literature and Section 6 concludes. Except for the main characterization result, all proofs are in the Appendix.

2 Framework

2.1 Stochastic choice dataset

Let $\mathcal{A}$ be a finite set of *alternatives* and $\mathcal{M}^* = 2^{\mathcal{A}} \backslash \{\emptyset\}$ be the collection of all nonempty *menus* of alternatives. We call a non-empty collection of menus $\mathcal{M} \subseteq \mathcal{M}^*$ an *experiment* and $\mathcal{A}(\mathcal{M}) = \bigcup_{M \in \mathcal{M}} M$ the *range* of the experiment.

For an experiment $\mathcal{M}$, let $\mathcal{S}(\mathcal{M}) = \{(a, M) \in \mathcal{A} \times \mathcal{M} \mid a \in M\}$ be the set of possible *choice observations*, where (a, M) is interpreted as the observation that, when offered menu M, a decision maker chose the alternative $a \in M$. A *stochastic choice function* (SCF) on experiment $\mathcal{M}$ is a mapping $P : \mathcal{S}(\mathcal{M}) \mapsto [0, 1]$ such that $\sum_{a \in M} P(a, M) = 1$ for all $M \in \mathcal{M}$.

A *stochastic choice dataset* is a pair $\theta = (\mathcal{M}, P)$ consisting of an experiment $\mathcal{M}$ and a SCF P on $\mathcal{M}$. The experiment $\mathcal{M}$ is the set of menus offered to the decision maker(s), which is chosen by nature or the experimental designer. For each menu $M \in \mathcal{M}$, $P(a, M)$ is the empirical frequency or proportion of times that the alternative a was chosen from M by the decision maker(s). When $\mathcal{M} \neq \mathcal{M}^*$ we say that menus have been *restricted* in the sense that choice data has not been collected for all possible menus.

A stochastic choice dataset $(\mathcal{M}, P)$ fulfills *Positivity* if $P(a, M) > 0$ for all $(a, M) \in \mathcal{S}(\mathcal{M})$. This is an axiom in the original characterizations of the Luce Choice Rule (Luce, 1959; Luce and Suppes, 1965). We will not, however, assume Positivity, i.e., we explicitly consider stochastic choice datasets where some alternatives are chosen with zero probability in some menus. In this case, we speak of *censored alternatives*, possibly in addition to restricted menus.

For an experiment $\mathcal{M}$, a *choice correspondence* is a mapping $C : \mathcal{M} \to \mathcal{M}^*$ such that $C(M) \subseteq M$ for all $M \in \mathcal{M}$; in that case, $\mathcal{A}^C(\mathcal{M}) = \mathcal{A}(\{C(M) \mid M \in \mathcal{M}\})$ is the C-censored range. In particular, for a stochastic choice dataset $\theta = (\mathcal{M}, P)$, the choice correspondence $C_\theta(M) := \{a \in M \mid P(a, M) > 0\}$ is the *support* of P on $\mathcal{M}$. The support captures the empirical observation of which alternatives have

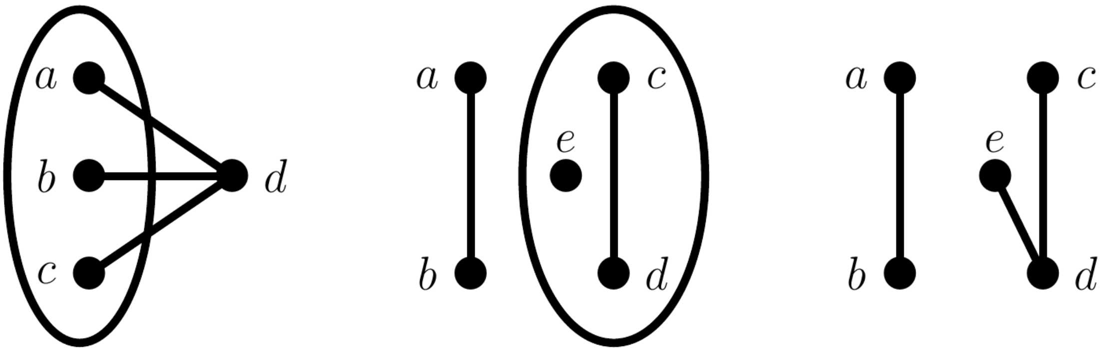

Figure 1: Graphical depictions of the three experiments in Example 1.

actually been chosen in each menu. The support is the identity mapping if and only if θ satisfies Positivity.

Example 1. Experiments can be represented graphically by depicting alternatives as nodes, binary menus as edges connecting nodes, and other menus with outlines containing nodes. For instance, Figure 1(left) illustrates the example from the introduction: an experiment $\mathcal{M}$ with range $\mathcal{A}(\mathcal{M}) = \{a, b, c, d\}$ and menus $\mathcal{M} = \{\{a, b, c\}, \{a, d\}, \{b, d\}, \{c, d\}\}$. Figure 1(center) represents a different experiment $\mathcal{M}'$ with range $\mathcal{A}(\mathcal{M}') = \{a, b, c, d, e\}$ and menus $\mathcal{M}' = \{\{a, b\}, \{c, d\}, \{c, d, e\}\}$. Figure 1(right) represents the C-censored experiment on $\mathcal{M}'$ when $C(\{a, b\}) = \{a, b\}$, $C(\{c, d\}) = \{c, d\}$, and $C(\{c, d, e\}) = \{d, e\}$. □

2.2 Luce Choice Rule

Given an experiment $\mathcal{M}$, a choice correspondence $C : \mathcal{M} \to \mathcal{M}^*$, and a utility function $v : \mathcal{A}(\mathcal{M}) \mapsto \mathbb{R}_{++}$, we denote by

$$P_v^C(a, M) = \begin{cases} \dfrac{v(a)}{\displaystyle\sum_{b \in C(M)} v(b)} & \text{if } a \in C(M) \\ 0 & \text{if } a \notin C(M) \end{cases} \tag{1}$$

the C-censored Luce SCF generated by v on the experiment $\mathcal{M}$.[1]

[1]Using the transformation $u(a) = \ln v(a)$, the first part of Equation (1) can be rewritten as $P(a, M) = e^{u(a)} \left(\sum_{b \in M} e^{u(b)}\right)^{-1}$ with $u : \mathcal{A}(\mathcal{M}) \mapsto \mathbb{R}$ a (not necessarily positive) real-valued function. This corresponds to the well-known *(multinomial) logit model* (e.g., McFadden, 2001). By virtue of the logarithmic transformation, the Luce Choice Rule and the logit model are equivalent in the sense that a stochastic choice dataset can be rationalized by one if and only if it can be rationalized by the other (see, e.g., Anderson et al., 1992, Chapter 1).

A stochastic choice dataset $\theta = (\mathcal{M}, P)$ can be rationalized by a (censored) *Luce Choice Rule* (or is *Luce rationalizable*), if there exists a utility function $v : \mathcal{A}(\mathcal{M}) \mapsto \mathbb{R}_{++}$ such that

$$P(a, M) = P_v^{C_\theta}(a, M) \quad \forall\, (a, M) \in \mathcal{S}(\mathcal{M}).$$

We denote the set of utility functions that can Luce rationalize the stochastic choice dataset $\theta = (\mathcal{M}, P)$ by

$$\mathcal{V}(\theta) = \left\{ v : \mathcal{A}(\mathcal{M}) \to \mathbb{R}_{++} \;\middle|\; P = P_v^{C_\theta} \right\}.$$

Let $\mathcal{P}(\mathcal{M})$ be the set of all SCFs on experiment $\mathcal{M}$, and $\mathcal{L}(\mathcal{M})$ be the set of SCFs $P \in \mathcal{P}(\mathcal{M})$ such that the stochastic choice dataset $(\mathcal{M}, P)$ is Luce-rationalizable; hence $P \in \mathcal{L}(\mathcal{M})$ if and only if $\mathcal{V}(\mathcal{M}, P) \neq \emptyset$. It is also sometimes useful to classify SCFs in terms of their support. Given an experiment $\mathcal{M}$ and a choice correspondence C defined on any experiment that contains $\mathcal{M}$, we denote by $\mathcal{P}^C(\mathcal{M}) = \left\{ P \in \mathcal{P}(\mathcal{M}) \;\middle|\; C_{(\mathcal{M},P)}(M) = C(M)\; \forall\, M \in \mathcal{M} \right\}$ the set of C-censored SCFs on $\mathcal{M}$ and by $\mathcal{L}^C(\mathcal{M}) = \mathcal{P}^C(\mathcal{M}) \cap \mathcal{L}(\mathcal{M})$ the C-Luce rationalizable SCFs.

2.3 Connected alternatives and overlapping cycles

Given an experiment $\mathcal{M}$ and a choice correspondence C on $\mathcal{M}$, an array of alternatives and menus $\phi = [a_1, \ldots, a_{n+1}; M_1, \ldots, M_n]$ is an *overlapping C-sequence (of length $n \geq 1$) in $\mathcal{M}$* if $M_i \in \mathcal{M}$ and $a_i, a_{i+1} \in C(M_i)$ with $a_i \neq a_{i+1}$ for all $i \in \{1, \ldots, n\}$. For convenience, we denote the i-th alternative by $a_i(\phi)$ and the i-th menu by $M_i(\phi)$.

Alternatives a and b are *C-connected in $\mathcal{M}$*, written $a \sim_{\mathcal{M}}^C b$, if there is an overlapping C-sequence ϕ of length n in $\mathcal{M}$ such that $a = a_1(\phi)$ and $b = a_{n+1}(\phi)$. Lemma 1 in the Appendix shows that $\sim_{\mathcal{M}}^C$ is an equivalence relation on $\mathcal{A}^C(\mathcal{M})$, and we denote the corresponding quotient set by $\mathcal{Q}^C(\mathcal{M})$.

If ϕ is an overlapping C-sequence of length n in $\mathcal{M}$ and $a_1(\phi) = a_{n+1}(\phi)$, then ϕ is an *overlapping C-cycle* in $\mathcal{M}$. An overlapping cycle ϕ of length n is degenerate if $M_i(\phi) = M_j(\phi)$ for all $i, j = 1, \ldots, n$, and non-degenerate otherwise.

Example 1 (continued). The experiment $\mathcal{M}$ in Figure 1(left) has, among others, the nondegenerate overlapping cycle $[a, b, c, d, a; \{a, b, c\}, \{a, b, c\}, \{c, d\}, \{a, d\}]$ and all alternatives are connected. The experiment $\mathcal{M}'$ in Figure 1(center) has the nondegenerate overlapping cycle $[c, d, e, c; \{c, d\}, \{c, d, e\}, \{c, d, e\}]$ and two equivalence classes $\mathcal{Q}(\mathcal{M}') = \{\{a, b\}, \{c, d, e\}\}$. The C-censored experiment on $\mathcal{M}'$ in Figure

1(right) has the same equivalence classes as the uncensored experiment but has no non-degenerate overlapping cycles. □

Remark 1. Luce (1959, Definition 1, p. 25) introduced the concept of a *finitely connected domain*, which required that all a, b with $P(a, \{a, b\}) > 1/2$ be linked by a chain with $P(a_i, \{a_i, a_{i+1}\}) \in [1/2, 1)$ for all $i = 1, \dots, n$. Suppes et al. (1989, Chapter 17, Theorem 7), Horan (2021), and also the working-paper version of Echenique and Saito (2019) consider the related concept of a *linked domain*: assuming that $\mathcal{M}$ contains all binary menus, every pair of alternatives a, b are linked by a chain of alternatives $a = a_1, a_2, \dots, a_{n+1} = b$ such that a_i, a_{i+1} are imperfectly discriminated for all $i = 1, \dots, n$, meaning that $P(a_i, \{a_i, a_{i+1}\}) \in (0, 1)$ (and hence $a_i \neq a_{i+1}$). If C is the support of P, an overlapping C-sequence is essentially a *chain of imperfect discrimination* where the binary menus are replaced by arbitrary ones. We do not impose a linked (or finitely connected) domain but merely use overlapping C-sequences to state conditions for our results.

3 Luce rationalizable datasets

Our main result characterizes the set of Luce rationalizable stochastic choice datasets for *any* experiment. To this end, we say that a stochastic choice dataset $\theta = (\mathcal{M}, P)$ satisfies the *General Product Rule (GPR)* if the following holds.

Axiom 1 (GPR)**.** For any overlapping C_θ-cycle $[a_1, \dots, a_{n+1}; M_1, \dots, M_n]$,

$$\prod_{i=1}^{n} P(a_i, M_i) = \prod_{i=1}^{n} P(a_{i+1}, M_i). \tag{2}$$

Each side of Equation (2) describes products of non-zero choice probabilities along an intransitive choice cycle in $\theta = (\mathcal{M}, P)$. On the left-hand side, a_1 is chosen when a_2 is available, a_2 is chosen when a_3 is available, and so on, but a_n is chosen when a_1 is available; hence, the intransitive choice cycle $a_1 \to a_2 \to \dots \to a_n \to a_1$. On the right-hand side, a_2 is chosen when a_1 is available, a_3 is chosen when a_2 is available, and so on, but a_1 is chosen when a_n is available; hence, the counterclockwise choice cycle $a_1 \leftarrow a_2 \leftarrow \dots \leftarrow a_n \leftarrow a_1$. The GPR, therefore, requires that products of choice probabilities be equal for the clockwise and counterclockwise cycles; intuitively, violations of transitivity are not systematic in the sense that one direction of an intransitive cycle is not more likely than another.

