Multivalued logit: Revealing bundle utility from choice*

Henrik Petri*

Abstract. We introduce a logit type model of bundle choice that allows for complementarities between alternatives. A utility-value is assigned to each bundle of alternatives. The probability of choosing a bundle is proportional to the utility-value assigned to it. The data available to the researcher is in the form of product-level market shares, i.e. a marginal distribution. We characterize the model and show that its underlying parameters are unique (up to multiplication by a scalar). An implication of uniqueness is that any complementarities between alternatives can be revealed from data.

Keywords: stochastic choice, bundle choice, marginal data, complementarity, substitutability, logit.

JEL codes: D01, D90.

1. Introduction

A standing assumption in much of the stochastic/discrete choice literature is that the set of alternatives compromising a choice set are *mutually exclusive*. However, this assumption is at odds with economic environments where consumers (may) choose multiple (bundles of) alternatives. In particular, this assumption precludes complementarities between alternatives in a choice set. Consider a customer visiting a grocery store. Among coffees, she might not just choose one type of coffee but several ones (for instance differing in their strengths, such as a regular one and a stronger one). Or, it might be, that she enjoys her coffee more with a pastry, or more with milk, and hence purchases a combination of these alternatives. Due to complementarities, some alternatives are enjoyed more when purchased together as a bundle than

* *First version:* January 31, 2023. *This version:* March 07, 2024. I thank John Rehbeck and Kai Wang for comments that significantly improved the manuscript.

*Department of Economics, University of Gothenburg, Sweden, E-mail: henrik@petri.se.

alone. Such complementarities are, however, difficult to infer from standard discrete choice data due to the mutual exclusivity of alternatives.

Ideally, to analyse bundle choice the researcher would need to observe the frequencies by which each bundle/subset of a set of alternatives is chosen. For instance, many of the existing econometric methods on bundle choice require individual level choice data on bundles such as scanner-data. However such data can be costly and hard to come by in practice. In this paper we try to side-step this issue and propose to analyse bundle choices using the more accessible "product-level market share" choice data (for a related discussion see Wang (2021) who is motivated by similar concerns as us and propose econometric methods to identify demand in bundle models with observable product-level market shares). That is, we assume that the researcher observes the aggregate market shares of each alternative (i.e. the fraction of the population that chooses a particular alternative) but not the individual bundle choice probabilities (i.e. how often each bundle is chosen).

Besides being a more common type of data,[1] an advantage of studying marginal choice data over bundle data is that it is more parsimonious. A probability distribution over bundles from a set of n alternatives requires $2^n - 1$ parameters (one for each bundle). Whereas describing the marginal distribution only requires n parameters, i.e. the marginal probability of each alternative.[2] To illustrate, consider data on purchases from a grocery store. If the grocery store only records the number of customers that chooses each item, but not the exact bundles that they choose, this gives access to the marginal distribution, but not the distribution over bundles.

In this paper we study a bundle version of the classical logit rule that allows complementarities between alternatives to be inferred from choice data. The proposed model is a natural extension of the classical (single-valued) Luce rule. The main ingredient is a utility function on bundles of alternatives, where the probability of choosing a bundle is proportional to the utility value assigned to the bundle. We offer a characterization of the model and show that the underlying parameters are

---

[1]I.e. it is a theoretically less demanding notion as every multivalued choice rule induces a marginal probability distribution.

[2]See also Ahn, Echenique, and Saito (2018) who are motivated by similar concerns in the context of average choice data.

identified. An implication is that any complementarities between alternatives are revealed by data.

We give necessary and sufficient conditions on marginal stochastic choice data to be consistent with (i.e. arise as the marginal probability of) a multivalued logit rule. Our main property is path independence, which is reminiscent of Luce's choice axiom (we refer to section 3.1 for a precise statement and discussion of this property). We also show that the representation of any marginal data set by a multivalued logit rule is unique (up to multiplication by a scalar). Perhaps surprisingly, this means that the model has just as good identification properties as the classical logit rule, despite the larger number of parameters of the model. Thus, in the case of logit, we can use the less rich marginal choice data to identify the underlying distribution on bundles.

The representation by a multivalued logit rule gives a baked in definition of complementarity and substitutability. Roughly speaking, two alternatives a, b are complements (substitutes) if the utility value of the bundle $\{a, b\}$, is larger (smaller) than the utility of the alternatives a and b in isolation, i.e. if $\sigma(\{a, b\}) > \sigma(\{a\}) + \sigma(\{b\})$. Section 5 establishes necessary and sufficient conditions on observed choice data for there to be complementarities between pairs of alternatives. We also introduce higher order versions of complementarity/substitutability applicable to bundles of size larger than two. Essentially, there is complementarity at a bundle A if there is additional complementarity between alternatives in A compared to any complementarity between alternatives in subsets of A. Our identification result implies that complementarity and substitutability between alternatives in a bundle can be inferred from (marginal) choice data alone.

We also consider a special case of multivalued logit where only binary bundles (containing at most two alternatives) are chosen by the agent. This model deviates minimally from the standard single-valued/univalent logit rule in that the choices are "almost" univalent. The binary-valued logit model is characterized by a relaxation of the Luce choice axiom. The standard Luce choice axiom says that the probability of choosing an alternative $a \in A$ equals the probability of choosing a subset $A \setminus c$ (where $c \neq a$) times the probability of choosing a in $A \setminus c$. Roughly, our property says that the probability of choosing a bundle containing both a and b in a menu A

equals the probability of choosing a subset of the form $A \setminus c$ (where $c \neq a$ and $c \neq b$) times the probability of choosing a bundle containing both a and b in $A \setminus c$.

The paper is concluded with a discussion of the relation to the classical Luce rule for stochastic choice functions. As a corollary to our characterization of multivalued logit we obtain a (to the best of our knowledge) new characterization of the classical Luce rule. The model is characterized by our path independence property. Although path independence is a more permissive condition than Luce's choice axiom for marginal choice data, we show that these two conditions are equivalent for (positive) stochastic choice functions.

1.1. **Related literature.** This paper relates to an econometric literature that considers estimation and identification of complementarities in discrete choice models of bundle choice (Gentzkow (2007), Allen and Rehbeck (2019), Iaria and Wang (2020), Wang (2021), Allen and Rehbeck (2022)). [3]Particularly close to the present paper is a subset of this literature that establish identification in bundle models with only aggregate product-level market shares available to the analyst (see e.g. Allen and Rehbeck (2019), Wang (2021) and Allen and Rehbeck (2022)). A challenge when establishing identification in these models is that there are usually many more parameters than data points. I.e. observing a high demand for one alternative could either be due to consumers' preference for this good, but also be due to complementarities with other products. These two sources of demand need to be disentangled. We show how to identify bundle utilities, and hence complementarities, in a logit model with only marginal data on choices available.

The notion of marginal data considered in this paper is inspired by Manzini, Mariotti, and Ülkü (2022) who propose to analyze approval behavior using a "stochastic approval function". A stochastic approval function is similar to the type of data considered in this paper in that it does not impose any constraint on choice probabilities to sum to one in menus. They also consider the notion of marginalization and view a stochastic approval function as the marginalization of approval probabilities. Azrieli and Rehbeck (2022) also study marginal choice data. However, their notion of marginal data is different from ours. They use marginal choice data to infer properties

[3]For a related decision-theoretic paper that studies model-free and behavioral notions of complementarity see Manzini, Mariotti, and Ülkü (2019).

of standard (univalent) stochastic choice models, whereas we consider the marginal bundle data and apply it to study a bundle version of the standard logit choice model. Wang (2022) studies similar data but in a different context. He interprets the data as a binary authorization/categorization probability. Similarly to us, he characterizes a logit type rule, but the interpretation of his model is different from ours.

