

# Law of Demand and Stochastic Choice\*

S. Cerreia-Vioglio<sup>a</sup>, F. Maccheroni<sup>a</sup>, M. Marinacci<sup>a</sup>, and A. Rustichini<sup>b</sup>

<sup>a</sup>Università Bocconi and IGIER

<sup>b</sup>University of Minnesota

Draft of June 2018

## Abstract

We characterize consistent random choice rules in terms of the optimality of their support. We then study stochastic choice in a consumer theory setting. We prove a law of individual demand for stochastic choice that generalizes the standard deterministic law.

---

\*Part of the material of this paper was first circulated under the title “Law of Demand and Forced Choice”, IGIER WP 593, 2016. We thank Chris Chambers, Federico Echenique, Ryota Iijima, Kota Saito, and Tomasz Strzalecki for helpful discussions. Simone Cerreia-Vioglio and Massimo Marinacci gratefully acknowledge the financial support of ERC (grants SDDM-TEA and INDIMACRO, respectively).

# 1 Introduction

Choices are often stochastic. The strands of literature that in different disciplines – such as economics, neuroscience and psychology – have dealt with stochasticity in choices have mostly focused on pure choice behavior. Here we study how stochastic choice may affect consumer behavior, the most basic economic choice problem.

We first characterize, in Theorem 1, random choice rules that satisfy the assumption of consistency, which is a form of Luce’s choice axiom (see, Luce 1959). Differently from Luce, we do not require that these rules have full support: a conceptually problematic assumption in a consumer theory framework where it is natural to assume that dominated options will never be selected.

We then show that for this important class of stochastic choices, the law of demand for normal goods – arguably the main result of traditional consumer theory – continues to hold on average when strictly dominated alternatives are dismissed. A “certainty equivalence” principle for stochastic consumer theory thus emerges from our analysis. It ensures that choice stochasticity a la Luce does not alter qualitatively the findings of traditional deterministic consumer theory – which is, indeed, a special case of our analysis. All foundations of Luce’s rule proposed in the literature, be they behavioral or neural, thus share this significant economic consequence.<sup>1</sup>

Finally, some relevant references are discussed as the analysis develops.

## 2 Preliminaries: random choice rules and optimality

Let  $\mathcal{A}$  be the collection of all non-empty finite choice sets  $A$  of an all inclusive set of alternatives  $X$ . Each choice set  $A$  represents a decision problem in which an agent has to choose an alternative in  $A$ .<sup>2</sup>

**Definition 1** *A random choice rule is a function  $p : \mathcal{A} \rightarrow \Delta(X)$  such that  $p(\cdot, A) \in \Delta(A)$  for all  $A \in \mathcal{A}$ .*

We interpret  $p(a, A)$  as the probability that an agent chooses alternative  $a \in A$  within the choice set  $A$ . In a (ergodic) long run setup, this probability can be viewed as the long run frequency with which  $a$  is chosen across repetitions of the decision

---

<sup>1</sup>See, e.g., Woodford (2014) and Matejka and McKay (2015) for recent foundations of Luce’s rule, as well as Krajbich and Dean (2015) for a recent overview.

<sup>2</sup>Throughout the paper, for each set  $E \subseteq X$ , finite or not, we denote by  $\Delta(E)$  the set of all finitely supported probabilities on  $E$ . So,  $p(\cdot, A) \in \Delta(A)$  if and only if  $p(\cdot, A) : A \rightarrow [0, 1]$ , with  $p(a, A) > 0$  for finitely many  $a$  in  $A$ , and  $\sum_{a \in A} p(a, A) = 1$ . With a slight abuse of notation, we regard  $\Delta(A)$  as included in  $\Delta(X)$ , since the former is naturally embedded in the latter.

problem represented by  $A$ . We regard  $p$  as a purely behavioral notion that accounts for agent's choices, a way to organize choice data without any mental interpretation per se. Our analysis is thus in the spirit of Pareto (1900) and Samuelson (1938).

**Definition 2** A random choice rule  $p$  is consistent if, for all  $a \in A \subseteq B$ ,<sup>3</sup>

$$p(a, B) = p(a, A) p(A, B) \quad (1)$$

This condition is a form of the classic Luce's choice axiom (see Luce, 1959). It ensures that  $p(\cdot, A)$  and  $p(\cdot, B)$  are linked via conditioning a la Renyi (1955). We denote by  $\sigma_p : \mathcal{A} \rightrightarrows X$  the support correspondence defined by  $\sigma_p(A) = \{a \in A : p(a, A) > 0\}$ .

**Example 1 (Luce)** Given a function  $\varphi : X \rightarrow (0, \infty)$ , define  $p : \mathcal{A} \rightarrow \Delta(X)$  by for each  $A \in \mathcal{A}$

$$p(a, A) = \frac{\varphi(a)}{\sum_{b \in A} \varphi(b)} \quad \forall a \in A \quad (2)$$

This function  $p$  is a consistent random choice rule with full support, i.e.,  $\sigma_p(A) = A$  for all  $A \in \mathcal{A}$ . Luce (1959) proved that this is the general form of consistent random choice rules that have full support. In particular, the uniform rule  $p(a, A) = 1/|A|$  is the special case with  $\varphi = 1$ . ▲

**Example 2 (Optimization)** A correspondence  $\sigma : \mathcal{A} \rightrightarrows X$  is a *choice correspondence* if  $\emptyset \neq \sigma(A) \subseteq A$  for all  $A \in \mathcal{A}$ . By a classic result of Arrow (1959),  $\sigma$  is *optimal* (or *rational*) – i.e., it represents alternatives that are optimal according to some weak order – if and only if it satisfies the following version of WARP:

$$A \subseteq B \text{ and } \sigma(B) \cap A \neq \emptyset \implies \sigma(B) \cap A = \sigma(A) \quad (C)$$