Theorem 1. *A stochastic choice dataset $\theta = (\mathcal{M}, P)$ for an arbitrary collection of menus $\mathcal{M} \subseteq \mathcal{M}^*$ can be rationalized by a (censored) Luce Choice Rule if and only if it satisfies the General Product Rule.*

Proof. For necessity, suppose $\theta = (\mathcal{M}, P)$ can be rationalized by a Luce Choice Rule: there is $v : \mathcal{A}(\mathcal{M}) \to \mathbb{R}_{++}$ such that (1) is satisfied for all $(a, M) \in \mathcal{S}(\mathcal{M})$. Let $[a_1, \ldots, a_{n+1}; M_1, \ldots, M_n]$ be an overlapping C_θ-cycle in $\mathcal{M}$. Then,

$$\begin{aligned}\prod_{i=1}^{n} P(a_i, M_i) &= \prod_{i=1}^{n} \left(\frac{v(a_i)}{\sum_{b \in M_i} v(b)} \right) = \frac{\prod_{i=1}^{n} v(a_i)}{\prod_{i=1}^{n} \left(\sum_{b \in M_i} v(b) \right)} \\ &= \frac{\prod_{i=1}^{n} v(a_{i+1})}{\prod_{i=1}^{n} \left(\sum_{b \in M_i} v(b) \right)} = \prod_{i=1}^{n} \left(\frac{v(a_{i+1})}{\sum_{b \in M_i} v(b)} \right) = \prod_{i=1}^{n} P(a_{i+1}, M_i),\end{aligned}$$

and so the GPR is satisfied.

The argument for sufficiency is as follows. Lemma 2 in the Appendix shows that $\theta = (\mathcal{M}, P)$ can be rationalized by a Luce Choice Rule if and only if each equivalence class $Q \in \mathcal{Q}^{C_\theta}(\mathcal{M})$ can be, independently, rationalized by a Luce Choice Rule. Therefore, it is without loss of generality to focus on a stochastic choice dataset $\theta = (\mathcal{M}, P)$ where $a \sim_M^{C_\theta} b$ for all $a, b \in \mathcal{A}(\mathcal{M})$ (if not, the same argument can be applied to each equivalence class separately).

Now fix an arbitrary $a^* \in \mathcal{A}^{C_\theta}(\mathcal{M})$ and define $v(a^*) = 1$. For $a \in \mathcal{A}^{C_\theta}(\mathcal{M})$, let $[a_1, \ldots, a_{n+1}; M_1, \ldots, M_n]$ be an overlapping C_θ-sequence with $a = a_1$ and $a_{n+1} = a^*$. Define

$$v(a) = \prod_{i=1}^{n} \frac{P(a_i, M_i)}{P(a_{i+1}, M_i)}.$$

We need to show that $v(a)$ is well-defined, that is, $v(a)$ is independent of the chosen sequence. Let $[b_1, \ldots, b_{n+1}; M'_1, \ldots, M'_n]$ be another overlapping C_θ-sequence with $a = b_1$ and $b_{n+1} = a^*$. Then, we observe that

$$[a_1, \ldots, a_{n+1}, b_n, \ldots, b_2, b_1; M_1, \ldots, M_n, M'_n, \ldots, M_1]$$

is an overlapping C_θ-cycle. By Equation (2),

$$\prod_{i=1}^{n} P(a_i, M_i) \cdot P(a_{n+1}, M'_n) \prod_{i=2}^{n} P(b_i, M'_{i-1}) = \prod_{i=1}^{n} P(a_{i+1}, M_i) \prod_{i=1}^{n} P(b_i, M'_i)$$

and, since $a_{n+1} = a^* = b_{n+1}$,

$$\prod_{i=1}^{n} \frac{P(a_i, M_i)}{P(a_{i+1}, M_i)} = \prod_{i=1}^{n} \frac{P(b_i, M'_i)}{P(b_{i+1}, M'_i)}$$

and hence $v(a)$ is well-defined.

Consider any $M \in \mathcal{M}$ and any $a \in C_\theta(M)$. Let $[a_1, \ldots, a_{n+1}; M_1, \ldots, M_n]$ be an overlapping C_θ-sequence with $a = a_1$ and $a_{n+1} = a^*$. For any $b \in C_\theta(M)$ with $b \neq a$, $[b, a_1, \ldots, a_{n+1}; M, M_1, \ldots, M_n]$ is an overlapping C_θ-sequence connecting b and a^*, and hence

$$v(b) = \frac{P(b, M)}{P(a, M)} \prod_{i=1}^{n} \frac{P(a_i, M_i)}{P(a_{i+1}, M_i)} = \frac{P(b, M)}{P(a, M)} v(a).$$

It follows that

$$\sum_{b \in c_\theta(M)} v(b) = \frac{v(a)}{P(a, M)} \sum_{b \in c_\theta(M)} P(b, M) = \frac{v(a)}{P(a, M)},$$

implying that

$$P(a, M) = \frac{v(a)}{\sum_{b \in c_\theta(M)} v(b)},$$

which completes the proof. □

3.1 Special cases

We briefly discuss some special cases that highlight the relation to prior characterizations of the Luce Choice Rule.

Unrestricted menus and noncensored alternatives. When all menus are available, Luce (1959) shows that a stochastic choice dataset can be rationalized by a Luce Choice Rule if and only if it satisfies Positivity and the Luce Choice Axiom. Adapted to an arbitrary collection of menus, $(\mathcal{M}, P)$ satisfies the Luce Choice Axiom if

$$P(a, A) = P(a, B) \sum_{b \in B} P(b, A).$$

whenever $a \in B \subseteq A$ for menus $A, B \in \mathcal{M}$. With Positivity, the Luce Choice Axiom is equivalent to the Independence of Irrelevant Alternatives (IIA):[2] if $a, b \in A \cap B \in \mathcal{M}$ for menus $A, B \in \mathcal{M}$, then

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

[2] Luce (1959) shows the equivalence between the Luce Choice Axiom and IIA when all menus are available but, with the adapted definition of the axioms, it is straightforward to adapt the argument to an arbitrary collection of menus.

Under Positivity, the GPR implies IIA and hence the Luce Choice Axiom because, if $a, b \in A \cap B \in \mathcal{M}$ for $A, B \in \mathcal{M}$, then the GPR implies $P(a, A)P(b, B) = P(b, A)P(a, B)$.

Binary menus and noncensored alternatives. When the experiment consists of all binary menus $\mathcal{B} = \{M \in \mathcal{M}^* \mid |M| = 2\}$, Luce and Suppes (1965) show that a stochastic choice dataset can be rationalized by a Luce Choice Rule if and only if it satisfies Positivity and the Luce Product Rule. Adapted to an arbitrary collection of menus, $(\mathcal{M}, P)$ satisfies the Luce Product Rule if

$$P(a, \{a, b\})P(b, \{b, c\})P(c, \{a, c\}) = P(b, \{a, b\})P(c, \{b, c\})P(a, \{a, c\})$$

whenever $a, b, c \in \mathcal{A}$ are distinct and $\{a, b\}, \{a, c\}, \{b, c\} \in \mathcal{M}$. Clearly, under Positivity, the GPR implies the Luce Product Rule. Moreover, it follows immediately from the argument in Luce and Suppes (1965, footnote 9, p. 341) that the two axioms are equivalent when $\mathcal{M} = \mathcal{B}$.

Unrestricted menus and censored alternatives. The case of censored alternatives, but with unrestricted menus, has been studied in Ahumada and Ülkü (2018), Echenique and Saito (2019), and Horan (2021). All those works consider a product rule similar to the GPR, called *Axiom 1* by Ahumada and Ülkü (2018), *Cyclical Independence* by Echenique and Saito (2019), and *Strong Product Rule* by Horan (2021). For any given support, the special case of Theorem 1 where $\mathcal{M} = \mathcal{M}^*$ encompasses the results in this prior literature, specifically Ahumada and Ülkü (2018, Theorem 1) and Echenique and Saito (2019, Theorem 1). However, these works provide additional results characterizing specific models for how the choice correspondence C_θ arises, which is not our focus (see Section 5).

Remark 2. An alternative proof of Theorem 1 could be given starting from Theorem 1 of Ahumada and Ülkü (2018) or Theorem 1 of Echenique and Saito (2019). The idea is to extend a stochastic choice dataset $(\mathcal{M}, P)$ to another stochastic choice dataset $(\mathcal{M}^*, P')$ by defining $P'(\cdot, M) = P(\cdot, M)$ for all $M \in \mathcal{M}$, and setting $P'(a, M) = 1$ for some arbitrary $a \in M$ whenever $M \in \mathcal{M}^* \setminus \mathcal{M}$. It can be shown that the menus in $\mathcal{M}^* \setminus \mathcal{M}$ do not create any additional overlapping cycles in the censored experiment. As a result, the GPR on $(\mathcal{M}, P)$ boils down to the axioms of Ahumada and Ülkü (2018) and Echenique and Saito (2019) on $(\mathcal{M}^*, P')$. Applying their theorems to the extended stochastic choice dataset (and appropriately undoing the transformation afterwards) would yield our characterization. Hence, results when not all menus are available and when not all menus are observed can be derived from each other. However, our approach is conceptually simpler as the dataset in our characterization corresponds to the actual empirical dataset and not to an artificial extension of it.

Restricted menus and noncensored alternatives. An important special case of Theorem 1 is the characterization of the noncensored Luce Choice Rule for an arbitrary collection of menus; the proof follows immediately because C_θ is the identity if and only if θ satisfies Positivity.

Corollary 1. *A stochastic choice dataset $\theta = (\mathcal{M}, P)$ for an arbitrary collection of menus $\mathcal{M} \subseteq \mathcal{M}^*$ can be rationalized by an noncensored Luce Choice Rule if and only if it satisfies Positivity and the General Product Rule.*

3.2 Characterization by Elementary Cycles

By Theorem 1, the Luce Choice Rule provides a testable restriction for every overlapping C_θ-cycle. However, many of these restrictions are redundant. To provide a tighter characterization, we say that an overlapping C-cycle ϕ is *elementary* if it is non-degenerate and $a_i(\phi) \neq a_j(\phi)$ whenever $i \neq j$. Write $\mathcal{E}^C(\mathcal{M})$ for the set of elementary C-cycles on experiment $\mathcal{M}$.

If Equation (2) is satisfied for all overlapping C_θ-cycles, then it is clearly satisfied for the elementary cycles. Moreover, all non-elementary overlapping cycles can be obtained as the juxtaposition of two or more elementary cycles, and it is therefore sufficient to verify that Equation (2) is satisfied for elementary cycles.

Proposition 1. *A stochastic choice dataset $\theta = (\mathcal{M}, P)$ is Luce rationalizable if and only if Equation (2) is satisfied for every elementary C_θ-cycle in $\mathcal{E}^{C_\theta}(\mathcal{M})$.*

Proposition 1 reduces the number of conditions that need to be checked to verify if a stochastic choice dataset is Luce rationalizable. It also permits a characterization of experiments on which it is possible to verify the Luce Choice Rule independently. Here, it is useful to classify SCFs according to their support. In addition, we require notation for the restriction of SCF to a smaller experiment. If $P \in \mathcal{P}^C(\mathcal{M})$, we denote by $P_{\mathcal{M}'}$ the restriction of P to the smaller experiment $\mathcal{M}' \subset \mathcal{M}$, defined by $P_{\mathcal{M}'}(a, M) = P(a, M)$ for all $(a, M) \in \mathcal{S}(\mathcal{M}')$. Given a choice correspondence C describing the support, the following definition then formalizes when two experiments are independent.

Definition 1. Two disjoint experiments $\mathcal{M}_1$ and $\mathcal{M}_2$ are *C-Luce independent* if, for any SCF $P \in \mathcal{P}^C(\mathcal{M}_1 \cup \mathcal{M}_2)$,

$$P \in \mathcal{L}^C(\mathcal{M}_1 \cup \mathcal{M}_2) \quad \Longleftrightarrow \quad P_{\mathcal{M}_i} \in \mathcal{L}^C(\mathcal{M}_i) \quad \text{for } i = 1, 2.$$

When two experiments are Luce independent, one can therefore verify whether the stochastic choice dataset on the union is Luce rationalizable by, independently, verifying whether its restrictions to $\mathcal{M}_1$ and $\mathcal{M}_2$, respectively, are Luce rationalizable. Properties of the set of elementary C-cycles characterizes when two experiments are Luce independent.

Proposition 2. *Two disjoint experiments $\mathcal{M}_1$ and $\mathcal{M}_2$ are C-Luce independent if and only if there is a partition of the set of elementary cycles on $\mathcal{M}_1 \cup \mathcal{M}_2$ into elementary C-cycles that are only on $\mathcal{M}_1$ and elementary cycles that are only on $\mathcal{M}_2$; that is, $\mathcal{E}^C(\mathcal{M}_1 \cup \mathcal{M}_2) = \mathcal{E}^C(\mathcal{M}_1) \cup \mathcal{E}^C(\mathcal{M}_2)$.*[3]

Whether two experiments are Luce independent or not therefore depends on whether the union of the experiments introduces new elementary cycles that do not already belong to one of the two experiments. As a result, the set of elementary cycles can also be used to partition any experiment into smaller parts that are Luce independent of one another other. We say that $\{\mathcal{M}_1, \ldots, \mathcal{M}_K\}$ is an *independent C-partition* of experiment $\mathcal{M}$ if, for all $k = 1, \ldots, K$, the following are satisfied:

- $\mathcal{M}_k \subseteq \mathcal{M}$ and $\mathcal{M}_k$ is C-Luce independent of $\mathcal{M} \backslash \mathcal{M}_k$, and
- if $\mathcal{M}'_k \subsetneq \mathcal{M}_k$, then $\mathcal{M}'_k$ is not C-Luce independent of $\mathcal{M}_k \backslash \mathcal{M}'_k$.

That is, any two equivalence classes are C-Luce independent of each other and it is not possible to divide any equivalence class further into two experiments that are C-Luce independent of each other.

Corollary 2. *Every experiment has a unique independent C-partition. Moreover, if $\{\mathcal{M}_1, \ldots, \mathcal{M}_K\}$ is the independent C-partition for experiment $\mathcal{M}$, then $\mathcal{E}^C(\mathcal{M}_k) = \emptyset$ if and only if $|\mathcal{M}_k| = 1$, and $\mathcal{E}^C(\mathcal{M}) = \bigcup_{k:|\mathcal{M}_k|>1} \mathcal{E}^C(\mathcal{M}_k)$.*

Example 2. To illustrate, consider the three experiments in Figure 2 when C is the identity. The experiment on the left-hand-side has no elementary cycles, and every menu is a singleton equivalence class in the independent C-partition. In the experiment in the center, the independent C-partition has three equivalence classes:

[3] It is straightforward to show that, if two experiments are C-Luce independent when C is the identity mapping, then the experiments are C'-Luce independent for any given choice correspondence C'. However, the converse does not hold in general.