Our paper is related to a strand of literature following Manzini and Mariotti (2014) that studies stochastic consideration set formation in choice. Particularly close is Brady and Rehbeck (2016) who consider a random conditional choice set rule (RCCSR). The RCCSR assumes deterministic preferences and that feasible sets are drawn according to a multivalued logit rule (as studied in this paper). Preference maximizing alternatives are then chosen from each feasible set of alternatives. Similar to Brady and Rehbeck (2016) our proof technique makes use of Möbius inversion to define the utility values of subsets of alternatives. A notable difference between our approaches is that we do not need to assume an outside default option. The default option is needed to identify the underlying preferences in an RCCSR. Since we only model the feasibility correspondence, the default option is not needed. Another related paper is Ravid and Steverson (2019). In subsection 4.1 we consider a binary version of multivalued logit (where only bundles of size smaller than two are chosen by the DM). This version of our model can be viewed as an additive counterpart of their model.

Finally, this paper also relates to a recent strand of literature that explores issues that arise when inferring multivalued behaviors from single-valued choice data. Inferring underlying multivalued behaviors using single-valued data calls for various auxiliary assumptions on how agents arrive at unique choices (such as the tie-breaking rule used by individuals). This can in turn cause issues with identification (see e.g. Petri (2022) for a discussion). Only a few recent papers propose methods to elicit (deterministic) multivalued choices directly using (non-) forced multivalued choice tasks (see e.g. Costa-Gomes, Cueva, and Gerasimou (2021) and Gerasimou (2022)). However, these papers are mainly situated in a deterministic (non-repeated) setting.

2. Data

We denote by X a finite set of alternatives. A nonempty subset $A \subseteq X$ is called a *menu*. Let $\mathcal{A}$ denote the collection of nonempty subsets of X, i.e. $\mathcal{A}$ is the collection of all menus. Ideally, to describe bundle choices the researcher would need to observe, for each menu of choices, the probability by which each bundle in the menu is chosen. Formally, such a dataset can be described by a *stochastic choice rule*, which is a mapping $R : \mathcal{A} \times \mathcal{A} \to [0, 1]$ such that $\sum_{B \subseteq A} R(B, A) = 1$ for all menus $A \in \mathcal{A}$ and such that $R(B, A) > 0$ only if $B \subseteq A$. For each $A \in \mathcal{A}$, $R(\cdot, A)$ is thus a probability distribution on subsets of A and $R(B, A)$ can be interpreted as the probability of choosing bundle B in A. However, as explained in the introduction, this type of bundle data is rare in practice. We instead assume that the researcher observes a marginal SCR, i.e. the market-shares of each alternative.

Definition 2.1. A function $\rho : X \times \mathcal{A} \mapsto [0, 1]$ is a *marginal stochastic choice rule (marginal SCR)* if there is a stochastic choice rule R such that

$$\rho(a, A) = \sum_{a \in B : B \subseteq A} R(B, A)$$

for all $a \in A$ and $A \subseteq X$. ◁

We wish to emphasize that even though the adding-up constraint $\sum_{a \in A} \rho(a, A) = 1$ does not hold for marginal SCRs, the numbers $\rho(a, A)$ have a clear probabilistic interpretation. To see this, suppose that we have a given probability distribution on bundles in a menu A. The probability of selecting alternative a in menu A is then simply the sum of probabilities over all bundles B containing a, i.e. $\sum_{a \in B : B \subseteq A} R(B, A)$. Another interpretation of $\rho(a, A)$ is as the aggregate market-share of alternative a in A. I.e. it is obtained by summing the market shares $R(B, A)$ of all bundles containing a.

Note that the underlying SCR R in definition 2.1 is not observed by the researcher, but has to be inferred from ρ. Without further information, there may be several SCRs R inducing the same (observed) marginal SCR. To clarify this we may also define a marginal SCR without reference to an underlying SCR R. The next proposition shows that a marginal SCR can equivalently be defined as a mapping $\rho : X \times \mathcal{A} \mapsto [0, 1]$ such that for all $A \subseteq X$: $\sum_{a \in A} \rho(a, A) \geq 1$.

Proposition 2.2. *A function $\rho : X \times \mathcal{A} \to [0,1]$ is a marginal SCR if and only if for each menu $A \subseteq X$:*

$$\sum_{a \in A} \rho(a, A) \geq 1.$$

As mentioned, marginal SCRs are generally not rich enough to identify underlying multivalued choices R. However, as we will see in section 3, this type of data does not lose any power in case of multivalued logit (i.e. identification is possible).

3. Multivalued logit

This section presents and discusses the main model studied in this paper.

Definition 3.1. A marginal SCR ρ is a *multivalued logit rule* if there is a function $\sigma : 2^X \to [0, \infty)$ such that

$$\rho(a, A) = \frac{\sum\limits_{B \subseteq A : a \in B} \sigma(B)}{\sum\limits_{B \subseteq A} \sigma(B)} \tag{1}$$

for all $a \in A$ and $A \subseteq X$ and $\sigma(\{a\}) > 0$ for all $a \in A$. $\triangleleft$

The resemblance with the classical (univalent) logit choice rule is clear. A stochastic choice function is a *univalent logit rule* (or classical logit rule) if there is a utility function $u : X \to (0, \infty)$ such that $\rho(a, A) = \frac{u(a)}{\sum_{b \in A} u(b)}$ for all $a \in A$ and $A \subseteq X$. The univalent logit rule can be viewed as a special case of multivalued logit, corresponding to a function σ s.t. $\sigma(A) = 0$ for all $A \subseteq X$ with $|A| \geq 2$.

The parameter σ in a multivalued logit rule can be interpreted as a measure of the desirability/utility value of a bundle of alternatives. Faced with a menu of alternatives A the decision maker chooses a specific bundle of alternatives $B \subseteq A$ with probability proportional to its utility value $\sigma(B)$. I.e. with probability $R(B, A) = \frac{\sigma(B)}{\sum\limits_{B \subseteq A} \sigma(B)}$. It is worth repeating that the bundle choice probabilities $R(B, A)$ are not observable in our framework, but rather the marginal probabilities $\rho(a, A) = \sum\limits_{a \in B \subseteq A} R(B, A)$. If bundle probabilities $R(B, A)$ were observable then identification and characterization of the model would follow along similar lines of reasoning as for the classical logit rule.

3.1. **Characterization and identification.** The multivalued logit rule is characterized by three different properties on choice probabilities. The first property is a weak version of the positivity/full support assumption of the standard (univalent) logit rule. Roughly speaking, it excludes the existence of monopolistic goods in menus, i.e. goods such that each and every individual chooses them.

Varied support. A marginal SCR ρ satisfies *varied support* if $\rho(a, A) < 1$ for all $a \in A$ and $A \subseteq X$.