Given an optimal choice correspondence  $\sigma : \mathcal{A} \rightrightarrows X$ , the function  $p : \mathcal{A} \rightarrow \Delta(X)$  defined by

$$p(a, A) = \begin{cases} \frac{1}{|\sigma(A)|} & \text{if } a \in \sigma(A) \\ 0 & \text{else} \end{cases} \quad (3)$$

is easily seen to be a consistent random choice rule – without full support and not in the Luce-Renyi form unless  $\sigma(A) = A$  for all  $A \in \mathcal{A}$ . When  $\sigma$  is an optimal *choice function*, so  $\sigma(A)$  is a singleton for all  $A \in \mathcal{A}$ , this rule takes the deterministic form:<sup>4</sup>

$$p(a, A) = \delta_{\sigma(A)}(a) \quad \forall a \in A$$

Optimal choice functions can thus be viewed as special, deterministic, rules of the form (3). ▲

---

<sup>3</sup>Given  $A \subseteq B$ , we denote by  $p(A, B)$  the quantity  $\sum_{a \in A} p(a, B)$  with the convention that  $p(\emptyset, B) = 0$ .

<sup>4</sup>Here  $\delta_x$  denotes the (Dirac) probability at  $x \in X$ , that is,  $\delta_x(a) = 1$  if  $x = a$  and 0 otherwise.

In view of the previous example, we say that a choice correspondence  $\sigma$  is *optimal* if it satisfies (C). This example seems to provide a very specific rule, (3), where optimality of  $\sigma$  implies consistency. The next result, proved in Appendix A, makes this observation formal and much more general. Indeed, it characterizes consistent random choice rules in terms of the optimality of their support correspondence.

**Theorem 1** *A function  $p : \mathcal{A} \rightarrow \Delta(X)$  is a consistent random choice rule if and only if it has the form*

$$p(a, A) = \begin{cases} \frac{\varphi(a)}{\sum_{b \in \sigma(A)} \varphi(b)} & \text{if } a \in \sigma(A) \\ 0 & \text{else} \end{cases} \quad (4)$$

where  $\varphi : X \rightarrow (0, \infty)$  and  $\sigma : \mathcal{A} \rightrightarrows X$  is an optimal choice correspondence. Moreover,  $\sigma$  is unique and coincides with  $\sigma_p$ .

This characterization shows that consistency of random choice rules is the stochastic counterpart of the optimality behavior, characterized by WARP, for choice correspondences. Moreover, Luce's rule (2) corresponds to the special case  $\sigma(A) = A$  for all  $A \in \mathcal{A}$  (which is trivially optimal), while the optimization rule (3) is the special case in which either  $\varphi = 1$  or  $\sigma$  is single-valued. So, the two previous examples are both special cases of the random choice rule (4).<sup>5</sup> In light of Theorem 1, it is also immediate to show that Luce's rule is characterized by consistency plus the following weak form of positivity:  $p(a, \{a, b\}) > 0$  for all  $a, b \in X$ . In words, the agent must always select with strictly positive probability each element from the set  $\{a, b\}$ .

In keeping with our behavioral approach, we do not give any mental interpretation of Theorem 1. For us, it is a characterization of consistent random choice rules that, in particular, shows that optimal choice functions are a special, deterministic, class of such rules. Because of this remarkable property, our study of stochastic choice in a consumer theory framework – our main object of interest – will be able to generalize the standard deterministic theory.

A final remark. A random choice rule  $p$  is *uniform* if, given any  $A \in \mathcal{A}$ , all alternatives in  $A$  that have a chance to be chosen are equally likely. Formally, for each  $a \in A$ , we have

$$p(a, A) = \begin{cases} \frac{1}{|\sigma_p(A)|} & \text{if } a \in \sigma_p(A) \\ 0 & \text{else} \end{cases}$$

This uniform rule is close in spirit to the analysis of Becker (1962). By Theorem 1, a uniform random choice rule  $p$  is consistent if and only if its support correspondence

---

<sup>5</sup>Theorem 1 also shows that consistent random choice rules are a special case of what Echenique and Saito (2015) call general Luce rules. These rules, axiomatically characterized by Echenique and Saito (2015), take the form (4) where  $\varphi : X \rightarrow (0, \infty)$  but the choice correspondence  $\sigma$  is not necessarily optimal.

$\sigma_p$  is optimal. This equivalence completes the analysis of Example 2, in which the “if” was considered.

### 3 Random consumption

Our aim here is to develop a behavioral (i.e., non-preferential, so “non-mental”) consumer theory in a random choice setting that, *inter alia*, encompasses as a special case the traditional deterministic behavioral consumer theory presented, for example, in chapter 2 of Mas-Colell et al. (1995). A relevant related work is Mossin (1968), which outlined a stochastic theory of consumption (with a different framework and motivation). A less relevant related work is Halldin (1974). It proposes an alternative theory which is a mixture of stochastic choice over pairs and choice over budget sets via a propensity function.

#### 3.1 Individual stochastic demand

Let  $X = \mathbb{R}_+^n$  be the space of all bundles of goods and  $B : \mathbb{R}_+^n \times \mathbb{R}_{++} \rightrightarrows X$  the budget correspondence defined by  $B(q, w) = \{x \in X : q \cdot x \leq w\}$  for each price and wealth pair  $(q, w)$ . Now  $\mathcal{A}$  is replaced with a larger class  $\mathcal{B}$  that contains  $\mathcal{A}$  and all budget sets  $B(q, w)$ . So, in this section  $\mathcal{B}$  is the domain of our analysis.<sup>6</sup>

**Definition 3** A function  $d : \mathbb{R}_+^n \times \mathbb{R}_{++} \rightarrow \Delta(X)$  is an (individual) stochastic demand induced by a consistent random choice rule  $p : \mathcal{B} \rightarrow \Delta(X)$  if  $d(q, w)(x) = p(x, B(q, w))$ .