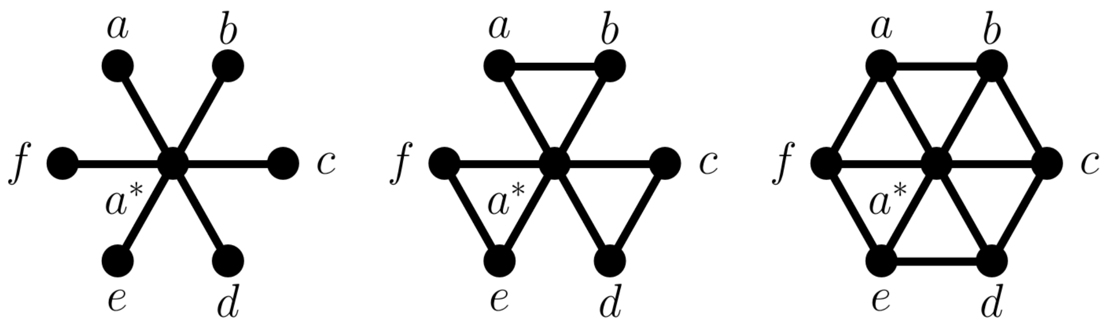

Figure 2: Experiments discussed in Example 2.

$\{\{a^*, a\}, \{a^*, b\}, \{a, b\}\}$, $\{\{a^*, c\}, \{a^*, d\}, \{c, d\}\}$, and $\{\{a^*, e\}, \{a^*, f\}, \{e, f\}\}$. Finally, the independent C-partition for the experiment on the right-hand-side has only one equivalence class that contains all menus. □

4 Experimental Design

In an experimental setting, a stochastic choice dataset consists of the experiment, which is chosen by the experimenter, and a SCF, which is the actual choice data collected from the experimental participants. Experimental design is concerned with the choice of an experiment, and practical considerations (e.g., monetary or time constraints) often dictate a judicious restriction of the menus for which choice data will be collected. We consider implications of our characterization in terms of three criteria that may factor into the design of an experiment: falsification (the power to reject the Luce Choice Rule), identification (the power to identify, up to rescaling, a unique utility for each alternative), and prediction (the power to predict Luce choice uniquely for out-of-sample menus).

There is, however, a caveat: the experiment is chosen *without* knowing the support, i.e., which alternatives will be censored. As our results emphasize, the implications of the Luce Choice Rule pertain to the stochastic choice data on the *censored* experiment. In an experimental design problem, we can view the restriction of menus as exogenous, but the censoring of menus is endogenous. This limitation is unavoidable.

Our approach will be to state conditions on experiments given censoring for an arbitrary but fixed choice correspondence. When the choice correspondence is the identity, the censored experiment coincides with the experiment chosen at the experimental design stage. Hence, our results apply directly when the SCF satisfies Positivity, since in this case the support is the identity. This will be our leading interpretation in the discussion. It is, however, useful to allow for the possibility of censored alternatives in the statement of results because there are various approaches in the literature to explicitly model the support in a censored Luce Choice Rule (see Section 5). In the conditions we state, additional restrictions coming from a specific model of the support (or any other prior information about the support) can be incorporated in the choice correspondence to adjust the conditions accordingly for the censored experiment.

To state our conditions, we require notation to describe the extension of a SCF to a larger domain. Suppose $\mathcal{M}_1$ and $\mathcal{M}_2$ are disjoint experiments, C is a choice correspondence on $\mathcal{M}_1 \cup \mathcal{M}_2$, and $P_1 \in \mathcal{P}^C(\mathcal{M}_1)$ is a C-censored SCF on $\mathcal{M}_1$. Then, we denote by $\mathcal{P}^C(\mathcal{M}_1 \cup \mathcal{M}_2 | P_1) = \{P \in \mathcal{P}^C(\mathcal{M}_1 \cup \mathcal{M}_2) \mid P_{\mathcal{M}_1} = P_1\}$ the set of C-censored extensions of P_1 to the experiment $\mathcal{M}_1 \cup \mathcal{M}_2$, and by $\mathcal{L}^C(\mathcal{M}_1 \cup \mathcal{M}_2 | P_1) = \mathcal{P}^C(\mathcal{M}_1 \cup \mathcal{M}_2 | P_1) \cap \mathcal{L}^C(\mathcal{M}_1 \cup \mathcal{M}_2)$ the set of C-Luce rationalizable extensions.

4.1 Falsification

Before choice data is collected, any experiment $\mathcal{M}$ can give rise to a stochastic choice dataset consistent with the Luce Choice Rule because $P_v \in \mathcal{L}(\mathcal{M})$ for *any* utility function $v : \mathcal{A}(\mathcal{M}) \to \mathbb{R}_{++}$. However, the Luce Choice Rule is not *falsifiable* for every experiment. We call Luce unfalsifiable an experiment on which every stochastic choice dataset can be rationalized by Luce choice.

Definition 2. An experiment $\mathcal{M}$ is *C-Luce unfalsifiable* if every C-censored stochastic choice dataset on $\mathcal{M}$ is Luce rationalizable, i.e., $\mathcal{L}^C(\mathcal{M}) = \mathcal{P}^C(\mathcal{M})$.

A corollary of Proposition 1 characterizes Luce unfalsifiable experiments for an arbitrary choice correspondence.

Corollary 3. *An experiment $\mathcal{M}$ is C-Luce unfalsifiable if and only if it has no elementary C-cycles, i.e., $\mathcal{E}^C(\mathcal{M}) = \emptyset$.*

An experiment without elementary cycles cannot, therefore, provide compelling evidence that choice behavior follows the Luce Choice Rule because *every* SCF on the experiment can be rationalized by a Luce Choice Rule.

Example 2 (continued). In a "star" experimental design, a *reference option* a^* is fixed, and participants are given binary menus of the form $\{a^*, a\}$ for all alternatives $a \in \mathcal{A}(\mathcal{M})$, $a \neq a^*$. An example of a star design is given on the left-hand side of Figure 2. This experiment is Luce unfalsifiable: Axiom GPR has no bite because there are no elementary cycles. However, if the experimenter added any additional menu to the design creating at least one elementary cycle, as in the examples on the center and right-hand side of Figure 2, the experiment would become Luce falsifiable. □

Star designs are common in experimental economics and psychology. For instance, Clithero (2018) collected choice frequencies only for a star design, and hence the dataset obtained there (also used in Alós-Ferrer et al., 2021) contains no elementary cycles (falsification was not the article's objective; further choices not involving the star center were made only once, in a second part). Davis-Stober et al. (2015) (also used in Alós-Ferrer and Garagnani, 2024) based their design on a star-shaped set of comparisons across lotteries, but (for different reasons than those explained here) also included additional comparisons creating elementary cycles.

To falsify the Luce Choice Rule, a stochastic choice dataset for a star experiment must be complemented with choice data from additional menus. However, it is not always the case that choice data from *any* additional menus can falsify the Luce Choice Rule on an existing stochastic choice dataset. The question is, therefore, given a stochastic choice dataset for an initial experiment, for which additional menus should data be collected to ensure falsifiability? Suppose that $\theta = (\mathcal{M}_1, P_1)$ is a Luce rationalizable stochastic choice dataset and $\mathcal{M}_2$ is an experiment that is disjoint from $\mathcal{M}_1$. Collecting additional choice data for the experiment $\mathcal{M}_2$ could falsify the Luce Choice Rule in one of two ways. First, if the experiment $\mathcal{M}_2$ is not Luce unfalsifiable, then the choice data restricted to $\mathcal{M}_2$ could be inconsistent with the Luce Choice Rule. This falsifies the Luce Choice Rule on $\mathcal{M}_1 \cup \mathcal{M}_2$ but not because of additional restrictions related to the initial stochastic choice dataset. Second, even when the choice data restricted to $\mathcal{M}_2$ is Luce rationalizable, the combined dataset on $\mathcal{M}_1 \cup \mathcal{M}_2$ may not be. This case is of interest because it highlights cross restrictions that choice data from the new experiment $\mathcal{M}_2$ imposes for the existing dataset on $\mathcal{M}_1$.

Definition 3. Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be two disjoint experiments, C be a choice correspondence on $\mathcal{M}_1 \cup \mathcal{M}_2$, and $P_1 \in \mathcal{L}^C(\mathcal{M}_1)$. Then, $\mathcal{M}_2$ can *C-Luce falsify* $\theta = (\mathcal{M}_1, P_1)$ if, for all extensions $P \in \mathcal{P}^C(\mathcal{M}_1 \cup \mathcal{M}_2 | P_1)$,

$$P \in \mathcal{L}^C(\mathcal{M}_1 \cup \mathcal{M}_2) \iff P_{\mathcal{M}_2} \in \mathcal{L}^C(\mathcal{M}_2).$$

The following proposition characterizes when a new experiment can Luce falsify an existing stochastic choice dataset.

Proposition 3. *The experiment* $\mathcal{M}_2$ *can* C*-Luce falsify* $\theta = (\mathcal{M}_1, P_1)$ *if and only if there is some elementary* C*-cycle on* $\mathcal{M}_1 \cup \mathcal{M}_2$ *that contains menus from both* $\mathcal{M}_1$ *and* $\mathcal{M}_2$*, i.e.,* $\mathcal{E}^C(\mathcal{M}_1 \cup \mathcal{M}_2) \neq \mathcal{E}^C(\mathcal{M}_1) \cup \mathcal{E}^C(\mathcal{M}_2)$*.*

For a given support, whether a new experiment can Luce falsify the stochastic choice data from an existing experiment is, therefore, a property only of the two experiments and does not depend on the particular SCF on the initial experiment. In particular, Luce falsifiability depends only on the structural properties of the two experiments, and not on the given SCF P_1: the new experiment $\mathcal{M}_2$ can Luce falsify stochastic choice data on the initial experiment $\mathcal{M}_1$ if and only if the elementary cycles on $\mathcal{M}_1 \cup \mathcal{M}_2$ cannot be partitioned into elementary cycles only on $\mathcal{M}_1$ and elementary cycles only on $\mathcal{M}_2$. In particular, this means that experiments are not Luce independent. As a result, the conditions for falsification are symmetric: if $\mathcal{M}_2$ cannot Luce falsify choice on $\mathcal{M}_1$, then $\mathcal{M}_1$ also cannot Luce falsify choice on $\mathcal{M}_2$.

Example 3. Let $\mathcal{A} = \{a, b, c, d, e\}$ and consider experiment $\mathcal{M}_1 = \{\{a, c\}, \{b, c\}\}$, depicted by solid lines on the left of Figure 3. This experiment contains no elementary cycle, hence it is Luce unfalsifiable. However, the experiment $\mathcal{M}_2 = \{\{c, d\}, \{c, e\}, \{d, e\}\}$, depicted by dashed lines on the left of Figure 3 is Luce falsifiable, because it contains elementary cycle $\phi = [c, d, e, c; \{c, d\}, \{d, e\}, \{c, e\}]$.

Now consider the joint experiment $\mathcal{M}_1 \cup \mathcal{M}_2$. The addition of $\mathcal{M}_2$ to $\mathcal{M}_1$ does generate a new condition for falsifiability, the one corresponding to ϕ. However, this new condition does not impose any restriction on the data on $\mathcal{M}_1$. Hence, neither of the two experiments can falsify each other.

However, consider instead the example in the center of Figure 3, where $\mathcal{A} = \{a, b, c, d\}$, $\mathcal{M}_1 = \{\{a, b\}, \{a, c\}\}$ (depicted by solid lines), and $\mathcal{M}_2 = \{\{b, d\}, \{c, d\}\}$ (depicted by dashed lines). Separately, both experiments $\mathcal{M}_1$ and $\mathcal{M}_2$ are Luce unfalsifiable, but taken together the larger experiment $\mathcal{M} = \mathcal{M}_1 \cup \mathcal{M}_1$ is Luce falsifiable, as it contains an elementary cycle. That is, neither experiment would be able to provide data falsifying the Luce Choice Rule, but when one is added to the other, the Luce choice has testable implications.

The example on the right of Figure 3 is as the previous one, but the experiment $\mathcal{M}_1$ includes the menu $\{b, c\}$. Hence $\mathcal{M}_1$ is Luce falsifiable in itself, as it includes an elementary cycle. Further, $\mathcal{M}_2$ can Luce falsify $\mathcal{M}_1$ (and vice versa), as its addition creates a new elementary cycle which was also not present in $\mathcal{M}_2$. □

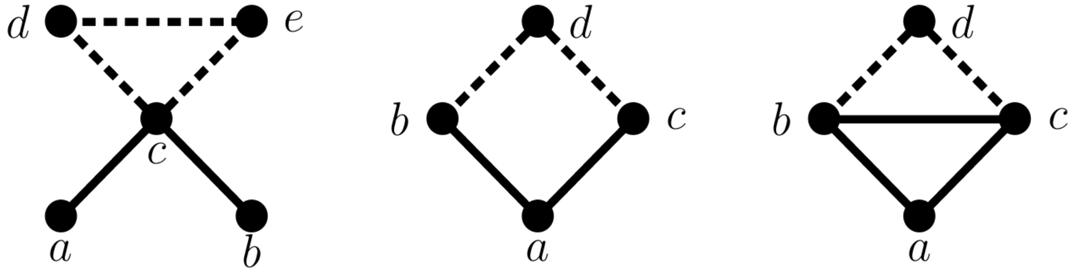

Figure 3: Experiments discussed in Example 3.