If a marginal SCR satisfies the varied support assumption then there is for every $a \in A$ an alternative $b \neq a$ such that $\rho(b, A) > 0$. This condition is clearly necessary for a multivalued logit since it assumes that all singletons are chosen with positive probability. The next property, path independence, is the main property in our characterization. It is related to (and in our setting weaker than) Luce's choice axiom. We discuss the relation to Luce's choice axiom further in section 6.

Path independence. A marginal SCR ρ satisfies *path independence* if

$$[1 - \rho(a, A)][1 - \rho(b, A \setminus a)] = [1 - \rho(b, A)][1 - \rho(a, A \setminus b)] \tag{2}$$

for all $a, b \in A$ and $A \subseteq X$.

Recall that $\rho(a, A)$ can be interpreted as the fraction of times that a decision maker chooses (a bundle containing) a. The quantity $1 - \rho(a, A)$ can hence be interpreted as the fraction of times that a bundle not containing a is chosen, or as the fraction of times that alternative a is not chosen in A. To see this, note that

$$1 - \rho(a, A) = \sum_{B \subseteq A} R(B, A) - \sum_{a \in B : B \subseteq A} R(B, A) = \sum_{B \subseteq A \setminus a} R(B, A).$$

To understand the intuition behind path independence consider the probability of choosing a bundle in $A \setminus \{a, b\}$ when all alternatives in A are available. Roughly speaking, the probability that an agent chooses a bundle $B \subseteq A \setminus \{a, b\}$ equals the probability that alternative a is not chosen in A (equivalently: the probability that a bundle not containing A is chosen) times the probability that b is not chosen when only the alternatives in $A \setminus a$ are available. Path independence says that this probability is independent on whether we first eliminate a from A and then b from $A \setminus a$, or, if we first eliminate b from A and then a from $A \setminus b$.

We next introduce our final property, which is a weak positivity assumption reminiscent of the eponymous Block-Marschak polynomials (Block and Marschak, 1960) characterizing the random utility model. To state it, we define an auxiliary function $f^\rho : 2^X \to (0, \infty)$ for each $A \subseteq X$ with $A \neq X$ by setting

$$f^\rho(A) = \prod_{i=1}^{k} [1 - \rho(x_i, X \setminus \{x_1, \ldots, x_{i-1}\})] \tag{3}$$

where $\{x_1, \ldots, x_k\} = X \setminus A$ is an enumeration of $X \setminus A$. Set $f^\rho(X) = 1$. Lemma A.1 in the appendix implies that this function is well-defined, i.e. that $f^\rho(A)$ is independent of the enumeration of $X \setminus A$. Loosely, the function $f^\rho(A)$ gives the probability of choosing *some* bundle in A when the set of available alternatives is X.

Positivity. A marginal SCR ρ satisfies *positivity* if

$$\delta f^\rho(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} f^\rho(B) \geq 0$$

for all $A \subseteq X$.

Note that $f^\rho(A)$ gives the probability that *some* bundle in A is chosen. Suppose that we would like to find the probability that exactly the bundle A is chosen. Then we subtract from $f^\rho(A)$ the probabilities $f^\rho(A \setminus b)$ that some bundle in $A \setminus b$ is chosen (for each $b \in A$). But, then the probabilities of the bundles in the pairwise intersections of the sets $A \setminus b$ have been subtracted too often, so we add them back, and so on. This combinatorial argument, which is similar to the inclusion-exclusion principle, shows that the probability of choosing bundle A is equal to $\sum_{B \subseteq A} (-1)^{|A \setminus B|} f^\rho(B)$. Positivity is thus a quite weak requirement as it only requires each bundle to be chosen with weakly positive probability. Indeed, if ρ is a multivalued logit, then one can show that there is a scalar $c > 0$ s.t. $\delta f^\rho(A) = c\sigma(A)$ for all menus A, so positivity of δf^ρ is equivalent to positivity of σ for multivalued logit.

The following recursive formulation of positivity perhaps clarifies some of the intuition behind it. For all $A \subseteq X$ and $a \in A$, define $\Delta_a f^\rho(A) = f^\rho(A) - f^\rho(A \setminus a)$ and for any sequence $a_1, ..., a_n \in A$ define

$$\Delta_{a_n} \dots \Delta_{a_1} f^\rho(A) = \Delta_{a_n} \dots \Delta_{a_2} f^\rho(A) - \Delta_{a_n} \dots \Delta_{a_2} f^\rho(A \setminus a_1).$$

Then positivity is equivalent to the condition that $\Delta_{a_n} \dots \Delta_{a_1} f^\rho(A) \geq 0$ for all $A \subseteq X$ and any $a_1, \dots, a_n \in A$. Thus positivity says that the probability of choosing some bundle of a set A increases as A gets larger, and it does so at an increasing rate. This interpretation of positivity resembles other positivity conditions considered in the literature such as those considered in Brady and Rehbeck (2016) and Aguiar (2017).

Theorem 3.2. *A marginal SCR is a multivalued logit rule if and only if it satisfies varied support, path independence and positivity. Moreover, the representation is unique (up to multiplication of σ by a scalar).*

The proof of theorem 3.2 is in appendix A.3 and uses Möbius inversion to define a function $\sigma : X \to (-\infty, \infty)$ from marginal choice probabilities ρ. Positivity is then used to assure that σ is weakly positive. If ρ is a multivalued logit then there is a constant $\alpha > 0$ such that $\sigma(A) = \alpha \cdot \delta f^\rho(A)$ for all $A \subseteq X$, so positivity of $\delta f^\rho(A)$ implies positivity of σ. The previous equation also implies that the function σ is unique up to multiplication by a scalar. This relationship is useful in establishing testable implications of versions of multivalued logit obtained by imposing restrictions on σ. Note, for instance, that we do not impose any monotonicity conditions on σ at the outset. I.e., besides being positive, σ is completely unrestricted. But, since σ is revealed by $\delta f^\rho(\cdot)$, we can test any properties on σ we wish by using $\delta f^\rho(\cdot)$. We may, for instance, postulate that σ is larger for smaller sets reflecting a smaller cost of purchasing smaller bundles. A multivalued logit ρ with function $\sigma : X \to (0, \infty)$ is called *monotone* if $\sigma(A) \geq \sigma(B)$ if $A \subseteq B$. From the previous discussion, we thus see that a multivalued logit with monotone σ is characterized by imposing the condition that $\delta f^\rho(A) \geq \delta f^\rho(B)$ if $A \subseteq B$.

3.2. **Binary menus.** In empirical and experimental applications limited menu collections are common. A very common type of domain is the domain consisting of binary choice menus, i.e. all menus A with $|A| = 2$. The characterization of the model is particularly simple when restricted to the collection of binary menus (in the remainder of the paper we will use the short-hand notation $\rho(a, b)$ for $\rho(a, ab)$).

Product rule. A marginal SCR ρ satisfies the product rule if for all $a, b, c \in X$:

$$\frac{1-\rho(b,a)}{1-\rho(a,b)} \times \frac{1-\rho(c,b)}{1-\rho(b,c)} = \frac{1-\rho(c,a)}{1-\rho(a,c)}.$$

Note that $1-\rho(b,a)$ reveals (up to a scalar) the fraction of times an agent chooses a bundle containing the single item a in $\{a,b\}$. This is because $1-\rho(b,a) = \frac{\sigma(a)}{\sigma(a)+\sigma(b)+\sigma(ab)}$. Hence, similarly as for the classical logit rule, it follows that $\frac{1-\rho(b,a)}{1-\rho(a,b)}$ reveals the ratio $\frac{\sigma(a)}{\sigma(b)}$. By this it is immediate that any multivalued logit satisfies the product rule. Below we show that this property characterizes multivalued logit restricted to the binary menus.