We interpret  $d(q, w)(x)$  as the probability that bundle  $x \in B(q, w)$  is chosen at price  $q$  with wealth  $w$ . In particular, the average cost function  $c : \mathbb{R}_+^n \times \mathbb{R}_{++} \rightarrow \mathbb{R}_+$  of the bundle demanded is  $c(q, w) = \sum_{x \in B(q, w)} (q \cdot x) d(q, w)(x)$ , while the (individual) average demand function  $\bar{d} : \mathbb{R}_+^n \times \mathbb{R}_{++} \rightarrow \mathbb{R}_+^n$  is  $\bar{d}(q, w) = \sum_{x \in B(q, w)} x d(q, w)(x)$ .<sup>7</sup> Clearly,  $c(q, w) = q \cdot \bar{d}(q, w)$ . In what follows, we will consider only stochastic demands  $d$  induced by consistent random choice rules  $p$ . For such a reason, we omit to mention  $p$ .

The stochastic demand  $d(q, w)$  has finite support, i.e., only finitely many bundles of the budget set  $B(q, w)$  can be selected with strictly positive probability. This assumption eases our analysis, yet still substantially broadens the scope of the traditional deterministic analysis that, typically, assumes unique optimal bundles to best carry out

<sup>6</sup>See Appendix B for more details on this richer domain. Theorem 1 can be easily extended to this domain, as we show in Theorem 12 of Appendix C.

<sup>7</sup>Observe that  $x$  is a vector while  $d(q, w)(x)$  is a scalar. Thus, we slightly abuse notation by denoting the scalar product of these two objects by inverting the role of scalars and vectors.

comparative statistics exercises. In this regard, note that if  $p$  is the random choice rule in (4), then

$$d(q, w)(x) = \begin{cases} \frac{\varphi(x)}{\sum_{y \in \sigma(B(q, w))} \varphi(y)} & \text{if } x \in \sigma(B(q, w)) \\ 0 & \text{else} \end{cases} \quad (5)$$

provided  $\sigma$  is an optimal choice correspondence that is finitely valued, i.e.,  $\sigma(B) \in \mathcal{A}$  for all  $B \in \mathcal{B}$ . This example is important for our analysis because it shows that standard demand functions are included in our setup: when  $\sigma$  is a choice function, the stochastic demand (5) becomes

$$d(q, w)(x) = \delta_{\sigma(B(q, w))}(x) \quad \forall x \in B(q, w)$$

and so it is a classic Walrasian demand function. That said, if  $\varphi = 1$  the stochastic demand (5) takes a uniform form close in spirit to the analysis of Becker (1962), as already remarked. Besides standard demand functions, also uniform demand functions a la Becker (1962) are thus included in our setup.

### 3.2 Walras' law

Given a stochastic demand, by construction we have only  $c(q, w) \leq w$ . So, in general only the following weak form of Walras' law holds.

**Proposition 2** *Let  $d$  be a stochastic demand. If  $w < w'$ , then  $c(q, w) \leq c(q, w')$ .*

Equality, and more, holds under the following monotonicity condition.

**Definition 4** *A random choice rule  $p : \mathcal{B} \rightarrow \Delta(X)$  is stochastically monotone if  $p(x, \{x, y\}) = 0$  whenever  $x \ll y$ .<sup>8</sup>*

Comparisons between two bundles of goods  $x$  and  $y$  are much easier to make when one of them is strictly dominant, say  $x \ll y$ . Stochastic monotonicity captures this comparative easiness.

Under stochastic monotonicity, only alternatives that are not strictly dominated may have a chance to be selected. Specifically, for any choice set  $B \in \mathcal{B}$  let

$$\partial^+ B = \{x \in B : \nexists x' \in B, x' \gg x\}$$

be the collection of all elements of  $B$  that are not strictly dominated.

**Lemma 3** *If a consistent random choice rule  $p : \mathcal{B} \rightarrow \Delta(X)$  is stochastically monotone, then  $p(\partial^+ B, B) = 1$  for all  $B \in \mathcal{B}$ .*

---

<sup>8</sup>As usual,  $x \ll y$  means that  $x_i < y_i$  for all  $i = 1, \dots, n$ .

A stochastic demand  $d$  is *stochastically monotone* if it is induced by a stochastically monotone consistent random choice rule  $p$ . Since  $\partial^+ B(q, w) = \{x \in B(q, w) : q \cdot x = w\}$ , by the last lemma we have  $d(q, w)(x) > 0$  only if  $x \in \partial^+ B(q, w)$ . In turn, this implies the following sharp Walras' law for stochastic choice: under stochastic monotonicity, consumers always exhaust their budgets.<sup>9</sup>

**Proposition 4** *If the stochastic demand  $d$  is stochastically monotone, then  $c(q, w) = w$  (so,  $w < w'$  implies  $c(q, w) < c(q, w')$ ).*

By construction, both the stochastic and the average demands are homogeneous of degree zero, so there is no nominal illusion. In our analysis the consumer is always able to assess correctly whether a bundle of goods is affordable.

### 3.3 Law of demand

We can now study wealth and price effects. As to wealth effects, say that a good  $k$  is *normal* if its average demand increases as wealth increases:

$$w' > w \implies \bar{d}_k(q, w') \geq \bar{d}_k(q, w) \quad \forall q \in \mathbb{R}_{++}^n$$

By Proposition 4, under stochastic monotonicity, we have  $q \cdot \bar{d}(q, w') > q \cdot \bar{d}(q, w)$  if  $w' > w$ . So, intuitively, some of the goods have to be normal, at least locally. As to price effects, we have the following preliminary result.