4.2 Identification

When choice data for either all menus or all binary menus are available, the utility function in a Luce Choice Rule is unique up to rescaling by a strictly positive constant (see, e.g., Luce and Suppes, 1965, Chapter 5.2). That means that any Luce rationalizable dataset pins down uniquely the ratio of utilities for any two alternatives. However, for an arbitrary collection of menus, the utilities are only unique up to rescaling *within* each C_θ-equivalence class. Recall that, for an experiment $\mathcal{M}$ and any choice correspondence C on $\mathcal{M}$, $\mathcal{Q}^C(\mathcal{M})$ is the quotient set corresponding to the equivalence relation $\sim^C_\mathcal{M}$.

Proposition 4. *Let $\theta = (\mathcal{M}, P)$ be a stochastic choice dataset with $\mathcal{Q}^{C_\theta}(\mathcal{M}) = \{Q_1, \dots, Q_m\}$ and $v \in \mathcal{V}(\theta)$. Then, $w \in \mathcal{V}(\theta)$ if and only if there exists a collection of strictly positive scalars $(\lambda_1, \dots, \lambda_m)$ such that, for $i = 1, \dots, m$,*

$$v(a) = \lambda_i w(a) \quad \forall\, a \in Q_i.$$

As a result, there are experiments where, for any Luce rationalizable dataset, it is not possible to compare the utilities for some of the alternatives. When such comparisons are of economic interest, the ability to identify (up to rescaling) the utilities of all alternatives may therefore also factor into an experimental design.

Definition 4. Let $\mathcal{M}$ be an experiment and C a choice correspondence on $\mathcal{M}$. Then, $\mathcal{M}$ is *C-Luce identified* if, for any $P \in \mathcal{L}^C(\mathcal{M})$, whenever $v, w \in \mathcal{V}(\mathcal{M}, P)$, there exists $\lambda > 0$ such that $v(a) = \lambda w(a)$ for all $a \in \mathcal{A}^C(\mathcal{M})$.

The following corollary characterizes Luce identified experiments for an arbitrary choice correspondence.

Corollary 4. *Experiment $\mathcal{M}$ is C-Luce-identified if and only if all alternatives are C-connected, i.e, $a \sim^C_{\mathcal{M}} b$ for all $a, b \in \mathcal{A}^C(\mathcal{M})$.*

Since alternatives can be connected by overlapping sequences that are not part of elementary cycles, the conditions that allow for identification are distinct from the conditions that allow for falsification. Fixing a choice correspondence C, Proposition 4 and Corollary 4 show that the identification for an experiment $\mathcal{M}$ depends on the partition $\mathcal{Q}^C(\mathcal{M})$, with C-Luce identification being achieved when $\mathcal{Q}^C(\mathcal{M}) = \{\mathcal{A}^C(\mathcal{M})\}$. In contrast, falsifiable implications of the Luce Choice Rule depend on set of elementary C-cycles $\mathcal{E}^C(\mathcal{M})$ and can often be described in terms of the independent C-partition $\{\mathcal{M}_1, \dots, \mathcal{M}_K\}$ for the experiment $\mathcal{M}$.

To illustrate, consider the experiments in Example 2 (Figure 2). When C is the identity, $\mathcal{Q}^C(\mathcal{M}) = \{\mathcal{A}^C(\mathcal{M})\}$ in all three experiments and so these experiments are Luce identified. However, in the experiment on the left of Figure 2, $|\mathcal{M}_k| = 1$ for all k in the independent C-partition: the experiment is Luce unfalsifiable because no menu is part of an elementary cycle. In contrast, the experiments in the center and the right of Figure 1 are also Luce identified but $|\mathcal{M}_k| > 1$ for all k in the independent C-partition: every menu belongs to an elementary cycle. By adding some additional menus, these experiments achieve identification together with conditions that allow for a falsification of the Luce Choice Rule, which now has testable implications for every menu.

The experiment on the right of Figure 2 has the additional property that the independent C-partition is $\{\mathcal{M}\}$; that is the experiment cannot be further divided into separate parts that Luce independent of one another. By Proposition 3, the stochastic choice data on any subset of menus in $\mathcal{M}$ can be used to Luce falsify stochastic choice on the remaining menus. As such, the Luce Choice Rule has more testable implications than for the experiment in the center of Figure 2, where the independent C-partition has three distinct equivalence classes. However, in terms of an experimental design, these additional testable restrictions come at the cost of collecting data for an additional set of menus. In contrast, the experiment in the center of Figure 2 is Luce identified, has testable implications of the Luce Choice Rule on every menu, and has the additional property that, if any of the menus are removed, either the experiment is no longer Luce identified or some of the remaining menus are no longer part of an elementary cycle. In that sense, the experiment in the center is efficient because any smaller experiment either fails to identify unique utilities (up to rescaling) for all of the alternatives or has an independent C-partition with singleton menus that do not contribute data to falsify the Luce Choice Rule.

These examples highlight the importance of considering both identification and falsification as part of an experimental design. As a final illustration, consider the experiment from Example 1 (Figure 1, left) but without the menu $\{b, d\}$. The discussion in the introduction applies also when the menu $\{b, d\}$ is removed; that is, the Luce Choice Axiom and Luce Product Rule are vacuous for this experiment. However, for an experimental design, this experiment actually has desirable properties. First, note that (when C is the identity) the experiment is Luce identified because all alternatives are connected by overlapping sequences (as with all the experiments in Figure 2). Second, the independent C-partition contains only one equivalence class, namely the entire experiment. That means, every subset of menus can Luce falsify choice on the remaining menus (as in Figure 2, right). However, the experiment is also efficient: if any subset of menus is removed, either the new experiment is not Luce identified or has an independent C-partition containing singleton menus that do not contribute conditions to falsify the Luce Choice Rule (as in Figure 2, center). In contrast, the experiment that includes menu $\{b, d\}$ is Luce identified, cannot be divided into Luce independent parts, but is not efficient because there is a smaller experiment (for instance, where menu $\{b, d\}$ is removed) that is also Luce identified and where every menu contributes data that can falsify the Luce Choice Rule.

4.3 Prediction

When practical considerations dictate a restriction of the menus, a third criterion that may factor into an experimental design is the power to predict Luce rationalizable choice behavior on menus that are *not* in the sample for which data will be collected. Suppose, again, that $\mathcal{M}_1$ and $\mathcal{M}_2$ are two experiments, and the experimenter wants to use choice data from experiment $\mathcal{M}_1$ to predict behavior in experiment $\mathcal{M}_2$. For any Luce rationalizable SCF $P_1 \in \mathcal{L}(\mathcal{M}_1)$, there is an extension to $\mathcal{M}_1 \cup \mathcal{M}_2$ that is Luce rationalizable.[4] However, we are interested in cases where this extension is *unique*: if the experimenter is interested in choice behavior on a collection of menus $\mathcal{M}$, for which subset of menus $\mathcal{M}' \subseteq \mathcal{M}$ must data be collected in order to uniquely predict Luce choice behavior on the remaining menus in $\mathcal{M} \backslash \mathcal{M}'$.

[4]Suppose $v \in \mathcal{V}(\mathcal{M}_1, P_1)$ and define $w : \mathcal{A}(\mathcal{M}_1 \cup \mathcal{M}_2) \to \mathbb{R}_{++}$ by $w(a) = v(a)$ for all $a \in \mathcal{A}(\mathcal{M}_1)$ and $w(b) = 1$ for all $b \in \mathcal{A}(\mathcal{M}_1 \cup \mathcal{M}_2) \backslash \mathcal{A}(\mathcal{M}_1)$; then P_w is a Luce rationalizable extension of P_1 to $\mathcal{M}_1 \cup \mathcal{M}_2$.

Definition 5. Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be disjoint experiments and C be a choice correspondence on $\mathcal{M}_1 \cup \mathcal{M}_2$ such that $C(\mathcal{M}_2)$ contains no singleton menus.[5] Then, $\mathcal{M}_1$ can *C-Luce predict* choice on $\mathcal{M}_2$ if $|\mathcal{L}^C(\mathcal{M}_1 \cup \mathcal{M}_2 | P)| = 1$ for some $P \in \mathcal{L}^C(\mathcal{M}_1)$.

The following proposition characterizes when Luce rationalizable choice from one experiment can predict Luce choice in another experiment.

Proposition 5. *The experiment $\mathcal{M}_1$ can C-Luce predict choice on $\mathcal{M}_2$ if and only if the range of $\mathcal{M}_1$ contains the range of $\mathcal{M}_2$ and any two alternatives that are C-connected in $\mathcal{M}_2$ are already C-connected in $\mathcal{M}_1$; that is $\mathcal{A}^C(\mathcal{M}_2) \subseteq \mathcal{A}^C(\mathcal{M}_1)$ and, for all $a, b \in \mathcal{A}^C(\mathcal{M}_2)$, if $a \sim^C_{\mathcal{M}_2} b$, then $a \sim^C_{\mathcal{M}_1} b$.*

Analogously to falsification, Proposition 5 shows that, for a given choice correspondence, prediction is a feature of the experiments and not of the stochastic choice functions because it depends only on the structural relation between the experiments. That is, if $\mathcal{M}_1$ can C-Luce predict choice on $\mathcal{M}_2$, then $|\mathcal{L}^C(\mathcal{M}_1 \cup \mathcal{M}_2 | P)| = 1$ for *any* $P \in \mathcal{L}^C(\mathcal{M}_1)$. First, it is only possible to predict Luce choice behavior on $\mathcal{M}_2$ if its range is contained in the range of $\mathcal{M}_1$, since stochastic choice data from $\mathcal{M}_1$ imposes no restrictions on the relative frequency of alternatives outside its range. Second, *any* two alternatives that are connected in $\mathcal{M}_2$ must already be connected in $\mathcal{M}_1$; that is, each C-equivalence classes on $\mathcal{M}_2$ is contained in a C-equivalence class on $\mathcal{M}_1$. In that case, each C-equivalence class in $\mathcal{M}_1$ identifies a utility (up to rescaling) for the Luce Choice Rule which determines uniquely the Luce rationalizable stochastic choices on $\mathcal{M}_2$. That does not mean that the experiment $\mathcal{M}_2$ does not yield potentially useful choice information because, of course, the experiment $\mathcal{M}_2$ may still Luce falsify the stochastic choice dataset from $\mathcal{M}_1$.

Example 4. Consider the two experiments in Figure 4 (left), i.e. experiments $\mathcal{M}_1 = \{\{a, b\}, \{c, d\}\}$ and $\mathcal{M}_2 = \{\{a, c\}, \{b, d\}\}$. These experiments have the same range. However, even assuming that all alternatives are chosen with positive probability, none of the experiments can be used to predict the choice frequencies in the other, since there is no relation between connections in the experiments. In contrast, in the example in the center of the figure, experiment $\mathcal{M}_1 = \{\{a, b\}, \{c, d\}, \{b, d\}\}$ can predict choice frequencies in the second experiment $\mathcal{M}_2 = \{\{a, c\}\}$, but not vice versa. Last, in the example on the right of the figure, experiment $\mathcal{M}_1 = \{\{a, b\}, \{b, d\}, \{c, d\}, \{a, c\}\}$ and $\mathcal{M}_2 = \{\{a, b\}, \{b, c\}, \{c, d\}, \{a, d\}\}$ can predict

[5] We impose the restriction that $C(\mathcal{M}_2)$ does not contain singleton menus to avoid distinguishing cases later and because, for a singleton menu, the choice probability is always 1 and so there is nothing to predict.

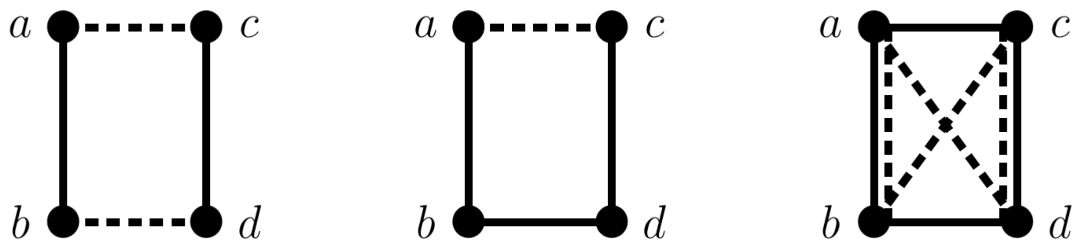

Figure 4: Experiments discussed in Example 4.

Luce choice in each other. That is, an experimenter only needs to collect choice data for one of these experiments in order to make a unique prediction for Luce choice on the other. □

5 Related Literature

The Luce Choice Rule is a fundamental building block of the discrete choice literature (McFadden, 1974; Anderson et al., 1992), and the empirical literature is too extensive to review. Reflecting the importance of the model for applied work, a large theoretical literature has examined extensions of the Luce Choice Rule and their axiomatic properties.

The seminal characterization in Luce (1959) and Luce and Suppes (1965) both require a rich domain, that is, they do not allow for an arbitrary collection of menus. Further, they both require Positivity (the requirement that all choice probabilities are strictly positive), and hence do not allow for censored alternatives. However, the original Luce Choice Axiom was formulated including a condition for the case of zero probabilities (Luce, 1959, p. 6), which obviously becomes void with Positivity. Several works have characterized explicit extensions of the Luce rule where Positivity is dropped, hence giving rise to models with censored alternatives (usually called censored Luce or censored logit models), where zero probabilities are allowed. An interpretation of those extensions is that a first decision stage eliminates certain alternatives and a second decision stage corresponds to a Luce Choice Rule over a restricted set of alternatives.[6]

[6] An early application of this idea is the random-demand model of McCausland (2009), which considers a specific domain where alternatives are consumption bundles and those which provide relatively less of all goods (and strictly less of some good) are eliminated in a first stage.