Proposition 3.3. *Let $\mathcal{A}$ be the collection of binary menus. A marginal SCR $\rho : X \times \mathcal{A} \to [0,1]$ is a multivalued logit if and only if it satisfies the product rule and varied support.*

Proof. Fix $a_0 \in X$ and define $\sigma(a) = \frac{1-\rho(a_0,a)}{1-\rho(a,a_0)}$ for all $a \in X$. For all $a, b \in X$ define

$$\sigma(\{a,b\}) = \frac{\rho(a,b)[\sigma(a)+\sigma(b)] - \sigma(a)}{1-\rho(a,b)}.$$

To show that σ is well-defined we need to show that

$$\frac{\rho(a,b)[\sigma(a)+\sigma(b)] - \sigma(a)}{1-\rho(a,b)} = \frac{\rho(b,a)[\sigma(b)+\sigma(a)] - \sigma(b)}{1-\rho(b,a)}$$

for all $a, b \in X$. But, one can show that the previous equality holds if and only if $\frac{\sigma(b)}{\sigma(a)} = \frac{1-\rho(a,b)}{1-\rho(b,a)}$. By the product rule and the definition of σ it follows that

$$\frac{1-\rho(a,b)}{1-\rho(b,a)} = \frac{1-\rho(a_0,b)}{1-\rho(b,a_0)} \times \frac{1-\rho(a,a_0)}{1-\rho(a_0,a)} = \frac{\sigma(b)}{\sigma(a)}.$$

We next show that ρ is a multivalued logit. Let $a, b \in X$. By definition of σ we have $\sigma(\{a,b\}) = \frac{\rho(a,b)[\sigma(a)+\sigma(b)]-\sigma(a)}{1-\rho(a,b)}$. Solving for $\rho(a,b)$ we then get the desired representation. □

4. Multivalued logit with limited domain

We next discuss special cases of the model obtained by varying the domain of the function σ. We first consider a model where the agent is restricted to choose binary bundles. We then consider the classical univalent logit rule and discuss the relation to it.

4.1. **Binary-valued logit.** Call a marginal SCR ρ a binary-valued logit rule if there is a *symmetric*[4] function $\sigma : X \times X \to (0, \infty)$ such that

$$\rho(a, A) = \frac{\sum_{b \in A} \sigma(a, b)}{\sum_{\{c,d\} \subseteq A} \sigma(c, d)}$$

for all $a \in A$ and $A \subseteq X$. This model is clearly a special case of multivalued logit. I.e. a binary logit rule is a multivalued logit rule with function σ s.t. $\sigma(A) = 0$ for all $A \subseteq X$ with $|A| \geq 3$. It turns out that the characterization of multivalued logit is considerably simplified by imposing this constraint. For all menus $A \subseteq X$ define $O(a, b, A)$ for all $a, b \in A$ by

$$O(a, b, A) = \rho(a, A) - [1 - \rho(b, A)]\rho(a, A \setminus b).$$

The number $O(a, b, A)$ can be interpreted as the probability of choosing a bundle containing *both* a and b in menu A. Consider the case when ρ is a multivalued logit, then

$$O(a, b, A) = \frac{\sum_{a,b \in B : B \subseteq A} \sigma(B)}{\sum_{B \subseteq A} \sigma(B)}.$$

Thus, $O(a, b, A)$ is the probability by which a, b are chosen together in a bundle in menu A. The case $O(a, b, A) = 0$ corresponds to a case where a and b are completely independent, i.e. a, b are never chosen together in a bundle in A. The higher is $O(a, b, A)$ the higher is the degree of dependence between a and b. Our next property is reminiscent of the classical Luce choice axiom (LCA).

Bundle LCA. For all menus $A \subseteq X$ and alternatives $a, b, c \in A$ it holds that

$$O(a, b, A) = [1 - \rho(c, A)]O(a, b, A \setminus c).$$

Roughly speaking, bundle LCA says that the probability of choosing a bundle $a, b \in A$ equals the probability of not choosing $c \in A$ multiplied with the probability

[4]The function σ is symmetric if $\sigma(a, b) = \sigma(b, a)$ for all $a, b \in X$.

of choosing $a, b \in A \setminus c$ (i.e. when c is not available in A). Note that bundle LCA implies path independence. This is because $O(a, b, \{a, b\}) = \rho(a, b) + \rho(b, a) - 1 = O(b, a, \{a, b\})$ for all $a, b \in A$. Using bundle LCA this in turn implies that $O(a, b, A) = O(b, a, A)$ for all $a, b \in A$ and for all menus A with $|A| = 3$. By induction and bundle LCA it then follows that $O(a, b, A) = O(b, a, A)$ for all $a, b \in A$ and menus A. It is then straightforward to show that these symmetry conditions on the $O(\cdot, \cdot, A)$ are equivalent to path independence.

Proposition 4.1. *A marginal SCR ρ is a binary-valued logit if and only if it satisfies varied support and bundle LCA.*

The proof of proposition 4.1 uses bundle LCA (hence path independence) and Möbius inversion to show that ρ is a *signed* multivalued logit rule, i.e. ρ satisfies equation (1) in definition 3.1 with a function $\sigma : X \to (-\infty, \infty)$. Bundle LCA then assures that σ vanishes for all menus with three or more alternatives. To show that σ is positive for all binary menus, we note that $\delta f^{\rho}(ab) \geq 0$ holds for all $a, b \in X$. I.e. that positivity vacuously holds for binary menus. We refer to the appendix for a complete proof of proposition 4.1.

4.2. **Univalent logit.** Recall that a marginal SCR is a *univalent logit rule* if there is a function $\sigma : X \to (0, \infty)$ such that $\rho(a, A) = \frac{\sigma(a)}{\sum_{b \in A} \sigma(b)}$ for all $a \in A$ and $A \subseteq X$. A *stochastic choice function (SCF)* ρ is a marginal SCR with the adding-up constraint $\sum_{a \in A} \rho(a, A) = 1$. It is well-known that Luce's choice axiom (LCA) characterizes SCFs with a univalent logit representation.

LCA. A marginal SCR ρ satisfies LCA if

$$\rho(a, A) = [1 - \rho(b, A)]\rho(a, A \setminus b)$$

for all $A \subseteq X$ and $a, b \in A$.

We next show that this characterization extends to marginal SCRs. An implication of this is hence that Luce's choice axiom implies that choice probabilities sum to one.