**Lemma 5** *Let  $d$  be a stochastic demand. If  $q < q'$ , then  $\bar{d}(q', w) \not\geq \bar{d}(q, w)$ .*

Next we show that a classic compensated law of demand continues to hold “on average”.

**Lemma 6** *Let  $(q', w')$  and  $(q, w)$  be in  $\mathbb{R}_{++}^n \times \mathbb{R}_{++}$  and let  $d$  be a stochastic demand. If  $q' \cdot \bar{d}(q, w) = c(q', w')$  and  $d$  is stochastically monotone, then*

$$(q' - q) \cdot (\bar{d}(q', w') - \bar{d}(q, w)) \leq 0 \tag{6}$$

When a sharp Walras' law holds, condition  $q' \cdot \bar{d}(q, w) = c(q', w') = w'$  becomes a standard Slutsky wealth compensation. In this case the (individual) law of demand for normal goods – arguably the most important result of consumer theory – continues to hold on average.

**Theorem 7 (Law of Average Demand)** *Let the stochastic demand  $d$  be stochastically monotone. If wealth and other prices do not change, an increase (decrease) in the price of a normal good  $k$  decreases (increases) its average demand  $\bar{d}_k$ .*

---

<sup>9</sup>For instance, Gabaix (2014) assumes that the consumer he studies “is boundedly rational, but smart enough to exhaust his budget.”

Under consistency, on average the behavior of consumers continues to satisfy the law of demand for normal goods provided they are able to select strictly dominant alternatives. The standard consumer theory result for Walrasian demand functions, first stated on p. 14 of Slutsky (1915), is the special case that corresponds to the deterministic demand function  $d(q, w) = \delta_{\sigma(B(q,w))}$ . Indeed, as previously remarked, in this case its average demand is the Walrasian demand function, i.e.,  $\bar{d}(q, w) = \sigma(q, w)$ . Our result thus generalizes the most important finding of classical behavioral consumer theory. At the same time, it goes well beyond that. For instance, it includes the purely random choice of Becker (1962) (viewed as the uniform case).

To sum up, choice stochasticity a la Luce, along with a monotonicity property, leads to an individual demand that on average has the same qualitative properties of the special, yet standard, deterministic case. A “certainty equivalence” principle for stochastic consumer theory results.

A final remark. As we emphasized throughout the paper, our purely behavioral analysis has abstracted from any preferential, so mental, notion. It is then natural to close with a few words on a possible, complementary, preferential approach to stochastic consumer theory that, instead, may generalize the classical preferential consumer theory presented, for example, in chapter 3 of Mas-Colell et al. (1995). To this end, a random utility interpretation of random choice rules can be adopted, in which there exists a probability measure  $\pi$ , a stochastic preference, defined on a class  $\mathcal{P}$  of preference relations on  $X$  such that  $p(a, B) = \pi(P \in \mathcal{P} : \forall b \in B, aPb)$ . Via this representation, some results of classical preferential consumer theory that hold for elements of  $\mathcal{P}$  may have stochastic counterparts. An analysis along these lines is, however, beyond the scope of the present paper and is left for future research.<sup>10</sup>

## A Appendix: proof of Theorem 1

The main goal of this appendix is to prove Theorem 1. We start by introducing some new notation and a preparatory result. Recall that for each  $a, b \in X$

$$p(a, b) = p(a, \{a, b\}) \quad \text{and} \quad r(a, b) = \frac{p(a, b)}{p(b, a)}$$

**Lemma 8** *If  $p : \mathcal{A} \rightarrow \Delta(X)$  is a consistent random choice rule, then  $\sigma_p$  is optimal.*

---

<sup>10</sup>In this vein, Bandyopadhyay et al. (1999) study a stochastic demand which is generated by a random utility model. This alternative approach might also have an intersection (mainly technical) with the literature that studies average demand in a large economy (see, for a review, Jerison and Quah, 2006).



**Proof** Clearly,  $\emptyset \neq \sigma_p(A) \subseteq A$  for all  $A \in \mathcal{A}$ . Let  $A, B \in \mathcal{A}$  such that  $A \subseteq B$ . Assume that  $\sigma_p(B) \cap A \neq \emptyset$ . We want to show that  $\sigma_p(B) \cap A = \sigma_p(A)$ . Since  $p$  is consistent, if  $a \in \sigma_p(B) \cap A$ , then  $0 < p(a, B) = p(a, A)p(A, B)$ . It follows that  $p(a, A) > 0$ , that is,  $a \in \sigma_p(A)$ . Thus,  $\sigma_p(B) \cap A \subseteq \sigma_p(A)$ . As to the converse inclusion, let  $a \in \sigma_p(A) \subseteq A \subseteq B$ , that is,  $p(a, A) > 0$ . By contradiction, assume that  $a \notin \sigma_p(B)$ , that is,  $p(a, B) = 0$ . Since  $p$  is consistent, we then have  $0 = p(a, B) = p(a, A)p(A, B)$ . Since  $p(a, A) > 0$ , this implies that  $p(A, B) = 0$ , that is,  $\sigma_p(B) \cap A = \emptyset$ . This contradicts  $\sigma_p(B) \cap A \neq \emptyset$ , proving the opposite inclusion and completing the proof. ■

**Proof of Theorem 1** “If”. Let  $p$  be given by (4) with  $\sigma$  optimal and  $\varphi : X \rightarrow (0, \infty)$ . Since  $\varphi$  is strictly positive and  $\sigma$  a choice correspondence,  $p$  is a well defined random choice rule. Let  $A, B \in \mathcal{A}$  such that  $A \subseteq B$  and  $a \in A$ . We have two cases:

1.  $\sigma(B) \cap A \neq \emptyset$ . Since  $\sigma$  is optimal,  $\sigma(B) \cap A = \sigma(A)$ . On the one hand, by (4), if  $a \in \sigma(A)$ , then  $a \in \sigma(B)$  and  $p(a, A) = \varphi(a) / \sum_{b \in \sigma(A)} \varphi(b)$ . We can conclude that

$$p(a, B) = \frac{\varphi(a)}{\sum_{b \in \sigma(B)} \varphi(b)} = \frac{\varphi(a)}{\sum_{b \in \sigma(A)} \varphi(b)} \frac{\sum_{b \in \sigma(B) \cap A} \varphi(b)}{\sum_{b \in \sigma(B)} \varphi(b)} = p(a, A)p(A, B)$$

On the other hand, if  $a \notin \sigma(A)$ , we have that  $a \in A \setminus \sigma(B)$ , so  $p(a, A) = 0 = p(a, B)$ . In both cases (1) holds.