In this spirit, Ahumada and Ülkü (2018) study *Luce rules with limited consideration* where all menus are available, and characterize it through a single axiom (Axiom 1). This axiom is equivalent to the GPR when all menus are available, thus the characterization (Ahumada and Ülkü, 2018, Theorem 1) can be derived from Theorem 1, as mentioned in Section 3.1. Echenique and Saito (2019) also characterize censored Luce models where all menus are available. Their *cyclical independence* axiom is equivalent to the GPR for this case, and thus their characterization (Echenique and Saito, 2019, Theorem 1) can also be derived from Theorem 1. These authors further study *threshold models,* a particular class of censored Luce rules where alternatives become censored when their relative utility falls below a fixed threshold. Horan (2021) studies a *lexicographic choice rule* with the same interpretation (and also requiring all menus to be available), but also asks when the first-stage selection can be rationalized by a binary relation. The resulting *stochastic semi-order model* generalizes the threshold models of Echenique and Saito (2019). For characterization purposes, however, the *Strong Product Rule* used in Horan (2021) is again equivalent to the GPR if all menus are available.

Also related to this approach is Doğan and Yıldız (2021), which considers a new axiom called *odds supermodularity,* requiring that the odds against an existing alternative increase at least additively as new alternatives are added to the choice set. This axiom characterizes *preference-oriented Luce rules,* which are a particular case of censored Luce rules where the first-stage eliminates alternatives which do not maximize a given, fixed preference relation. Again, this result requires that all menus are available. Of course, if censoring is dropped, the original Luce Choice Rule is obtained, that is, odds supermodularity and Positivity together provide an alternative characterization of the Luce Choice Rule (Echenique and Saito, 2019, Corollary 1). Relating directly to the revealed preference literature and to two-stage Luce models, Cerreia-Vioglio et al. (2021) show that, in a setting where all menus are available but Positivity is dropped, a random choice rule satisfies Luce's Choice Axiom if and only if the choice correspondence defined by its support satisfies the Weak Axiom of Revealed Preference and random choice then follows by a tie breaking rule that satisfies Rényi's (1955) Conditioning Axiom.

The recent literature has also examined other generalizations of the Luce Choice Rule, which can accommodate received behavioral anomalies, but has done so maintaining the assumption that all or at least a rich class of menus are available. Gul et al. (2014) study extensions of the Luce Choice Rule that allow for violations of the weak axiom of revealed preference. They consider *rich choice rules* where the set of available menus is not necessarily the collection of all possible menus, but it is still “rich” in the sense that the choice probability for a given menu as a subset of other menus can be varied continuously, even when the superset menus are required to exclude another, fixed menu. In particular, this implies an infinite number of available menus. A preliminary result in this work (Gul et al., 2014, Theorem 1) identifies a condition which characterizes the Luce Choice Rule within the class of rich choice rules. Faro (2023) consider discrete choice models where the set of alternatives might contain replicas. The objective is to address the well-known *duplicates problem* (Debreu, 1960) (an issue also addressed by Gul et al., 2014). This work provides an axiomatic characterization of the Luce model in a context with replicas, maintaining both Positivity and that all menus are available.

Other work has considered generalizations of the Luce Choice Rule capturing attentional and perceptual phenomena, with the aim of encompassing well-known behavioral anomalies. The *perception-adjusted Luce models* of Echenique et al. (2018) include a perception order, which allows for violations of independence of irrelevant alternatives. Their characterization assumes Positivity and that all binary and ternary menus are available. Relatedly, Tserenjigmid (2021) considers *order-dependent Luce models*, where the utility of alternatives depends on their relative ordering in the menu. These fulfill Positivity and are characterized by weakening the IIA and the Luce Product Rule. The characterization assumes that all menus are available. Kovach and Tserenjigmid (2022b) consider *focal Luce models* including a bias toward a menu-dependent set of focal alternatives. The model assumes that all binary menus are available. Focal Luce models are characterized by Positivity, the Luce Product Rule, and a modification of IIA. Heydari (2021) considers *random arbitration rules*, which extend the Luce Choice Rule when alternatives can be ranked along multiple attributes, specifying that choices depend on menu-specific reference points. Alternatives which are dominated attribute-wise in a menu are never chosen, violating Positivity. However, the model assumes that all menus are available.

Recently, Kovach and Tserenjigmid (2022a) have provided characterizations of *nested logit*, a widely-used generalization of logit choice which also allows for violations of IIA as the similarity effect (or the existence of duplicates; Debreu, 1960). Nested logit amounts to a two-step logit procedure, where first a set or *nest* of alternatives is chosen from a predetermined collection of such nests, and then an alternative from the chosen nest is itself chosen. The characterization of Kovach and Tserenjigmid (2022a) assumes Positivity and that all menus are available.

Other axiomatic work on the Luce Choice Rule has looked at more fine-grained characterizations which pin down not only the logit functional form, but also the characteristics of the utility functional form (in a multi-attribute setting). For example, Ahn et al. (2018) characterize the logit model with linear utility when only average population choices are observable. Breitmoser (2021) has studied the general axiomatic foundations of *conditional logit*, where the utilities are specified ex ante rather than being part of the characterization. More recently, Cerreia-Vioglio et al. (2023) have characterized a kind of logit choice rules where the probability-generating utility depends on time constraints and initial biases, allowing a dependence of choice probabilities on externally-imposed deadlines and initial anchors. Although the technical assumptions on underling choice spaces vary, these works maintain Positivity assumptions and the availability of all (finite) menus.

Lastly, although most work on the Luce Choice Rule has concentrated on static representations, some contributions have also looked at explicitly dynamic frameworks. For example, Gul et al. (2014) considered the extension of the Luce Choice Rule to dynamic problems, and Fudenberg and Strzalecki (2015) provided an axiomatic characterization of a *Discounted Adjusted Luce Model* in an intertemporal setting where the decision maker chooses from a menu of actions yielding an outcome for the current period and a menu of actions for the next one.

In summary, most of the characterizations of choice rules related to or generalizing the Luce Choice Rule maintain both Positivity and a rich set of menus. Some contributions weaken Positivity (allowing censored alternatives), but none of them addresses the problem of arbitrarily restricted menus. Most of the literature has imposed additional structure on the Luce Choice Rule or its generalization with censored alternatives, while essentially maintaining a framework with a rich set of menus. In contrast, we generalize the Luce Choice Axiom/Luce Product Rule in order to provide a characterization of the censored Luce Choice Rule that applies for any collection of menus.

Last, we remark that the fact that obtaining data for all choice menus is impractical has led to alternative approaches which abstract from questions of falsification and concentrate instead on fitting of models to data. In particular, *conjoint analysis* introduces attribute variation for choice problems across individuals, especially survey designs. This approach goes back to Luce and Tukey (1964) and is extensively used in marketing (Green and Rao, 1971; Hainmueller et al., 2014). The new tools for experimental and survey design that we provide can be used as a complement to conjoint analysis to, e.g., ensure falsifiability while improving the empirical fit of models and data.

6 Conclusion

The Luce Choice Rule, or multinomial logit model, is widely used in economics and beyond to model choice data. However, empirical work deals with limited datasets. For practical (e.g., cost) considerations, datasets rarely contain all potential menus, or even all binary menus. Yet, all previous characterizations of logit choice rest upon the assumption that data for a rich set of menus is observed. We provide a characterization of the Luce Choice Rule that allows for an arbitrary collection of menus, and hence can be applied to any actual stochastic choice dataset. Further, extending previous work, our characterization applies when not all alternatives are observed with positive probability, as will often be the case in actual datasets.

When an empirical researcher designs an experiment or survey, the objectives of research must be balanced against the practical limitations. Building upon our characterization, we provide further results that show how to determine whether an experiment will fulfill three criteria. The first is falsification, i.e. whether the data will allow for potentially falsifying the Luce Choice Rule. The second is identification, i.e. whether it will be possible to uniquely pin down a utility function underlying the Luce Choice Rule. The third is prediction, i.e. being able to predict choice frequencies for out of sample menus.

Our contribution is therefore twofold. On one hand, we provide a characterization of Luce choice that applies independently of the practical dataset limitations that are to be expected in empirical work. On the other hand, we provide tools that empirical researchers can use, before data collection, to ensure that data-collection designs efficiently accomplish some common research objectives.

References

Ahn, D. S., F. Echenique, and K. Saito (2018): "On Path Independent Stochastic Choice," *Theoretical Economics*, 13, 61–85.

Ahumada, A. and L. Ülkü (2018): "Luce Rule with Limited Consideration," *Mathematical Social Sciences*, 93, 52–56.

Alós-Ferrer, C., E. Fehr, and N. Netzer (2021): "Time Will Tell: Recovering Preferences when Choices are Noisy," *Journal of Political Economy*, 129, 1828–1877.

Alós-Ferrer, C. and M. Garagnani (2024): "Improving Risky-Choice Predictions Using Response Times," *Journal of Political Economy: Microeconomics*, 2(2), 335–354.

Anderson, S. P., J.-F. Thisse, and A. De Palma (1992): *Discrete Choice Theory of Product Differentiation*, Cambridge, MA: MIT Press.

Breitmoser, Y. (2021): "An Axiomatic Foundation of Conditional Logit," *Economic Theory*, 72, 245–261.

Cerreia-Vioglio, S., P. O. Lindberg, F. Maccheroni, M. Marinacci, and A. Rustichini (2021): "A Canon of Probabilistic Rationality," *Journal of Economic Theory*, 196.

Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and A. Rustichini (2023): "Multinomial Logit Processes and Preference Discovery: Inside and Outside the Black Box," *Review of Economic Studies*, 90, 1155–1194.

Clithero, J. A. (2018): "Improving Out-of-Sample Predictions Using Response Times and a Model of the Decision Process," *Journal of Economic Behavior and Organization*, 148, 344–375.

Davis-Stober, C. P., N. Brown, and D. R. Cavagnaro (2015): "Individual Differences in the Algebraic Structure of Preferences," *Journal of Mathematical Psychology*, 66, 70–82.

Debreu, G. (1960): "Review of 'Individual Choice Behavior: A Theoretical Analysis' by R. D. Luce." *American Economic Review*, 50, 186–188.

Doğan, S. and K. Yildiz (2021): "Odds Supermodularity and the Luce Rule," *Games and Economic Behavior*, 126, 443–452.

Echenique, F. and K. Saito (2019): "General Luce Model," *Economic Theory*, 68, 811–826.

Echenique, F., K. Saito, and G. Tserenjigmid (2018): "The Perception-Adjusted Luce Model," *Mathematical Social Sciences*, 93, 67–76.

Faro, J. H. (2023): "The Luce Model With Replicas," *Journal of Economic Theory*, 208, 105596.

Fudenberg, D., P. Strack, and T. Strzalecki (2018): "Speed, Accuracy, and the Optimal Timing of Choices," *American Economic Review*, 108, 3651–3684.

Fudenberg, D. and T. Strzalecki (2015): "Dynamic Logit with Choice Aversion," *Econometrica*, 83, 651–691.

Goeree, J. K., C. A. Holt, and T. R. Palfrey (2005): "Regular Quantal Response Equilibrium," *Experimental Economics*, 8, 347–367.

Green, P. E. and V. R. Rao (1971): "Conjoint Measurement for Quantifying Judgmental Data," *Journal of Marketing Research*, 8, 355–363.

Gul, F., P. Natenzon, and W. Pesendorfer (2014): "Random Choice as Behavioral Optimization," *Econometrica*, 82, 1873–1912.

Hainmueller, J., D. J. Hopkins, and T. Yamamoto (2014): "Causal Inference in Conjoint Analysis: Understanding Multidimensional Choices via Stated Preference Experiments," *Political Analysis*, 22, 1–30.

Heydari, P. (2021): "Luce Arbitrates: Stochastic Resolution of Inner Conflicts," *Games and Economic Behavior*, 126, 33–74.

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

Kovach, M. and G. Tserenjigmid (2022a): "Behavioral Foundations of Nested Stochastic Choice and Nested Logit," *Journal of Political Economy*, 130, 2411–2461.

——— (2022b): "The Focal Luce Model," *American Economic Journal: Microeconomics*, 14, 378–413.

Luce, R. D. (1959): *Individual Choice Behavior: A Theoretical Analysis*, New York: Wiley.

Luce, R. D. and P. Suppes (1965): "Preference, Utility and Subjective Probability," in *Handbook of Mathematical Psychology, Vol. 3*, ed. by R. D. Luce, R. R. Bush, and E. Galanter, New York: John Wiley & Sons, 249–410.

Luce, R. D. and J. W. Tukey (1964): "Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement," *Journal of Mathematical Psychology*, 1, 1–27.

Matejka, F. and A. McKay (2015): "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *American Economic Review*, 105, 272–98.

McCausland, W. J. (2009): "Random Consumer Demand," *Economica*, 76, 89–107.

McFadden, D. L. (1974): "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics*, ed. by P. Zarembka, New York: Academic Press, 105–142.

——— (2001): "Economic Choices," *American Economic Review*, 91, 351–378.

McKelvey, R. D. and T. R. Palfrey (1995): "Quantal Response Equilibria for Normal Form Games," *Games and Economic Behavior*, 10, 6–38.

——— (1998): "Quantal Response Equilibria for Extensive Form Games," *Experimental Economics*, 1, 9–41.

Moffatt, P. G. (2015): *Experimetrics: Econometrics for Experimental Economics*, London: Palgrave Macmillan.

Ratcliff, R. (1978): "A Theory of Memory Retrieval," *Psychological Review*, 85, 59–108.

Ratcliff, R. and J. N. Rouder (1998): "Modeling Response Times for Two-Choice Decisions," *Psychological Science*, 9, 347–356.

Rényi, A. (1955): "On a New Axiomatic Theory of Probability," *Acta Mathematica Academiae Scientiarum Hungaricae*, 6, 285–335.

Sims, C. A. (2003): "Implications of Rational Inattention," *Journal of Monetary Economics*, 50, 665–690.

Suppes, P., D. H. Krantz, R. D. Luce, and A. Tversky (1989): *Foundations of Measurement: Geometrical, Threshold, and Probabilistic Representations*, vol. 2, San Diego, California: Academic Press.

Tserenjigmid, G. (2021): "The Order-Dependent Luce Model," *Management Science*, 67, 6915–6933.

Appendix

We first establish two lemmata that are used in the proof of Theorem 1 and then provide proofs for the remaining results.

The following lemma establishes some properties of the relation $\sim^C_{\mathcal{M}}$.