Proposition 4.2. *A marginal SCR ρ is a univalent logit rule if and only if it satisfies varied support and LCA.*

Proof. It is clear that LCA implies Bundle LCA, since LCA implies that $O(a, b, A) = 0$ for all $a, b \in A$ and for all $A \subseteq X$. Hence, by proposition 4.1, ρ is a binary-valued logit rule with function $\sigma : X \times X \to (0, \infty)$. It thus suffices to show that $\sigma(a, b) = 0$ for all distinct $a, b \in X$. To see this, note that for all distinct $a, b \in X$ we have

$$\frac{\sigma(a,b)}{\sum_{\{c,d\} \subseteq A} \sigma(c,d)} = \rho(a, A) - [1 - \rho(b, A)]\rho(a, A \setminus b) = 0,$$

where the last equality follows by LCA. ∎

As a corollary to proposition 4.2, we note that LCA implies that probabilities sum to one (which is not assumed at the outset since we consider marginal data). To see this note that

$$\sum_{a \in A \setminus b} \rho(a, A) = (1 - \rho(b, A)) \sum_{a \in A \setminus b} \rho(a, A \setminus b) = 0 \tag{4}$$

Hence, since $\rho(a, a) = 1$ (which follows by definition of a marginal SCR) we have $\rho(a, b) = (1 - \rho(b, a))\rho(a, a) = 1 - \rho(b, a)$ for all $a, b \in A$. Using an induction argument and equation (4) it then follows that probabilities sum to one in each menu.

5. Complementarity/substitutability between alternatives

The representation in terms of a multivalued logit rule suggests a natural notion of complementarity/substitutability between alternatives. Roughly speaking, two alternatives are complements if they are enjoyed more when chosen/purchased as a bundle than in isolation. In the context of multivalued logit, this means that $\sigma(ab) > \sigma(a) + \sigma(b)$. Similarly, a and b are substitutes if $\sigma(ab) < \sigma(a) + \sigma(b)$. The following is a higher order generalization of substitutability/complementarity to bundles of larger size.

Definition 5.1. There is complementarity (substitutability) at the bundle $A \subseteq X$ if

$$\sum_{B \subseteq A} (-1)^{|A \setminus B|} \sigma(B) > (<)\, 0.$$

◁

As an example consider the bundle $\{a, b, c\}$. If there is complementarity at $\{a, b, c\}$ this means that $0 < \sigma(\{a, b, c\}) - \sigma(ab) - \sigma(ac) - \sigma(bc) + \sigma(a) + \sigma(b) + \sigma(c) =$

$\sigma(\{a,b,c\}) - (\sigma(ab) - \sigma(a) - \sigma(b)) - (\sigma(ac) - \sigma(a) - \sigma(c)) - (\sigma(bc) - \sigma(b) - \sigma(c)) - \sigma(a) - \sigma(b) - \sigma(c)$. I.e. this means that there is additional complementarity at $\{a,b,c\}$ relative to any of the pairwise or singleton bundles. By theorem 3.2 the bundle utility function σ is unique up to an affine transformation. An implication is hence that complementarity and substitutability at a bundle can be revealed from choice data. To see this note that if σ is such that there is complementarity at bundle A, then for any affine transformation $\sigma' = c \cdot \sigma$ of σ there will be complementarity at bundle A as well.

The following proposition gives a simple way to check whether there is complementarity at a small bundle of size two.

Proposition 5.2. *For all bundles $A \subseteq X$ with $|A| = 2$ there is complementarity at $A \subseteq X$ if and only if*

$$\sum_{a \in A} \rho(a, A) \geq \frac{3}{2}.$$

Proof. Let $A \subseteq X$ with $|A| = 2$. Assume that there is complementarity at $A \subseteq X$ so that $\sigma(ab) - \sigma(a) - \sigma(b) > 0$. Then, since ρ is a multivalued logit, it follows that $\sigma(A) = c \cdot \delta f^{\rho}(A)$ for all $A \subseteq X$. Using this we have that

$$\sigma(ab) - \sigma(a) - \sigma(b) = c[\delta f^{\rho}(ab) - \delta f^{\rho}(a) - \delta f^{\rho}(b)] =$$

$$= c[f^{\rho}(ab) - 2f^{\rho}(a) - 2f^{\rho}(b)] = cf^{\rho}(ab)[1 - 2(1 - \rho(b,a)) - 2(1 - \rho(a,b))] =$$

$$cf^{\rho}(ab)[2(\rho(a,b) + \rho(b,a)) - 3].$$

And, it is clear, that the last expression is strictly positive if and only if $\rho(a,b) + \rho(b,a) \geq \frac{3}{2}$. □

The borderline case, when there is zero additional complementarity or substitutability of purchasing a bundle relative to any smaller bundle, corresponds to a case of an additive function σ. The goods in the bundle A are in this case "independent". In the literature on bundle choice, two goods a, b are independent if the demand for alternative a is independent of the price of good b.

Proposition 5.3. *The function $\sigma : 2^X \to [0, \infty)$ is additive, i.e.*

$$\sigma(A) = \sum_{a \in A} \sigma(a)$$

for all $A \subseteq X$ if and only if

$$\sum_{B \subseteq A} (-1)^{|A \setminus B|} \sigma(B) = 0$$

for all $A \subseteq X$ with $|A| \geq 2$.

Proof. Define $g(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} \sigma(B)$ for all $A \subseteq X$. By the Möbius inversion formula in lemma A.2 it then follows that $\sigma(A) = \sum_{B \subseteq A} g(B)$ for all $A \subseteq X$. But, since $g(A) = 0$ for all $A \subseteq X$ with $|A| \geq 2$ and $g(a) = \sigma(a)$, it follows that $\sigma(A) = \sum_{B \subseteq A} g(B) = \sum_{a \in A} \sigma(a)$. Conversely, assume that σ is additive, then

$$\sum_{B \subseteq A} (-1)^{|A \setminus B|} \sigma(B) = \sum_{B : a \in B \subseteq A} (-1)^{|A \setminus B|} \left[\sigma(a) + \sigma(B \setminus a)\right] + \sum_{B \subseteq A \setminus a} (-1)^{|A \setminus B|} \sigma(B) =$$

$$= \sum_{B : a \in B \subseteq A} (-1)^{|A \setminus B|} \sigma(a) + \sum_{B \subseteq A \setminus a} (-1)^{|A \setminus B|-1} \sigma(B) + \sum_{B \subseteq A \setminus a} (-1)^{|A \setminus B|} \sigma(B) = 0.$$

□

6. Discussion

6.1. **Relation to classical logit and a new characterization.** In this section we discuss the relationship to the classical Luce rule for stochastic choice functions. A stochastic choice function is a marginal SCR $\rho : X \times \mathcal{A} \to [0, 1]$ satisfying the adding-up constraint $\sum_{a \in A} \rho(a, A) = 1$. Thus, a stochastic choice function imposes the extra constraint that alternatives are perfect substitutes already from the outset. In this section we show that imposing this constraint implies that bundle utilities are zero and hence cannot be inferred from data. As a corollary to this result, and perhaps of independent interest, we obtain a new characterization of the classical Luce rule.