2.  $\sigma(B) \cap A = \emptyset$ . It follows that  $a \notin \sigma(B)$  and  $p(A, B) = 0 = p(a, B)$ . Again, (1) holds.

Cases 1 and 2 prove that  $p$  is consistent.

“Only if”. Let  $p : \mathcal{A} \rightarrow \Delta(X)$  be a consistent random choice rule and set

$$a \succsim b \iff a \in \sigma_p(\{a, b\}) \tag{7}$$

By Lemma 8,  $\sigma_p$  is optimal. Since  $\sigma_p$  is optimal,  $\succsim$  is a weak order and

$$\begin{aligned} a \succsim b &\iff p(a, b) > 0 \\ a \succ b &\iff p(a, b) = 1 \\ a \sim b &\iff p(a, b) \in (0, 1) \end{aligned}$$

By Arrow (1959) and since  $\sigma_p$  is optimal, observe that

$$\sigma_p(A) = \{a \in A : a \succsim b \quad \forall b \in A\} \quad \forall A \in \mathcal{A} \tag{8}$$

Let  $\{X_i : i \in I\}$  be the family of all equivalence classes of  $X$  with respect to  $\sim$ . For each  $i \in I$  choose  $a_i \in X_i$ . Define  $\varphi : X \rightarrow (0, \infty)$  to be such that

$$\varphi(x) = r(x, a_i) = \frac{p(x, a_i)}{p(a_i, x)} \quad \forall x \in X_i, \forall i \in I \tag{9}$$

Consider  $i \in I$  and let  $x \sim a_i$ , we have that  $p(x, a_i), p(a_i, x) \in (0, 1)$ . Since  $i$  and  $x$  were arbitrarily chosen, it follows that  $\varphi$  is well defined. By (8), we have that

$$\sigma_p(S) = S \quad \forall S \in \mathcal{A} \text{ such that } S \subseteq X_i \text{ for some } i \in I \quad (10)$$

Since  $p$  is consistent, this implies that for each  $S \in \mathcal{A}$  such that  $S \subseteq X_i$  for some  $i \in I$  and for each  $a, b \in S$

$$\begin{aligned} p(a, S) &= p(a, b) p(\{a, b\}, S) > 0 \\ p(b, S) &= p(b, a) p(\{a, b\}, S) > 0 \end{aligned}$$

yielding that

$$\frac{p(a, S)}{p(b, S)} = \frac{p(a, b)}{p(b, a)} = r(a, b) \quad (11)$$

Next, consider  $a, b \in X$  such that  $a \sim b$ . We have that there exists  $i \in I$  such that  $a \sim b \sim a_i$ . By consistency and definition of  $\varphi$ , we can conclude that

$$\begin{aligned} p(a, \{a, b, a_i\}) &= p(a, a_i) p(\{a, a_i\}, \{a, b, a_i\}) = \frac{p(a, a_i)}{p(a_i, a)} p(a_i, a) p(\{a, a_i\}, \{a, b, a_i\}) \\ &= \frac{p(a, a_i)}{p(a_i, a)} p(a_i, \{a, b, a_i\}) = \varphi(a) p(a_i, \{a, b, a_i\}) \\ p(b, \{a, b, a_i\}) &= p(b, a_i) p(\{b, a_i\}, \{a, b, a_i\}) = \frac{p(b, a_i)}{p(a_i, b)} p(a_i, b) p(\{b, a_i\}, \{a, b, a_i\}) \\ &= \frac{p(b, a_i)}{p(a_i, b)} p(a_i, \{a, b, a_i\}) = \varphi(b) p(a_i, \{a, b, a_i\}) \end{aligned}$$

By (10), we have that  $p(a_i, \{a, b, a_i\}) > 0$  and  $p(b, \{a, b, a_i\}) > 0$ . By applying (11) twice, we can conclude that

$$\frac{p(a, S)}{p(b, S)} = r(a, b) = \frac{p(a, \{a, b, a_i\})}{p(b, \{a, b, a_i\})} = \frac{\varphi(a)}{\varphi(b)} \quad \forall S \in \mathcal{A} \text{ such that } a, b \in S \subseteq X_i \quad (12)$$

By consistency and since  $p(\sigma_p(A), A) = 1$ , we also have that

$$p(a, A) = p(a, \sigma_p(A)) p(\sigma_p(A), A) = p(a, \sigma_p(A)) \quad \forall a \in \sigma_p(A)$$

We are ready to conclude our proof, that is, proving (4) where  $\sigma = \sigma_p$ . We have two cases:

1.  $a \notin \sigma_p(A)$ . It trivially follows that  $p(a, A) = 0$ .
2.  $a \in \sigma_p(A)$ . By (8), all the elements in  $\sigma_p(A)$  are equivalent with respect to  $\succsim$  and therefore they are equivalent to some  $a_i$  with  $i \in I$ . It follows that  $\sigma_p(A) \cup \{a_i\} \in \mathcal{A}$  and it is such that  $\sigma_p(A) \cup \{a_i\} \subseteq X_i$ . By (10), we have that