Lemma 1. *Suppose $\mathcal{M}$ is an experiment and C is a choice correspondence on $\mathcal{M}$.*

(i) The binary relation $\sim^C_{\mathcal{M}}$ is an equivalence on $\mathcal{A}^C(\mathcal{M})$.

(ii) For each $Q \in \mathcal{Q}^C(\mathcal{M})$ and $M \in \mathcal{M}$, either $C(M) \subseteq Q$ or $C(M) \cap Q = \emptyset$.

(iii) For each $Q \in \mathcal{Q}^C(\mathcal{M})$ with $|Q| \geq 2$ there is an overlapping sequence $[a_1, \ldots, a_{n+1}; M_1, \ldots, M_n]$ on $C(\mathcal{M})$ such that $\bigcup_{i=1}^n C(M_i) = Q$.

Proof. For part (i), if $a \in \mathcal{A}^C(M)$, then $a \in C(M)$ for some $M \in \mathcal{M}$, and so $[a, a; M]$ is an overlapping sequence in $C(\mathcal{M})$. Hence, $a \sim^C_{\mathcal{M}} a$. If $a \sim^C_{\mathcal{M}} b$ for $a \neq b$, then there is an overlapping C-sequence $[a_1, \ldots, a_{n+1}; M_1, \ldots, M_n]$ with $a = a_1$ and $b = a_{n+1}$. Letting $a'_i = a_{n+1-i}$ and $M'_i = M_{n+1-i}$ for $i = 1, \ldots, n$, $[a'_1, \ldots, a'_{n+1}; M'_1, \ldots, M'_n]$ is an overlapping C-sequence with $b = a'_1$ and $a = a'_{n+1}$, and so $b \sim^C_{\mathcal{M}} a$. Now consider distinct $a, b, d \in \mathcal{A}^C(\mathcal{M})$ with $a \sim^C_{\mathcal{M}} b$ and $b \sim^C_{\mathcal{M}} d$, i.e., there are overlapping C-sequences $[a_1, \ldots, a_{m+1}; M_1, \ldots, M_m]$ and $[b_1, \ldots, b_{k+1}; N_1, \ldots, N_k]$ with $a = a_1$, $b = a_{m+1} = b_1$ and $d = b_{k+1}$. Then $[a_1, \ldots, a_m, b_1 \ldots, b_{k+1}; M_1, \ldots M_m, N_1, \ldots N_k]$ is an overlapping C-sequence with $a = a_1$ and $c = b_{k+1}$, and so $a \sim^C_{\mathcal{M}} d$. Hence, $\sim^C_{\mathcal{M}}$ is reflexive, symmetric and transitive, thus an equivalence relation.

Part (ii) follows because if $a \in C(M) \cap Q$ and $b \in C(M)$, then $[a, b; M]$ is an overlapping C-sequence, implying that $a \sim^C_{\mathcal{M}} b$. Hence $b \in Q$, and so $C(M) \subseteq Q$.

For part (iii), let $Q \in \mathcal{Q}^C(\mathcal{M})$ and consider any distinct $a, b \in Q$. Since $a \sim^C_{\mathcal{M}} b$, there is an overlapping C-sequence $[a_1, \ldots, a_{n+1}; M_1, \ldots, M_n]$ with $a = a_1$ and $b = a_{n+1}$. It follows from part (ii) that $\bigcup_{i=1}^n C(M_i) \subseteq Q$. If equality holds, the argument is complete. If not, there is $d \in Q \setminus \bigcup_{i=1}^n C(M_i)$, and the argument for transitivity in part (i) shows that there is an overlapping C-sequence $[a'_1, \ldots, a'_{m+1}; M'_1, \ldots, M'_m]$ such that $\{a, b, d\} \subseteq \bigcup_{j=1}^m C(M'_j) \subseteq Q$. Since all alternatives in Q are related by $\sim^C_{\mathcal{M}}$, repeating this argument eventually yields $Q \subseteq \bigcup_{j=1}^m C(M'_j) \subseteq Q$. □

The following lemma shows that a stochastic choice dataset is Luce rationalizable if and only if there is an independent Luce Choice Rule for each equivalence class on the experiment.

Lemma 2. *A stochastic choice dataset $\theta = (\mathcal{M}, P)$ can be rationalized by a Luce Choice Rule if and only if there exists a collection of utility functions $\big(v_i : Q_i \to \mathbb{R}_{++}\big)_{i=1}^m$, one for each equivalence class $Q_i \in \mathcal{Q}^{C_\theta}(\mathcal{M})$, such that, for $i = 1, \ldots, m$,*

$$P(a, M) = \begin{cases} \frac{v_i(a)}{\sum_{b \in C_\theta(M)} v_i(b)} & \text{if } a \in C_\theta(M) \\ 0 & \text{if } a \notin C_\theta(M) \end{cases} \tag{3}$$

whenever $(a, M) \in \mathcal{S}(\mathcal{M})$, $C_\theta(M) \subseteq Q_i$, where m is the number of equivalence classes in $\mathcal{Q}^{C_\theta}(\mathcal{M})$. Moreover, in that case, θ has a Luce representation with a utility function $v : \mathcal{A}(\mathcal{M}) \to \mathbb{R}_{++}$ given by $v(a) = v_i(a)$ for all $a \in Q_i$ and $i = 1, \ldots, m$ and $v(b)$ arbitrary for $b \in \mathcal{A}(\mathcal{M}) \setminus \mathcal{A}^{C_\theta}(\mathcal{M})$.

Proof. Suppose θ can be rationalized by a Luce Choice Rule: there exists $v : \mathcal{A}(\mathcal{M}) \to \mathbb{R}_{++}$ such that (1) is satisfied whenever $a \in M \in \mathcal{M}$. For $i = 1, \ldots, m$, define $v_i : Q_i \to \mathbb{R}_{++}$ by $v_i(a) = v(a)$. Then (3) is satisfied whenever $a_i \in M \subseteq Q_i$.

The converse follows because the equivalence relation $\sim_{\mathcal{M}}^{C_\theta}$ on alternatives also induces a partition of the menus by Lemma 1. For each equivalence class $Q_i \in \mathcal{Q}^{C_\theta}(\mathcal{M})$, let $\mathcal{M}_i = \{M \in \mathcal{M} \mid C_\theta(M) \subseteq Q_i\}$. Then $\mathcal{M}_i \cap \mathcal{M}_j = \emptyset$ for all $i \neq j$, and it follows by Lemma 1(ii) that $\bigcup_{i=1}^m \mathcal{M}_i$ contains the supports of all the relevant menus. Now suppose there is a collection of utility functions $\big(v_i : Q_i \to \mathbb{R}_{++}\big)_{i=1}^m$ such that (3) holds whenever $a \in C_\theta(M) \subseteq Q_i$. For $i = 1, \ldots, m$, let $v(a) = v_i(a)$ if $a \in Q_i$, and let $v(b)$ be arbitrary when $b \in \mathcal{A}(\mathcal{M}) \setminus \mathcal{A}^{C_\theta}(\mathcal{M})$. Then, $v : \mathcal{A}(\mathcal{M}) \to \mathbb{R}_{++}$ satisfies (1) whenever $a \in M \in \mathcal{M}$. □

Proof of Proposition 1

Proof. Let $\theta = (\mathcal{M}, P)$ be a stochastic choice dataset. The GPR trivially implies that Equation (2) holds for every elementary C_θ-cycle on $\mathcal{M}$. To see the converse, suppose that Equation 2 is satisfied for every elementary cycle in $\mathcal{E}^{C_\theta}(\mathcal{M})$. Now consider an arbitrary non-degenerate overlapping C_θ-cycle on $\mathcal{M}$, denoted $[a_1, \ldots, a_{n+1}; M_1, \ldots, M_n]$ with $a_{n+1} = a_1$. First, suppose only one alternative is repeated, $a_j = a_1$ (w.l.o.g). Then, consider alternatives $a_1, \ldots, a_{j-1}$ and menus $M_1, \ldots, M_{j-1}$. We have that $a_i, a_{i+1} \in C_\theta(M_i)$ for $i = 1, \ldots, j-2$, $a_{j-1} \in C_\theta(M_{j-1})$, and $a_1 = a_j \in C_\theta(M_{j-1})$. Hence, by (2) for elementary C_θ-cycles,

$$\prod_{i=1}^{j-1} P(a_i, M_i) = \prod_{i=1}^{j-2} P(a_{i+1}, M_i) \cdot P(a_1, M_{j-1}).$$

Now consider alternatives $a_j = a_1, a_{j+1}, \ldots, a_n$. We have that $a_i, a_{i+1} \in C_\theta(M_i)$ for $i = j, \ldots, n-1$, $a_n \in C_\theta(M_n)$, and $a_{n+1} = a_1 = a_j \in C_\theta(M_n)$. Hence, by (2) for elementary C_θ-cycles,

$$\prod_{i=j}^{n} P(a_i, M_i) = \prod_{i=j}^{n-1} P(a_{i+1}, M_i) \cdot P(a_1, M_n)$$

Thus,

$$\prod_{i=1}^{n} P(a_i, M_i) = \prod_{i=1}^{j-1} P(a_i, M_i) \prod_{i=j}^{n} P(a_i, M_i)$$
$$= \prod_{i=1}^{j-2} P(a_{i+1}, M_i) \cdot P(a_1, M_{j-1}) \cdot \prod_{i=j}^{n-1} P(a_{i+1}, M_i) \cdot P(a_1, M_n) = \prod_{i=1}^{n} P(a_{i+1}, M_i)$$

where the last equality follows because $a_{n+1} = a_1 = a_j$. The GPR then follows by induction on the number of repeated alternatives. □

Proof of Proposition 2

Proof. First, suppose that $\mathcal{E}^C(\mathcal{M}_1 \cup \mathcal{M}_2) = \mathcal{E}^C(\mathcal{M}_1) \cup \mathcal{E}^C(\mathcal{M}_2)$. Let $P \in \mathcal{P}^C(\mathcal{M}_1 \cup \mathcal{M}_2)$ and let $\theta = (\mathcal{M}_1 \cup \mathcal{M}_2, P)$. It is immediate that any utility that censored Luce rationalize θ also censored Luce rationalize $\theta_i = (\mathcal{M}_i, P_{\mathcal{M}_i})$ for $i = 1, 2$. Conversely, suppose that θ_i is C-Luce rationalizable for $i = 1, 2$. Then, by Theorem 1, Equation 2 is satisfied for every elementary C-cycle in $\mathcal{E}^C(\mathcal{M}_i)$ for $i = 1, 2$. Hence, because $\mathcal{E}^C(\mathcal{M}_1 \cup \mathcal{M}_2) = \mathcal{E}^C(\mathcal{M}_1 \cup \mathcal{M}_2) = \mathcal{E}^C(\mathcal{M}_1) \cup \mathcal{E}^C(\mathcal{M}_2)$, Equation 2 is satisfied for every elementary C-cycle in $\mathcal{E}^C(\mathcal{M}_1 \cup \mathcal{M}_2)$. Hence, by Proposition 1, θ is C-Luce rationalizable. Therefore, $\mathcal{M}_1$ and $\mathcal{M}_2$ are C-Luce independent.

Now suppose that $\mathcal{M}_1$ and $\mathcal{M}_2$ are C-Luce independent. We want to show that $\mathcal{E}^C(\mathcal{M}_1 \cup \mathcal{M}_2) = \mathcal{E}^C(\mathcal{M}_1) \cup \mathcal{E}^C(\mathcal{M}_2)$. By contradiction, suppose this is not true. Since $\mathcal{E}^C(\mathcal{M}_i) \subseteq \mathcal{E}^C(\mathcal{M}_1 \cup \mathcal{M}_2)$ for $i = 1, 2$, it follows that there exists some $\phi \in \mathcal{E}^C(\mathcal{M}_1 \cup \mathcal{M}_2) \setminus \big(\mathcal{E}^C(\mathcal{M}_1) \cup \mathcal{E}^C(\mathcal{M}_2)\big)$. Write $\phi = [a_1, \ldots, a_{n+1}; M_1, \ldots, M_n]$, $a_{n+1} = a_1$, $n \geq 2$. Since ϕ is neither an elementary C-cycle in $\mathcal{M}_1$ nor an elementary C-cycle in $\mathcal{M}_2$, some of the menus in the list $M_1, \ldots, M_n$ are in $\mathcal{M}_1$ and some are in $\mathcal{M}_2$. Hence, there exists $j^* \in \{1, \ldots, n\}$ such that $M_{j^*} \in \mathcal{M}_1$ and $M_{j^*-1} \in \mathcal{M}_2$ (modulo n) or vice versa (the symmetric case is identical). In particular, $a_{j^*} \in C(M_{j^*}) \cap C(M_{j^*-1}) \subseteq \mathcal{A}^C(\mathcal{M}_1) \cap \mathcal{A}^C(\mathcal{M}_2)$.

Define some $v_1 : \mathcal{A}(\mathcal{M}_1) \mapsto \mathbb{R}_{++}$ and then define $v_2 : \mathcal{A}(\mathcal{M}_2) \mapsto \mathbb{R}_+$ by

$$v_2(a) = \begin{cases} v_1(a_{j^*}) + 1 & \text{if } a = a_{j^*} \\ v_1(a) & \text{if } a \in \mathcal{A}^C(\mathcal{M}_1) \setminus \{a_{j^*}\} \\ 1 & \text{if } a \in \mathcal{A}^C(\mathcal{M}_2) \setminus \mathcal{A}^C(\mathcal{M}_1) \\ 0 & \text{otherwise} \end{cases}$$

Define a SCF P on $\mathcal{M}_1 \cup \mathcal{M}_2$ by $P(a, M) = P_{v_1}(a, M)$ for all $(a, M) \in \mathcal{S}(\mathcal{M}_1)$ and $P(a, M) = P_{v_2}(a, M)$ for all $(a, M) \in \mathcal{S}(\mathcal{M}_2)$. This is a well-defined SCF because $\mathcal{M}_1$ and $\mathcal{M}_2$ are disjoint. Then $(\mathcal{M}_i, P_{\mathcal{M}_i}) = (\mathcal{M}_i, P_{v_i})$, which is C-Luce rationalizable by construction. Since $\mathcal{M}_1$ and $\mathcal{M}_2$ are C-Luce independent, it follows that $P \in \mathcal{L}^C(\mathcal{M}_1 \cup \mathcal{M}_2)$. By Proposition 1, Equation (2) must hold for ϕ.