Proposition 6.1. *Let ρ be a stochastic choice function. Then ρ is a (signed) multivalued logit rule if and only if it is a univalent logit rule.*

Proof. Let ρ be a stochastic choice function represented by a multivalued logit with function σ. We claim that $\sigma(A) = 0$ for all $A \subseteq X$ with $|A| \geq 2$. Let $A \subseteq X$ with $|A| = 2$ so that $A = \{a, b\}$ for some $a, b \in X$. Then, since $\rho(a, b) + \rho(b, a) = 1$ we have

$$\frac{\sigma(a) + \sigma(b) + 2\sigma(\{a, b\})}{\sigma(a) + \sigma(b) + \sigma(\{a, b\})} = 1,$$

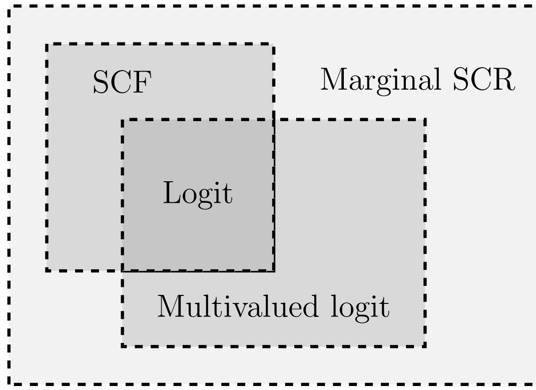

Figure 1. Figure showing the relationship between univalent and multivalued logit as established in proposition 6.1.

implying that $\sigma(\{a,b\}) = 0$. Assume that $\sigma(A) = 0$ for all $A \subseteq X$ with $2 \leq |A| \leq k$. Let $A \subseteq X$ with $|A| = k+1$. Then $\rho(a,A) = \frac{\sigma(A)+\sigma(a)}{\sigma(A)+\sum_{b\in A}\sigma(b)}$ for all $a \in A$. Since $\sum_{a\in A}\rho(a,A) = 1$, this implies that

$$\frac{|A|\sigma(A) + \sum_{a\in A}\sigma(a)}{\sigma(A) + \sum_{b\in A}\sigma(b)} = \sum_{a\in A}\frac{\sigma(A)+\sigma(a)}{\sigma(A)+\sum_{b\in A}\sigma(b)} = 1.$$

Thus $\sigma(A) = 0$, and the claim now follows by induction. □

As a corollary to proposition 6.1 we obtain a (to the best of our knowledge) new characterization of the classical Luce/logit rule.

Corollary 6.2. *A positive stochastic choice function ρ is a logit rule if and only if it satisfies path independence.*

6.2. **Zero probabilities in choice.** The preceding discussion, as summarized by corollary 6.2, implies that Luce's choice axiom and path independence are equivalent properties for *positive* stochastic choice functions. Recently, Cerreia-Vioglio, Lindberg, Maccheroni, Marinacci, and Rustichini (2021) obtain a characterization of a version of Luce's model that allows for zero probabilities in choice. They show that Luce's choice axiom is equivalent to a two stage model of choice, where in a first stage maximal alternatives are chosen using a weak order and in a second stage ties are broken using a standard Luce model. We conjecture that our path independence property is equivalent to Luce's choice axiom (for arbitrary, not necessarily positive, stochastic choice functions). It is straightforward to show that Luce's choice axiom implies path independence. To see this, note that $[1-\rho(a,A)][1-\rho(b,A\setminus a)] = [1-\rho(a,A)]\sum_{c\in A\setminus\{ab\}}\rho(c,A\setminus a) = \sum_{c\in A\setminus\{ab\}}\rho(c,A) = [1-\rho(b,A)]\sum_{c\in A\setminus\{ab\}}\rho(c,A\setminus b) = [1-\rho(b,A)][1-\rho(a,A\setminus b)]$.[5] The proof of the converse implication is less straightforward and has been confirmed for $|X|\leq 3$ (proof available on request).

Appendix A. Proofs

A.1. **Proof that the function f^ρ is well-defined.** Let $A\subseteq X$. For every enumeration of $X\setminus A$ as $\{x_1,\dots,x_k\}$ of X define

$$g(\{x_1,\dots,x_k\}) = \prod_{i=1}^{k}[1-\rho(x_i, X\setminus\{x_1,\dots,x_{i-1}\})],$$

where we use the convention that $X\setminus\{x_1,\dots,x_{i-1}\} = X$ for all $i\leq 1$. We define a binary relation $\sim$ on enumerations of $X\setminus A$ by $\{x_1,\dots,x_k\}\sim\{y_1,\dots,y_k\}$ if and only if $g(\{x_1,\dots,x_k\}) = g(\{y_1,\dots,y_k\})$. This relation is clearly transitive (since the canonical equality relation on real numbers is transitive). The following lemma shows that all pairs of sequences $\{x_1,\dots,x_k\}$ and $\{y_1,\dots,y_k\}$ are related by $\sim$ and hence that f^ρ is well-defined.

Lemma A.1. *Let $A\subseteq X$ and let $\{x_1,\dots,x_k\}$ and $\{y_1,\dots,y_k\}$ be enumerations of $X\setminus A$. If ρ is path independent then*

$$\prod_{i=1}^{k}[1-\rho(x_i, X\setminus\{x_1,\dots,x_{i-1}\})] = \prod_{i=1}^{k}[1-\rho(y_i, X\setminus\{y_1,\dots,y_{i-1}\})].$$

Proof. The proof is by induction on the cardinality k of $X\setminus A$. The base case $k=2$ follows directly by path independence. Suppose that the claim holds for all sets $A\subseteq X$ such that $|X\setminus A|\leq m$. Let A be a subset of X such that $|X\setminus A| = m+1$. Let $\{x_1,\dots,x_{m+1}\}$ and $\{y_1,\dots,y_{m+1}\}$ be enumerations of $X\setminus A$. If $x_{m+1} = y_{m+1}$ then the claim follows by the induction hypothesis. Assume that $x_{m+1}\neq y_{m+1} = a$. Then there is a $k < m+1$ with $x_k = a$. By path independence it follows that $\{x_1,\dots,x_k,x_{k+1},\dots,x_{m+1}\}\sim\{x_1,\dots,x_{k+1},x_k,\dots,x_{m+1}\}$. To see this note that

$$g(\{x_1,\dots,x_k,x_{k+1},\dots,x_{m+1}\}) =$$

[5] As discussed in subsection 4.2 Luce's choice axiom implies that choice probabilities sum to one and hence this implication even holds for marginal SCRs ρ.

$$\prod_{\substack{i=1\\ i\neq k, i\neq k+1}}^{m+1}[1-\rho(x_i,X\backslash\{x_1,\ldots,x_{i-1}\})][1-\rho(x_k,X\backslash\{x_1,\ldots,x_{k-1}\})][1-\rho(x_{k+1},X\backslash\{x_1,\ldots,x_{k-1},x_k\})] =$$

$$\prod_{\substack{i=1\\ i\neq k, i\neq k+1}}^{m+1}[1-\rho(x_i,X\backslash\{x_1,\ldots,x_{i-1}\})][1-\rho(x_{k+1},X\backslash\{x_1,\ldots,x_{k-1}\})][1-\rho(x_k,X\backslash\{x_1,\ldots,x_{k-1},x_{k+1}\})] =$$

$$g(\{x_1,\ldots,x_{k+1},x_k,\ldots,x_{m+1}\})$$

The last equality follows since $X\backslash\{x_1,\ldots,x_k,x_{k+1},\ldots,x_i\} = X\backslash\{x_1,\ldots,x_{k+1},x_k,\ldots,x_i\}$ for all $i > k+1$. Several applications of path independence then gives

$$\{x_1,\ldots,x_k,x_{k+1},\ldots,x_{m+1}\} \sim \{x_1,\ldots,x_{k+1},x_k,\ldots,x_{m+1}\} \sim$$

$$\{x_1,\ldots,x_{k+1},x_{k+2},x_k,\ldots,x_{m+1}\} \sim \cdots \sim \{x_1,\ldots,x_{m+1},x_k\}.$$