$\sigma_p(\sigma_p(A) \cup \{a_i\}) = \sigma_p(A) \cup \{a_i\}$ . By consistency and (11) and since  $a \in \sigma_p(A)$ , we can conclude that

$$\begin{aligned} p(a, A) &= p(a, \sigma_p(A)) = \frac{p(a, \sigma_p(A) \cup \{a_i\})}{p(\sigma_p(A), \sigma_p(A) \cup \{a_i\})} = \frac{\frac{p(a, \sigma_p(A) \cup \{a_i\})}{p(a_i, \sigma_p(A) \cup \{a_i\})}}{\frac{p(\sigma_p(A), \sigma_p(A) \cup \{a_i\})}{p(a_i, \sigma_p(A) \cup \{a_i\})}} \\ &= \frac{r(a, a_i)}{\sum_{b \in \sigma_p(A)} \frac{p(b, \sigma_p(A) \cup \{a_i\})}{p(a_i, \sigma_p(A) \cup \{a_i\})}} = \frac{r(a, a_i)}{\sum_{b \in \sigma_p(A)} r(b, a_i)} = \frac{\varphi(a)}{\sum_{b \in \sigma_p(A)} \varphi(b)} \end{aligned}$$

as wanted.

Uniqueness of  $\sigma$  trivially follows. ■

## B Appendix: mathematical tools

The class  $\mathcal{B}$  denotes a collection of non-empty subsets of  $X$  that includes all finite sets, that is,  $\mathcal{A} \subseteq \mathcal{B}$ . For example,  $\mathcal{B}$  is in Section 3 the collection of all non-empty finite sets as well as all the budget sets and  $X = \mathbb{R}_+^n$ . We denote by  $A$  and  $B$  generic elements of  $\mathcal{B}$ . Let  $p : \mathcal{B} \rightarrow \Delta(X)$  be a consistent random choice rule, that is,  $p(\cdot, A) \in \Delta(A)$  for all  $A \in \mathcal{B}$  and

$$p(a, B) = p(a, A)p(A, B) \quad \forall a \in A \subseteq B \quad (13)$$

By  $\sigma_p : \mathcal{B} \rightrightarrows X$ , we denote the support correspondence. Note that  $\sigma_p(B) \in \mathcal{A}$  for all  $B \in \mathcal{B}$ . Given  $B \in \mathcal{B}$ , for ease of notation, we might alternatively denote by  $\tilde{B}$  the support of  $p(\cdot, B)$ . By (13),  $p(a, B) = p(a, \tilde{B})$  for all  $a \in \tilde{B}$ . In particular, given  $C \in \mathcal{B}$  such that  $C \subseteq B$ ,

$$p(C, B) = p(C \cap \tilde{B}, \tilde{B}) \quad (14)$$

**Lemma 9** *Let  $A \subseteq B$  with  $A, B \in \mathcal{B}$ . The following statements are true:*

1.  $A \cap \tilde{B} = \tilde{A} \cap \tilde{B}$  and  $p(A, B) = p(\tilde{A} \cap \tilde{B}, \tilde{B})$ ;
2.  $p(A, B) > 0$  if and only if  $A \cap \tilde{B} = \tilde{A}$ .

**Proof 1.** By definition of  $\tilde{A}$ ,  $\tilde{A} \cap \tilde{B} \subseteq A \cap \tilde{B}$ . Vice versa, consider  $a \in A \cap \tilde{B}$ . By contradiction, assume that  $a \notin \tilde{A} \cap \tilde{B}$ . This implies that  $a \notin \tilde{A}$ , that is,  $p(a, A) = 0$ . By (13), we can conclude that  $p(a, B) = 0$ , that is,  $a \notin \tilde{B}$  a contradiction with  $a \in A \cap \tilde{B}$ . We conclude that  $A \cap \tilde{B} \subseteq \tilde{A} \cap \tilde{B}$ , proving the equality between the two sets. Moreover, by definition and since  $p(a, B) = p(a, \tilde{B})$  for all  $a \in \tilde{B}$ , we have that

$$p(A, B) = \sum_{a \in A \cap \tilde{B}} p(a, B) = \sum_{a \in A \cap \tilde{B}} p(a, \tilde{B}) = \sum_{a \in \tilde{A} \cap \tilde{B}} p(a, \tilde{B}) = p(\tilde{A} \cap \tilde{B}, \tilde{B})$$

2. By (13),  $p(a, B) = p(a, A)p(A, B)$  for all  $a \in A \subseteq B$ . This implies that if  $p(A, B) > 0$  and  $a \in A$ , then  $p(a, B) > 0$  if and only if  $p(a, A) > 0$ . It follows that

$A \cap \tilde{B} = \tilde{A}$ . As to the converse, assume that  $A \cap \tilde{B} = \tilde{A}$ . Since  $A, \tilde{A} \neq \emptyset$ , if  $A \cap \tilde{B} = \tilde{A}$ , then there exists  $a \in A$  which belongs to  $\tilde{B}$ . We can conclude that  $p(a, B) > 0$  and, in particular,  $p(A, B) > 0$ .  $\blacksquare$

Let  $f : X \rightarrow V$  be a function that takes values on a vector space  $V$ . It can be extended to  $\mathcal{B}$  by defining  $\phi : \mathcal{B} \rightarrow V$  as the average  $\phi(B) = \sum_{a \in \tilde{B}} f(a) p(a, B)$  of  $f$  with respect to  $p$ . Since  $p(a, B) = p(a, \tilde{B})$  for all  $a \in \tilde{B}$  and  $\sigma_p(B) = \sigma_p(\tilde{B})$ , we have that  $\phi(B) = \phi(\tilde{B})$ .