For each menu $M \in \mathcal{M}_1 \cup \mathcal{M}_2$, define $i(M_j) = 1$ if $M \in \mathcal{M}_1$ and $i(M_j) = 2$ if $M \in \mathcal{M}_2$. Define also $v(M) = \sum_{a \in M} v_{i(M_j)}(a)$. Again, this is well-defined because $\mathcal{M}_1$ and $\mathcal{M}_2$ are disjoint. Equation (2) for ϕ can be written as follows.

$$\prod_{j=1}^{n} \frac{v_{i(M_j)}(a_j)}{v(M_j)} = \prod_{j=1}^{n} \frac{v_{i(M_j)}(a_{j+1})}{v(M_j)} = \prod_{j=1}^{n} \frac{v_{i(M_{j-1})}(a_j)}{v(M_{j-1})},$$

where the subindices are modulo n. The product of the denominators is identical on both sides, hence the equation simplifies to

$$\prod_{j=1}^{n} v_{i(M_j)}(a_j) = \prod_{j=1}^{n} v_{i(M_{j-1})}(a_j).$$

Let $j \neq j^*$. If $a_j \in \mathcal{A}^C(\mathcal{M}_1) \cap \mathcal{A}^C(\mathcal{M}_2)$, then $v_1(a_j) = v_2(a_j)$ by construction of v_2. Hence the term for a_j is identical on both sides of the equality. If $a_j \in \mathcal{A}^C(\mathcal{M}_2) \setminus \mathcal{A}^C(\mathcal{M}_1)$, we must have that $i(M_j) = i(M_{j-1}) = 2$, since there is no menu in $\mathcal{M}_1$ containing a_j. Symmetrically, if $a_j \in \mathcal{A}^C(\mathcal{M}_1) \setminus \mathcal{A}^C(\mathcal{M}_2)$, we must have that $i(M_j) = i(M_{j-1}) = 1$, since there is no menu in $\mathcal{M}_2$ containing a_j. In both cases, again, the term for a_j is identical on both sides of the equality. Thus, the last equation reduces to

$$v_{i(M_{j^*})}(a_{j^*}) = v_{i(M_{j^*-1})}(a_{j^*}).$$

By the choice of a_{j^*}, $i(M_{j^*}) = 1$ and the left-hand side is equal to $v_1(a_{j^*})$, while $i(M_{j^*-1}) = 2$ and the right-hand side is equal to $v_2(a_{j^*}) = v_1(a_{j^*}) + 1$. This contradicts that $P \in \mathcal{L}^C(\mathcal{M}_1 \cup \mathcal{M}_2)$, and so $\mathcal{M}_1$ and $\mathcal{M}_2$ are not C-Luce independent. □

Proof of Corollary 2

With some abuse of notation, for an elementary C-cycle ϕ of length n, we write $M \in \phi$ if $M = M_i(\phi)$ for some $i \in \{1, \dots, n\}$.

Proof. Fix an experiment $\mathcal{M}$ and a choice correspondence C on $\mathcal{M}$. We construct the independent C-partition as follows. First let $\mathcal{M}' = \{M_1, ..., M_J\} = \big\{M \in \mathcal{M} \;\big|\; \forall\, \phi \in \mathcal{E}^C(\mathcal{M}), M \notin \phi\big\}$ be the set of all menus in $\mathcal{M}$ that are not part of any elementary C-cycle on $\mathcal{M}$. If $\mathcal{M}' = \mathcal{M}$, the independent C-partition is $\{\{M_j\} \mid j = 1, \dots, J\}$. Otherwise, for two menus $M, M' \in \mathcal{M}\backslash\mathcal{M}'$, write $M \sim_{\mathcal{E}} M'$ if there exist menus $N_1, \dots, N_n \in \mathcal{M}$ such that $M = N_1$, $M' = N_n$ and, for all $i = 1, \dots, n-1$, there exists some $\phi_i \in \mathcal{E}^C(\mathcal{M})$ such that $N_i, N_{i+1} \in \phi_i$. It is easily verified that $\sim_{\mathcal{E}}$ is an equivalence relation on $\mathcal{M}\backslash\mathcal{M}'$. Let $\mathcal{M}_{J+1}, \dots, \mathcal{M}_K$ be the corresponding equivalence classes. Then, $\{\mathcal{M}_1, \dots, \mathcal{M}_J, \mathcal{M}_{J+1}, \dots, \mathcal{M}_K\}$ is the independent C-partition.

By Proposition 2, to verify that any two equivalence classes in $\{\mathcal{M}_1, \dots, \mathcal{M}_K\}$ are C-Luce independent, it suffices to show that $\mathcal{E}^C(\mathcal{M}) = \bigcup_{k=1}^K \mathcal{E}^C(\mathcal{M}_k)$. Suppose not. Then, there exists $\phi \in \mathcal{E}^C(\mathcal{M})\backslash \bigcup_{k=1}^K \mathcal{E}^C(\mathcal{M}_k)$. Thus, there are menus $M, M' \in \phi$ such that $M \in \mathcal{M}_k$ and $M' \in \mathcal{M}_{k'}$ for some $k \neq k'$ (it must then be the case that $k, k' \in \{J+1, \dots, K\}$). But, since $M, M' \in \phi$, it follows that $M \sim_{\mathcal{E}} M'$, which contradicts that they are in different equivalence classes.

Finally, we need to verify that, for any $k \in \{J+1, \dots, K\}$, there does not exist $\mathcal{M}'_k \subsetneq \mathcal{M}_k$ such that $\mathcal{M}'_k$ and $\mathcal{M}_k\backslash\mathcal{M}'_k$ are C-Luce independent. Suppose this is not the case. Then, it follows by Proposition 2 that $\mathcal{E}^C(\mathcal{M}_k) = \mathcal{E}^C(\mathcal{M}'_k) \cup \mathcal{E}^C(\mathcal{M}_k\backslash\mathcal{M}'_k)$. Let $M \in \mathcal{M}'_k$ and $M' \in \mathcal{M}_k\backslash\mathcal{M}'_k$. Since $M \sim_{\mathcal{E}} M'$, there exists $N_1, \dots, N_n$ such that $M = N_1$, $M' = N_n$ and, for all $i = 1, \dots, n-1$, there exists some $\phi_i \in \mathcal{E}^C(\mathcal{M})$ such that $N_i, N_{i+1} \in \phi_i$. Since $M \in \mathcal{M}'_k$ and $M' \in \mathcal{M}_k\backslash\mathcal{M}'_k$, there exists some $i \in \{1, \dots, n-1\}$ such that $N_i \in \mathcal{M}'_k$, $N_{i+1} \in \mathcal{M}_k\backslash\mathcal{M}'_k$, and hence $\phi_i \in \mathcal{E}^C(\mathcal{M})$ with $N_i, N_{i+1} \in \phi_i$, implying that $\phi_i \in \mathcal{E}^C(\mathcal{M}_k)\backslash(\mathcal{E}^C(\mathcal{M}'_k) \cup \mathcal{E}^C(\mathcal{M}_k\backslash\mathcal{M}'_k))$, a contradiction. ☐

Proof of Corollary 3

Proof. If $\mathcal{E}^C(\mathcal{M}) = \emptyset$, by Proposition 1 the GPR is vacuously satisfied for any stochastic choice dataset $\theta = (\mathcal{M}, P)$. Thus, $\mathcal{M}$ is Luce unfalsifiable.

Conversely, suppose there is an elementary C-cycle ϕ on $\mathcal{M}$. Since ϕ is non-degenerate, there is an elementary C-cycle $\tilde{\phi} = [a_1, \dots, a_{n+1}; M_1, \dots, M_n]$ such that $M_1 \neq M_n$. Moreover, $|C(M_i)| \geq 2$ for each $i = 1, ..., n$. Let $P \in \mathcal{P}^C(\mathcal{M})$, and define P_ε by $P_\varepsilon(a, M) = P(a, M)$ if $M \neq M_1$, $P_\varepsilon(a_1, M_1) = 1 - \varepsilon$, $P_\varepsilon(b, M_1) =$

$\frac{\varepsilon}{|C(M_1)|-1}$ for $b \in C(M_1)\backslash\{a_1\}$, and $P(c, M_1) = 0$ for $c \in M_1\backslash C(M_1)$. Then, $P_\varepsilon \in \mathcal{P}^C(\mathcal{M})$ for $\varepsilon \in (0,1)$. However, $\lim_{\varepsilon\to 0}\prod_{i=1}^n P_\varepsilon(a_i, M_i) = \prod_{i=2}^n P(a_i, M_i)$ while $\lim_{\varepsilon\to 0}\prod_{i=1}^n P_\varepsilon(a_{i+1}, M_i) = 0$. Hence, there exists $\varepsilon^* \in (0,1)$ such that Equation (2) is not satisfied and so, by Theorem 1, $P_{\varepsilon^*} \notin \mathcal{L}^C(\mathcal{M})$. Thus, $\mathcal{M}$ is not Luce unfalsifiable. □

Proof of Proposition 3

Proof. Let $\mathcal{M}_1$ and $\mathcal{M}_2$ be disjoint experiments, C a choice correspondence on $\mathcal{M}_1 \cup \mathcal{M}_2$, and $P_1 \in \mathcal{L}^C(\mathcal{M}_1)$.

First, suppose $\mathcal{E}^C(\mathcal{M}_1\cup\mathcal{M}_2) = \mathcal{E}^C(\mathcal{M}_1)\cup\mathcal{E}^C(\mathcal{M}_2)$. Let $P \in \mathcal{P}^C(\mathcal{M}_1\cup\mathcal{M}_2|P_1)$ such that $P_{\mathcal{M}_2} \in \mathcal{L}^C(\mathcal{M}_2)$. Then, since $P_{\mathcal{M}_1} = P_1 \in \mathcal{L}(\mathcal{M}_1)$, it follows that Equation (2) is satisfied for all elementary C-cycles on $\mathcal{M}_1 \cup \mathcal{M}_2$ and so $P \in \mathcal{L}^C(\mathcal{M}_1 \cup \mathcal{M}_2)$. Hence, $\mathcal{M}_2$ cannot Luce falsify $(\mathcal{M}_1, P_1)$.

Now, suppose $\mathcal{E}^C(\mathcal{M}_1 \cup \mathcal{M}_2) \neq \mathcal{E}^C(\mathcal{M}_1) \cup \mathcal{E}^C(\mathcal{M}_2)$. Then, there exists an elementary C-cycle $\phi = [a_1, \ldots, a_{n+1}; M_1, \ldots, M_n]$ on $\mathcal{M}_1 \cup \mathcal{M}_2$ where $M_1 \in \mathcal{M}_1$ and $M_n \in \mathcal{M}_2$. Now define two SCFs P and P' on $\mathcal{M}_1 \cup \mathcal{M}_2$ by, for all $(a, M) \in \mathcal{S}(\mathcal{M}_1 \cup \mathcal{M}_2)$,

$$P(a, M) = \begin{cases} P_1(a, M) & \text{if } M \in \mathcal{M}_1 \\ \frac{1}{|C(M)|} & \text{if } M \in \mathcal{M}_2,\ a \in C(M) \\ 0 & \text{otherwise} \end{cases}$$

and

$$P'(a, M) = \begin{cases} P_1(a, M) & \text{if } M \in \mathcal{M}_1 \\ \frac{1+\mathbb{1}[a=a_1]}{|C(M)|+\mathbb{1}[a_1\in C(M)]} & \text{if } M \in \mathcal{M}_2,\ a \in C(M) \\ 0 & \text{otherwise} \end{cases}.$$

Hence, $P, P' \in \mathcal{P}^C(\mathcal{M}_1 \cup \mathcal{M}_2|P_1)$. Moreover, $(\mathcal{M}_2, P_{\mathcal{M}_2})$ can be C-censored Luce rationalized by the utility $v(a) = 1$ for all $a \in \mathcal{A}(\mathcal{M}_2)$ and $(\mathcal{M}_2, P'_{\mathcal{M}_2})$ can be C-censored Luce rationalized by the utility $v'(a) = 1 + \mathbb{1}[a = a_1]$ for all $a \in \mathcal{A}(\mathcal{M}_2)$. If P is not Luce rationalizable we are done, so suppose that $P \in \mathcal{L}^C(\mathcal{M}_1\cup\mathcal{M}_2|P_1)$. In that case, we show that $P' \notin \mathcal{L}^C(\mathcal{M}_1 \cup \mathcal{M}_2|P_1)$, which implies that $\mathcal{M}_2$ can

C-Luce falsify $(\mathcal{M}_1, P_1)$. By Theorem 1, it is sufficient to show that Equation 2 is not satisfied for ϕ. To show this, note that, on one hand,

$$\frac{\prod_{i=1}^n P'(a_i, M_i)}{\prod_{i=1}^n P'(a_{i+1}, M_i)} = \frac{\prod_{i:M_i\in\mathcal{M}_1} P_1(a_i, M_i) \frac{1}{\prod_{i:M_i\in\mathcal{M}_2}(|C(M_i)|+\mathbb{1}[a_1\in M_i])}}{\prod_{i:M_i\in\mathcal{M}_1} P_1(a_{i+1}, M_i) \frac{2}{\prod_{i:M_i\in\mathcal{M}_2}(|C(M_i)|+\mathbb{1}[a_1\in M_i])}}$$
$$= \frac{1}{2}\frac{\prod_{i:M_i\in\mathcal{M}_1} P_1(a_i, M_i)}{\prod_{i:M_i\in\mathcal{M}_1} P_1(a_{i+1}, M_i)}$$

because $M_1 \in \mathcal{M}_1$ but $M_n \in \mathcal{M}_2$. On the other hand,

$$\frac{\prod_{i=1}^n P(a_i, M_i)}{\prod_{i=1}^n P(a_{i+1}, M_i)} = \frac{\prod_{i:M_i\in\mathcal{M}_1} P_1(a_i, M_i) \frac{1}{\prod_{i:M_i\in\mathcal{M}_2}|C(M_i)|}}{\prod_{i:M_i\in\mathcal{M}_1} P_1(a_{i+1}, M_i) \frac{1}{\prod_{i:M_i\in\mathcal{M}_2}|C(M_i)|}} = \frac{\prod_{i:M_i\in\mathcal{M}_1} P_1(a_i, M_i)}{\prod_{i:M_i\in\mathcal{M}_1} P_1(a_{i+1}, M_i)},$$

and so

$$\frac{\prod_{i=1}^n P'(a_i, M_i)}{\prod_{i=1}^n P'(a_{i+1}, M_i)} \neq \frac{\prod_{i=1}^n P(a_i, M_i)}{\prod_{i=1}^n P(a_{i+1}, M_i)} = 1.$$