Since $x_k = y_{m+1}$ it follows by the induction hypothesis that $\{x_1,\ldots,x_{m+1},x_k\} \sim \{y_1,\ldots,y_{m+1}\}$. The claim then follows by transitivity of $\sim$. $\square$

A.2. **Proof of proposition 2.2.**

Proof. If ρ is a marginal SCR, then clearly

$$\sum_{a\in A}\rho(a,A) = \sum_{a\in A}\sum_{B:a\in B}R(B,A) \geq \sum_{B\subseteq A}R(B,A) = 1.$$

To prove the converse we will construct a stochastic choice rule R as follows. Let $A\in\mathcal{A}$. Order the alternatives in A as $\{a_1,\ldots,a_n\}$ such that $i>j$ imply $\rho(a_i,A)\geq\rho(a_j,A)$. For each $A\subseteq X$, define

$$\alpha(A) = \begin{cases} \frac{\sum_{a\in A}\rho(a,A)-1}{\sum_{a\in A}\rho(a,A)-\rho(a_n,A)} & \text{if } \rho(a_n,A)<1,\\ 1 & \text{otherwise.}\end{cases}$$

Then it is clear that $\alpha(A)>0$ for all $A\subseteq X$. Define $R(A,A)=\alpha(A)\rho(a_1,A)$ and $R(A\setminus\{a_1\},A)=\alpha(A)[\rho(a_2,A)-\rho(a_1,A)]$ and

$$R(A\setminus\{a_1,\ldots,a_{k-1}\},A) = \alpha(A)[\rho(a_k,A)-\rho(a_{k-1},A)]$$

for all $n > k \geq 3$. Set

$$R(\{a_k\}, A) = (1 - \alpha(A))\rho(a_k, A)$$

for all $1 \leq k < n$ and

$$R(\{a_n\}, A) = R(A\backslash\{a_1, \dots, a_{n-1}\}, A) = \alpha(A)[\rho(a_n, A) - \rho(a_{n-1}, A)] + (1-\alpha(A))\rho(a_n, A),$$

and $R(B, A) = 0$ otherwise. Then for all $k < n$ we have

$$\sum_{B:a_k \in B} R(B, A) = \sum_{i \leq k} R(A \setminus \{a_1, \dots, a_{i-1}\}, A) + R(\{a_k\}, A) =$$

$$= \sum_{i \leq k} \alpha(A)[\rho(a_i, A) - \rho(a_{i-1}, A)] + R(\{a_k\}, A) = \alpha(A)\rho(a_k, A) + (1-\alpha(A))\rho(a_k, A) = \rho(a_k, A),$$

and

$$\begin{aligned} \sum_{B:a_n \in B} R(B, A) &= \sum_{i \leq n} R(A \setminus \{a_1, \dots, a_{i-1}\}, A) = \\ &= \sum_{i \leq n-1} \alpha(A)[\rho(a_i, A) - \rho(a_{i-1}, A)] + R(\{a_n\}, A) = \\ = \alpha(A)\rho(a_{n-1}, A) &+ \alpha(A)[\rho(a_n, A) - \rho(a_{n-1}, A)] + (1 - \alpha(A))\rho(a_n, A) = \rho(a_n, A). \end{aligned}$$

Further, we note that

$$\sum_{B \subseteq A} R(B, A) = \sum_{i<n} R(A \setminus \{a_1, \dots, a_{i-1}\}, A) + \sum_{i<n} R(\{a_i\}, A) + R(\{a_n\}, A) = \tag{5}$$

$$\sum_{i<n} \alpha(A)[\rho(a_i, A) - \rho(a_{i-1}, A)] + \sum_{i<n} (1 - \alpha(A))\rho(a_i, A) + R(\{a_n\}, A) = \tag{6}$$

$$\alpha(A)\rho(a_n, A) + (1 - \alpha(A)) \sum_{i \leq n} \rho(a_i, A). \tag{7}$$

To show that that R is a stochastic choice rule, it only remains to prove that $\alpha(A)\rho(a_n, A) + (1-\alpha(A)) \sum_{i \leq n} \rho(a_i, A) = 1$. There are two cases. First, if $\rho(a_n, A) = 1$ then by definition we have $\alpha(A) = 1$ and we are done. Second, assume that $\rho(a_n, A) < 1$, then $\alpha(A) = \frac{\sum_{a \in A} \rho(a,A) - 1}{\sum_{a \in A} \rho(a,A) - \rho(a_n,A)}$ and rearranging this equality we obtain that $\alpha(A)\rho(a_n, A) + (1 - \alpha(A)) \sum_{i \leq n} \rho(a_i, A) = 1$. □

A.3. **Proof of theorem 3.2.** In the proof of theorem 3.2 we will employ a Möbius inversion formula for sets. Generally, Möbius inversion is a powerful combinatorial technique. The most general version of the Möbius inversion formula is formulated for partially ordered sets and it can be viewed as a generalization of the inclusion-exclusion formula. We will here apply a special case of the Möbius inversion formula for collections of sets partially ordered by set inclusion. The following statement is from Shafer (1976) (see also Brady and Rehbeck (2016, p.1215)).

Lemma A.2. *Let $f : 2^X \to \mathbb{R}$ and $g : 2^X \to \mathbb{R}$ be functions then*

$$f(A) = \sum_{B \subseteq A} g(B)$$

for all $A \subseteq X$ if and only if

$$g(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} f(B)$$

for all $A \subseteq X$.

Thus, the Möbius inversion formula allows the underlying distribution function g to be recovered when the cumulative distribution function g is available (and conversely).

Before proving theorem 3.2 we will show that path independence and varied support characterize a more general class of logit rules with negative σ allowed. A stochastic choice function ρ is called a *signed multivalued logit* if it satisfies equation (1) with a function $\sigma : 2^X \to (-\infty, \infty)$ (such that $\sigma(\{a\}) > 0$ for all $a \in A$). Most of our results, including theorem 3.2 will follow as corollaries to lemma A.3 below.

Lemma A.3. *A marginal SCR is a signed multivalued logit rule if and only if it satisfies varied support and path independence.*

Proof. For each $A \subseteq X$ define

$$\sigma(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} f^{\rho}(B),$$

where f^ρ is defined as in equation (3). Let $a \in A$ and $\{x_1, \ldots, x_k\} = X \setminus a$ an enumeration of $X \setminus a$. Then it follows that

$$\sigma(\{a\}) = f^\rho(\{a\}) = \prod_{i=1}^{k} [1 - \rho(x_i, X \setminus \{x_1, \ldots, x_{i-1}\})] > 0. \tag{8}$$

The strict inequality in equation (8) follows since $\rho(a, A) < 1$ for all $a \in A$ and $A \subsetneq X$ by varied support. The Möbius inversion formula in lemma A.2 gives

$$f^\rho(A) = \sum_{B \subseteq A} \sigma(B) \tag{9}$$

for all $A \subseteq X$. We use this formula to derive the desired representation of ρ. Let $A \subseteq X$ and $a \in A$. Let $\{x_1, \ldots, x_m\} = X \setminus A$ be an enumeration of $X \setminus A$ and let $x_{m+1} = a$. Then

$$[1 - \rho(a, A)] f^\rho(A) = [1 - \rho(x_{m+1}, X \setminus \{x_1, \ldots, x_m\})] f^\rho(A) =$$
$$\prod_{i=1}^{m+1} [1 - \rho(x_i, X \setminus \{x_1, \ldots, x_{i-1}\})] = f^\rho(A \setminus a).$$