**Proposition 10** *If the sets  $\{B_i\}_{i=1}^n \subseteq \mathcal{B}$  are pairwise disjoint and  $B = \bigcup_{i=1}^n B_i \in \mathcal{B}$ , then*

$$\phi(B) = \sum_{i=1}^n p(B_i, B) \phi(B_i)$$

**Proof** By Lemma 9 and since  $B_i \subseteq B$ , it follows that  $B_i \cap \tilde{B} = \tilde{B}_i \cap \tilde{B} \in \mathcal{A}$  for all  $i \in \{1, \dots, n\}$ . This implies that:<sup>11</sup> (a)  $\tilde{B}_i \cap \tilde{B}$  are pairwise disjoint and (b)

$$\left( \bigcup_{i=1}^n \tilde{B}_i \right) \cap \tilde{B} = \bigcup_{i=1}^n (\tilde{B}_i \cap \tilde{B}) = \bigcup_{i=1}^n (B_i \cap \tilde{B}) = \left( \bigcup_{i=1}^n B_i \right) \cap \tilde{B} = B \cap \tilde{B} = \tilde{B}$$

Let  $I = \{i : B_i \cap \tilde{B} \neq \emptyset\}$ . By the previous equality,  $I$  is non-empty. On the one hand, by Lemma 9 and since  $p(B_i, B) > 0$  for all  $i \in I$ , we have that  $B_i \cap \tilde{B} = \tilde{B}_i$  for all  $i \in I$ . On the other hand, if  $i \notin I$  then  $B_i \cap \tilde{B} = \emptyset$ , yielding that  $p(B_i, B) = 0$  for all  $i \notin I$ . Thus, if  $i \notin I$  there is no  $a \in B_i$  such that  $p(a, B) > 0$ . Moreover, by (14) we can conclude that for each  $i \in I$

$$0 < p(B_i, B) = p(B_i \cap \tilde{B}, \tilde{B}) = p(\tilde{B}_i \cap \tilde{B}, \tilde{B})$$

and by consistency and since  $\tilde{B} \supseteq \tilde{B}_i \cap \tilde{B} = \tilde{B}_i = B_i \cap \tilde{B} \neq \emptyset$  and  $\tilde{B}_i \cap \tilde{B} \in \mathcal{A}$  for all  $i \in I$ , we have that

$$p(a, \tilde{B}) = p(a, \tilde{B}_i \cap \tilde{B}) p(\tilde{B}_i \cap \tilde{B}, \tilde{B}) = p(a, \tilde{B}_i) p(\tilde{B}_i \cap \tilde{B}, \tilde{B})$$

---

<sup>11</sup>To ease notation, we write  $\tilde{B}_i$  in place of  $\tilde{B}_i$ .

for all  $a \in \tilde{B}_i \cap \tilde{B} = \tilde{B}_i$  and for all  $i \in I$ . By Lemma 9 and since the elements of  $\{\tilde{B}_i \cap \tilde{B}\}_{i \in I}$  are non-empty, pairwise disjoint, and finite, we have

$$\begin{aligned}
\phi(B) &= \phi(\tilde{B}) = \sum_{a \in \tilde{B}} f(a) p(a, \tilde{B}) = \sum_{a \in \bigcup_{i=1}^n (\tilde{B}_i \cap \tilde{B})} f(a) p(a, \tilde{B}) \\
&= \sum_{a \in \bigcup_{i \in I} (\tilde{B}_i \cap \tilde{B})} f(a) p(a, \tilde{B}) = \sum_{i \in I} \sum_{a \in \tilde{B}_i \cap \tilde{B}} f(a) p(a, \tilde{B}) \\
&= \sum_{i \in I} \sum_{a \in \tilde{B}_i} f(a) p(a, \tilde{B}) = \sum_{i \in I} \sum_{a \in \tilde{B}_i} f(a) p(a, \tilde{B}_i) p(\tilde{B}_i \cap \tilde{B}, \tilde{B}) \\
&= \sum_{i \in I} p(\tilde{B}_i \cap \tilde{B}, \tilde{B}) \sum_{a \in \tilde{B}_i} f(a) p(a, \tilde{B}_i) = \sum_{i \in I} p(\tilde{B}_i \cap \tilde{B}, \tilde{B}) \sum_{a \in \tilde{B}_i} f(a) p(a, B_i) \\
&= \sum_{i \in I} p(B_i \cap \tilde{B}, \tilde{B}) \phi(B_i) = \sum_{i \in I} p(B_i \cap \tilde{B}, B) \phi(B_i) \\
&= \sum_{i=1}^n p(B_i \cap \tilde{B}, B) \phi(B_i) = \sum_{i=1}^n p(B_i, B) \phi(B_i),
\end{aligned}$$

proving the statement. ■

## C Appendix: proofs and related analysis

Before starting the proofs, recall that  $X = \mathbb{R}_{++}^n$  and that  $\mathcal{B}$  is the collection of all non-empty finite sets and all the possible budget sets. To be consistent with the standard microeconomic literature, in this section we often denote alternatives in  $X$  – that is, bundles of goods – by  $x$  and  $y$  instead of  $a$  and  $b$ . We start by providing an ancillary fact.