Hence, Equation 2 is not satisfied for the elementary C-cycle ϕ on $\mathcal{M}_1 \cup \mathcal{M}_2$, and so $P' \notin \mathcal{L}^C(\mathcal{M}_1 \cup \mathcal{M}_2 | P_1)$. Hence, $\mathcal{M}_2$ can C-Luce falsify $(\mathcal{M}_1, P_1)$. ☐

Proof of Proposition 4

Proof. Let $\theta = (\mathcal{M}, P)$ be a stochastic choice dataset with $v \in \mathcal{V}(\theta)$, and $w : \mathcal{A}(\mathcal{M}) \to \mathbb{R}_{++}$. First, suppose that there exists a collection of strictly positive scalars $(\lambda_1, \ldots, \lambda_m)$ such that, for $i = 1, \ldots, m$, $v(a) = \lambda_i w(a)$ whenever $a \in Q_i \in \mathcal{Q}^{c_\theta}(\mathcal{M})$. Let $(a, M) \in \mathcal{S}(\mathcal{M})$; then, by Lemma 2, $C_\theta(M) \subseteq Q_i$ for some $i = 1, \ldots, m$ and $C_\theta(M) \cap Q_j = \emptyset$ for all $j \neq i$. Therefore,

$$P_v^{C_\theta}(a, M) = \mathbb{1}[a \in C_\theta(M)] \frac{v(a)}{\sum_{b\in C_\theta(M)} v(b)} = \mathbb{1}[a \in C_\theta(M)] \frac{\lambda_i w(a)}{\sum_{b\in C_\theta(M)} \lambda_i w(b)}$$
$$= \mathbb{1}[a \in C_\theta(M)] \frac{w(a)}{\sum_{b\in C_\theta(M)} w(b)} = P_w^{C_\theta}(a, M),$$

and so $w \in \mathcal{V}(\theta)$.

Now let $w \in \mathcal{V}(\theta)$. For each $i = 1, \ldots, m$, choose some $a_i^* \in Q_i \in \mathcal{Q}^{C_\theta}(\mathcal{M})$ and let $\lambda_i = \frac{v(a_i^*)}{w(a_i^*)}$. Fix some $i = 1, \ldots, m$ and let $a \in Q_i$. If $a = a_i^*$, then $\lambda_i w(a) = v(a)$ by construction. If $a \neq a_i^*$, then there exist an overlapping C_θ-sequence $[a_i^*, a_2, \ldots, a_n, a; M_1, \ldots, M_n]$ on $\mathcal{M}$ that C_θ-connects a_i^* and a. Therefore,

$$\frac{w(a_i^*)}{w(a_2)}\frac{w(a_2)}{w(a_3)}\cdots\frac{w(a_n)}{w(a)} = \frac{P(a_i^*, M_1)\sum_{b\in C_\theta(M_1)} w(b)}{P(a_2, M_1)\sum_{b\in C_\theta(M_1)} w(b)}\cdots\frac{P(a_n, M_n)\sum_{b\in C_\theta(M_n)} w(b)}{P(a, M_n)\sum_{b\in C_\theta(M_n)} w(b)}$$
$$= \frac{P(a_i^*, M_1)\sum_{b\in C_\theta(M_1)} v(b)}{P(a_2, M_1)\sum_{b\in C_\theta(M_1)} v(b)}\cdots\frac{P(a_n, M_n)\sum_{b\in C_\theta(M_n)} v(b)}{P(a, M_n)\sum_{b\in C_\theta(M_n)} v(b)}$$
$$= \frac{v(a_i^*)}{v(a_2)}\frac{v(a_2)}{v(a_3)}\cdots\frac{v(a_n)}{v(a)}$$

because $v, w \in \mathcal{V}(\theta)$. Hence,

$$\frac{w(a_i^*)}{w(a)} = \frac{v(a_i^*)}{v(a)} = \frac{\lambda_i w(a_i^*)}{v(a)},$$

and so $v(a) = \lambda_i w(a)$. □

Proof of Corollary 4

Proof. Let $\mathcal{M}$ be an experiment and C be a choice correspondence on $\mathcal{M}$. First, suppose that $a \sim_{\mathcal{M}}^C b$ for all $a, b \in \mathcal{A}^C(\mathcal{M})$. Then, for any stochastic choice dataset $\theta = (\mathcal{M}, P)$ that is C-censored, $|Q^{C_\theta}(\mathcal{M})| = 1$. Therefore, by Theorem 4, if $v, w \in \mathcal{V}(\theta)$, there exists $\lambda > 0$ such that $v(a) = \lambda w(a)$ for all $a \in \mathcal{A}^C(\mathcal{M})$, and so $\mathcal{M}$ is C-Luce identified.

Conversely, suppose that $\mathcal{Q}^C(\mathcal{M}) = \{Q_1, \ldots, Q_m\}$ where $m > 1$ and let $\theta = (\mathcal{M}, P)$ be a C-censored stochastic choice dataset with $v \in \mathcal{V}(\theta)$. Define $w : \mathcal{A}(\mathcal{M}) \to \mathbb{R}_{++}$ by

$$w(a) = \begin{cases} 2v(a) & \text{if } a \in Q_1 \\ v(a) & \text{otherwise} \end{cases}.$$

Then, by Proposition 4, $w \in \mathcal{V}(\theta)$ but there does not exist $\lambda > 0$ such that $v(a) = \lambda w(a)$ for all $a \in \mathcal{A}^C(\mathcal{M})$; hence $\mathcal{M}$ is not C-Luce identified. □

Proof of Proposition 5

Proof. First, suppose $\mathcal{M}_1$ can C-Luce predict stochastic choice on $\mathcal{M}_2$. We first show that $\mathcal{A}^C(\mathcal{M}_2) \subseteq \mathcal{A}^C(\mathcal{M}_1)$. For contradiction, suppose not. Then, there exists $a \in C(M) \in C(\mathcal{M}_2)$ with $a \in \mathcal{A}^C(\mathcal{M}_1 \cup \mathcal{M}_2) \setminus \mathcal{A}^C(\mathcal{M}_1)$. Moreover, since $C(\mathcal{M}_2)$ does not contain singletons, there is $b \in C(M) \setminus \{a\}$. Now let $P_1 \in \mathcal{L}^C(\mathcal{M}_1)$ such that there exists a unique extension $P \in \mathcal{L}^C(\mathcal{M}_1 \cup \mathcal{M}_2 | P_1)$. Let $v \in \mathcal{V}(\mathcal{M}_1 \cup \mathcal{M}_2, P)$ and define $v' : \mathcal{A}(\mathcal{M}_1 \cup \mathcal{M}_2) \to \mathbb{R}_{++}$ by

$$v'(d) = \begin{cases} v(d) & \text{if } a \in \mathcal{A}^C(\mathcal{M}_1) \\ v(d) + 1 & \text{if } a \in \mathcal{A}^C(\mathcal{M}_1 \cup \mathcal{M}_2) \setminus \mathcal{A}^C(\mathcal{M}_1) \\ 0 & \text{otherwise} \end{cases}$$

By construction, the restriction of $P_{v'}$ to $\mathcal{M}_1$ coincides with P_1. However, if $b \in \mathcal{A}^C(\mathcal{M}_1)$, then

$$\frac{P_v(a, M)}{P_v(b, M)} = \frac{v(a)}{v(b)} \neq \frac{v(a) + 1}{v(b)} = \frac{P_{v'}(a, M)}{P_{v'}(b, M)};$$

and if $b \in \mathcal{A}^C(\mathcal{M}_1 \cup \mathcal{M}_2) \setminus \mathcal{A}^C(\mathcal{M}_1)$, then

$$\frac{P_v(a, M)}{P_v(b, M)} = \frac{v(a)}{v(b)} \neq \frac{v(a) + 1}{v(b) + 1} = \frac{P_{v'}(a, M)}{P_{v'}(b, M)}.$$

Hence, $P_v \neq P_{v'}$ and so $|\mathcal{L}^C(\mathcal{M}_1 \cup \mathcal{M}_2 | P_1)| > 1$.

Now suppose there exists $P_1 \in \mathcal{L}^C(\mathcal{M}_1)$ such that $|\mathcal{L}^C(\mathcal{M}_1 \cup \mathcal{M}_2 | P_1)| = 1$, and let $a, b \in \mathcal{A}(\mathcal{M}_2)$ be such that $a \sim^C_{\mathcal{M}_2} b$ but, for contradiction, $a \nsim^C_{\mathcal{M}_1} b$. Since $P \in \mathcal{L}^C(\mathcal{M}_1)$, there exists a utility function $v : \mathcal{A}(\mathcal{M}_1) \mapsto \mathbb{R}_{++}$ such that $P = P^C_v$. Since $a \nsim^C_{\mathcal{M}_1} b$, a and b are in different equivalence classes of $\mathcal{M}_1$, which we denote Q_a and Q_b, respectively. Fix any $\lambda > 0$ and define a utility function $v_\lambda : \mathcal{A}(\mathcal{M}_1) \mapsto \mathbb{R}_{++}$ by

$$v_\lambda(c) = \begin{cases} \lambda v(c) & \text{if } c \in Q_a \\ v(c) & \text{if } c \in \mathcal{A}(\mathcal{M}_1) \setminus Q_a \end{cases}$$

Denote $P_\lambda = P_{v_\lambda}$, which is a SCF on $\mathcal{M}_1 \cup \mathcal{M}_2$, and denote by P^1_λ the restriction of P_λ to $\mathcal{M}_1$. It follows from Proposition 4 that $P^1_\lambda = P$ for all λ, that is, $P_\lambda \in \mathcal{L}^C(\mathcal{M}_1 \cup \mathcal{M}_2 | P)$ for all λ. However, $a \sim^C_{\mathcal{M}_2} b$ and hence $a \sim^C_{\mathcal{M}_1 \cup \mathcal{M}_2} b$, i.e. a and b are in the same equivalence class of $\sim^C_{\mathcal{M}_1 \cup \mathcal{M}_2}$. It follows from Proposition 4 that $P_\lambda \neq P_{\lambda'}$ whenever $\lambda \neq \lambda'$, i.e. $\mathcal{L}^C(\mathcal{M}_1 \cup \mathcal{M}_2 | P)$ contains infinitely many different C-Luce rationalizable SCFs extending P. This contradicts that $|\mathcal{L}^C(\mathcal{M}_1 \cup \mathcal{M}_2 | P_1)| = 1$, and so it must be the case that $a \sim^C_{\mathcal{M}_1} b$.

For the converse, suppose $\mathcal{A}^C(\mathcal{M}_2) \subseteq \mathcal{A}^C(\mathcal{M}_1)$ and, for all $a, b \in \mathcal{A}^C(\mathcal{M}_2)$, $a \sim^C_{\mathcal{M}_2} b$ implies $a \sim^C_{\mathcal{M}_1} b$. Now let $P_1 \in \mathcal{L}^C(\mathcal{M}_1)$, and suppose there exist two $P, P' \in \mathcal{L}^C(\mathcal{M}_1 \cup \mathcal{M}_2 | P_1)$. Since $\mathcal{A}^C(\mathcal{M}_2) \subseteq \mathcal{A}^C(\mathcal{M}_1)$ there exist $v, v' : \mathcal{A}(\mathcal{M}_1) \mapsto$

$\mathbb{R}_{++}$ such that $P = P_v^C$ and $P' = P_{v'}^C$ as SCFs on $\mathcal{M}_1 \cup \mathcal{M}_2$. Note that the restrictions of P, P' to $\mathcal{M}_1$ are both equal to P_1.

Let $Q_1, \ldots, Q_K$ be the equivalence classes of $\sim^C_{\mathcal{M}_1}$ on $\mathcal{M}_1$. Since $C(\mathcal{M}_2)$ has no singleton menus and $a \sim^C_{\mathcal{M}_2} b$ implies $a \sim^C_{\mathcal{M}_1} b$, it follows that $Q_1, \ldots, Q_K$ are also the equivalence classes of $\sim^C_{\mathcal{M}_1 \cup \mathcal{M}_2}$ on $\mathcal{M}_1 \cup \mathcal{M}_2$.

Let v_k, v'_k be the restrictions of v, v' to the equivalence class Q_k for $k = 1, \ldots, K$. Since the restrictions of P and P' to $\mathcal{M}_1$ are both equal to P_1 and v, v' are defined on $\mathcal{M}_1$, Theorem 4 implies that there exist $\lambda_1, \ldots, \lambda_K$ such that $v_k = \lambda_k v'_k$ for $k = 1, \ldots, K$. But since $Q_1, \ldots, Q_K$ are also the equivalence classes of $\sim^c_{\mathcal{M}_1 \cup \mathcal{M}_2}$, Proposition 4 implies that $P = P'$ as SCFs on $\mathcal{M}_1 \cup \mathcal{M}_2$. Hence $|\mathcal{L}(\mathcal{M}_1 \cup \mathcal{M}_2 | P_1)| = 1$ and, since P_1 was arbitrary, it follows that $\mathcal{M}_1$ can C-Luce predict $\mathcal{M}_2$. □