By the previous formula, and since $f^\rho(\{a\}) > 0$, it follows by an induction argument that $f^\rho(A) > 0$ for all $A \subseteq X$. Solving for $\rho(a, A)$ (which is possible since $f^\rho(A) > 0$) we thus obtain that

$$\rho(a, A) = \frac{f^\rho(A) - f^\rho(A \setminus a)}{f^\rho(A)} = \frac{\sum\limits_{B \subseteq A : a \in B} \sigma(B)}{\sum\limits_{B \subseteq A} \sigma(B)},$$

where the last equality follows by equation (9). We next show that σ is unique up to multiplication by a scalar α ($\alpha \geq 0$ if σ is positive). Let σ and σ' be two logit representations of ρ. Then

$$\frac{\sigma'(A)}{\sum_{B \subseteq X} \sigma'(B)} = \delta f^\rho(A) = \frac{\sigma(A)}{\sum_{B \subseteq X} \sigma(B)}$$

for all $A \subseteq X$. Set $\alpha = \frac{\sum_{B \subseteq X} \sigma'(B)}{\sum_{B \subseteq X} \sigma(B)}$ and it is then clear that $\sigma'(A) = \alpha \sigma(A)$ for all $A \subseteq X$. $\square$

We next derive theorem 3.2 as a corollary to proposition A.3.

Proof. Let ρ satisfy path independence and positivity. Define σ as in the proof of proposition A.3 above. By positivity it follows that $\sigma(A) \geq 0$ for all $A \subseteq X$. The remainder of the proof is similar to the proof of proposition A.3. $\square$

A.4. **Proof of proposition 4.1.** The following simple lemma is needed in the proof of proposition 4.1

Lemma A.4. *We have that $\delta f^\rho(ab) \geq 0$ for all $a, b \in X$.*

Proof. First note that $f^\rho(a) = (1-\rho(b,a))f^\rho(ab)$ and $f^\rho(b) = (1-\rho(a,b))f^\rho(ab)$. We then have that

$$\delta f^\rho(ab) = f^\rho(ab) - f^\rho(a) - f^\rho(b) =$$

$$= f^\rho(ab) - (1-\rho(b,a))f^\rho(ab) - (1-\rho(a,b))f^\rho(ab) =$$

$$= f^\rho(ab)[\rho(a,b) + \rho(b,a) - 1] \geq 0.$$

Clearly, $\delta f^\rho(a) = f^\rho(a) \geq 0$ for all $a \in X$. Thus, the claim follows. $\square$

We are now ready for the proof of proposition 4.1

Proof. By bundle LCA, hence path independence, and varied support there is by proposition A.3 a function $\sigma : 2^X \to (-\infty, \infty)$ such that

$$\rho(a, A) = \frac{\sum\limits_{B \subseteq A : a \in B} \sigma(B)}{\sum\limits_{B \subseteq A} \sigma(B)}$$

for all $a \in A$ and $A \subseteq X$. We next need to show that $\sigma(a,b) \geq 0$ fo all $a, b \in X$ and that $\sigma(A) = 0$ for all $A \subseteq X$ with $|A| \geq 3$. That $\sigma(a,b) \geq 0$ for all $a, b \in X$ follows from lemma A.4 and since ρ satisfies varied support (so $\delta f^\rho(a) > 0$ for all $a \in A$). We next show that $\sigma(A) = 0$ for all $A \subseteq X$ with $|A| \geq 3$. The proof is by induction on the cardinality of A. As a base case, let $A \subseteq X$ with $|A| = 3$, so that $A = \{a, b, c\}$ for some $a, b, c \in X$. We have that:

$$\frac{\sigma(abc) + \sigma(ab)}{\sum_{B \subseteq abc} \sigma(B)} = \rho(a, abc) - (1 - \rho(b, abc))\rho(a, ac) = O(a, b, A) =$$

$$[1-\rho(c, abc)]O(a, b, \{a, b\}) = [1-\rho(c, abc)][\rho(a, ab) - (1-\rho(b, ab))\rho(a, a)] = \frac{\sigma(ab)}{\sum_{B \subseteq abc} \sigma(B)}.$$

Thus, it follows that $\sigma(abc) = 0$. Assume that $\sigma(A) = 0$ for all $A \subset X$ with $|A| \leq k$. Let $A \subseteq X$ and $|A| = k+1$. Pick $a, b, c \in A$. By bundle LCA it then follows that

$$\frac{\sigma(A) + \sigma(ab)}{\sum_{B \subseteq A} \sigma(B)} = \rho(a, A) - (1 - \rho(b, A))\rho(a, A \setminus b) = O(a, b, A) =$$

$$[1-\rho(c, A)]O(a, b, A \setminus c) = [1-\rho(c, A)][\rho(a, A \setminus c) - (1-\rho(b, A \setminus c))\rho(a, A \setminus bc)] = \frac{\sigma(ab)}{\sum_{B \subseteq A} \sigma(B)}.$$

The first and last equality follows by the induction hypothesis. The second equality follows by bundle LCA. By the equation above we hence have that $\sigma(A) = 0$ and the claim follows. $\square$

References

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

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

R. Allen and J. Rehbeck. Identification with additively separable heterogeneity. *Econometrica*, 87(3):1021–1054, 2019.

R. Allen and J. Rehbeck. Latent complementarity in bundles models. *Journal of Econometrics*, 228(2):322–341, 2022.

Y. Azrieli and J. Rehbeck. Marginal stochastic choice. *arXiv preprint arXiv:2208.08492*, 2022.

H. Block and J. Marschak. Random orderings and stochastic theories of responses. In *Olkin, I., Gurye, S.G., Hoeffding, W., Madow, W.G., Mann, H.B. (Eds.), Contributions to Probability and Statistics. Stanford University Press, Stanford, CA.*, pages 172–217. 1960.

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

S. Cerreia-Vioglio, P. O. Lindberg, F. Maccheroni, M. Marinacci, and A. Rustichini. A canon of probabilistic rationality. *Journal of Economic Theory*, 196:105289, 2021.

M. Costa-Gomes, C. Cueva, and G. Gerasimou. Choice, deferral and consistency. *Quantitative Economics, forthcoming*, 2021.

M. Gentzkow. Valuing new goods in a model with complementarity: Online newspapers. *American Economic Review*, 97(3):713–744, 2007.

G. Gerasimou. Eliciting weak or incomplete preferences in the lab: A model-rich approach. *Working Paper*, 2022.

A. Iaria and A. Wang. Identification and estimation of demand for bundles. 2020.

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

P. Manzini, M. Mariotti, and L. Ülkü. Stochastic complementarity. *The Economic Journal*, 129(619):1343–1363, 2019.

P. Manzini, M. Mariotti, and L. Ülkü. A model of approval, 2022.

H. Petri. Random (ordered) multivalued choice. 2022.

D. Ravid and K. Steverson. Focus, then compare. 2019.

G. Shafer. *A mathematical theory of evidence*, volume 42. Princeton university press, 1976.

A. Wang. A blp demand model of product-level market shares with complementarity. *Available at SSRN 3785148*, 2021.

K. Wang. Logit function in stochastic categorizations. *Oxford Economic Papers*, 74 (2):610–622, 2022.