**Proposition 11** *Let  $p : \mathcal{B} \rightarrow \Delta(X)$  be a consistent random choice rule. If  $\succsim$  is the weak order that rationalizes  $\sigma_p$  restricted to  $\mathcal{A}$ ,<sup>12</sup> then*

$$\sigma_p(B) = \{x \in B : x \succsim y \quad \forall y \in B\} \quad \forall B \in \mathcal{B} \quad (15)$$

**Proof** Consider the random choice rule  $p$  restricted to  $\mathcal{A}$ . By Lemma 8 and since  $p$  is consistent on  $\mathcal{A}$ , we have that  $\sigma_p$  restricted to  $\mathcal{A}$  is optimal and is rationalized by the weak order  $\succsim$ , defined as in (7). Let  $B \in \mathcal{B}$ . By the proof of Theorem 1, if  $B \in \mathcal{A}$  then (15) holds. If  $B \notin \mathcal{A}$ , then  $B = B(q, w)$  for some  $(q, w) \in \mathbb{R}_{++}^n \times \mathbb{R}_{++}$ . As before, define  $\tilde{B} = \sigma_p(B)$ . We next prove (15) by proving both inclusions. By contradiction, assume that  $\sigma_p(B) \not\subseteq \{x \in B : x \succsim y \quad \forall y \in B\}$ . Since  $\sigma_p(B) \subseteq B$ , it follows that

<sup>12</sup>That is,  $x \succsim y$  if and only if  $x \in \sigma_p(\{x, y\})$ .

there exists  $x \in B$  such that  $x \succ y$  for some  $y \in \sigma_p(B)$ . Since  $\tilde{B} = \sigma_p(B) \in \mathcal{A}$  and  $p(y, \tilde{B}) = p(y, B)$  for all  $y \in \tilde{B}$ , we have  $\tilde{B} = \sigma_p(B) = \sigma_p(\tilde{B})$ . By Theorem 1, we can conclude that all the elements in  $\sigma_p(\tilde{B})$  are indifferent according to the weak order  $\succsim$ , thus,  $x \succ y$  for all  $y \in \tilde{B}$ , as well as  $x \notin \tilde{B}$ . By Theorem 1 and its proof, if we define  $A = \tilde{B} \cup \{x\} \subseteq B$ , then  $A \in \mathcal{A}$ ,  $p(x, A) = 1$ , and  $p(A, B) = 1$ . By consistency, this implies that  $p(x, B) = p(x, A)p(A, B) = 1$ , a contradiction with  $x \notin \tilde{B}$ , proving the “ $\subseteq$ ” inclusion and that  $\{x \in B : x \succsim y \ \forall y \in B\}$  is non-empty. Vice versa, assume that  $\bar{x} \in B$  is such that  $\bar{x} \succsim y$  for all  $y \in B$ . Define  $A = \tilde{B} \cup \{\bar{x}\} \in \mathcal{A}$ . By the previous part of the proof, we have that all the elements of  $A$  are indifferent according to  $\succsim$ . By Theorem 1 and consistency and since  $1 \leq p(\tilde{B}, B) \leq p(A, B) \leq 1$ , we have that

$$p(\bar{x}, B) = p(\bar{x}, A)p(A, B) = p(\bar{x}, A) > 0$$

proving that  $\bar{x} \in \sigma_p(B)$  and the opposite inclusion. ■

**Theorem 12** *Theorem 1 holds true whenever we replace  $\mathcal{A}$  with  $\mathcal{B}$  and  $\sigma$  is further assumed to be such that  $\sigma(B) \in \mathcal{A}$  for all  $B \in \mathcal{B}$ .*

**Proof** “Only if”. Since  $p$  is consistent,  $p$  is consistent when restricted to  $\mathcal{A}$ . By Theorem 1, it follows that there exist  $\varphi : X \rightarrow (0, \infty)$  and an optimal choice correspondence  $\sigma : \mathcal{A} \rightrightarrows X$  such that for each  $A \in \mathcal{A}$

$$p(a, A) = \begin{cases} \frac{\varphi(a)}{\sum_{b \in \sigma(A)} \varphi(b)} & \text{if } a \in \sigma(A) \\ 0 & \text{else} \end{cases} \quad \forall a \in A \quad (16)$$

Clearly,  $\sigma(A) = \sigma_p(A)$  for all  $A \in \mathcal{A}$ . Since  $p$  is consistent, we also have that  $p(a, A) = p(a, \tilde{A})$  for all  $a \in \tilde{A}$  and for all  $A \in \mathcal{B}$ . Since  $\tilde{A} \in \mathcal{A}$  and  $\sigma_p(\tilde{A}) = \tilde{A} = \sigma_p(A)$  for all  $A \in \mathcal{B}$ , we can conclude that for each  $A \in \mathcal{B}$  if  $a \in \sigma_p(A) = \sigma_p(\tilde{A}) = \tilde{A}$ , then

$$p(a, A) = p(a, \tilde{A}) = \frac{\varphi(a)}{\sum_{b \in \sigma_p(\tilde{A})} \varphi(b)} = \frac{\varphi(a)}{\sum_{b \in \sigma_p(A)} \varphi(b)}$$

while if  $a \in A \setminus \sigma_p(A)$ , then  $p(a, A) = 0$ . This proves that (16) holds with  $\sigma$  replaced by  $\sigma_p$  and for all  $A \in \mathcal{B}$ . By Lemma 11, we have that  $\sigma_p$  is an optimal choice correspondence such that  $\sigma_p(B) \in \mathcal{A}$  for all  $B \in \mathcal{B}$ , proving the implication. “If”. Assume that  $p$  can be represented as in (16) with  $\varphi : X \rightarrow (0, \infty)$  and  $\sigma : \mathcal{B} \rightrightarrows X$  an optimal choice correspondence such that  $\sigma(B) \in \mathcal{A}$  for all  $B \in \mathcal{B}$ . It is immediate to check that  $\sigma(B) = \sigma_p(B)$  for all  $B \in \mathcal{B}$ . By Theorem 1 and since  $\sigma$  is an optimal choice correspondence,  $p$  satisfies consistency on  $\mathcal{A}$ . Since  $\sigma = \sigma_p$  is optimal and  $\sigma_p(A) \subseteq A$  as well as  $\sigma_p(A) \cap \sigma_p(A) \neq \emptyset$  for all  $A \in \mathcal{B}$ , this implies that  $\sigma_p(\sigma_p(A)) = \sigma_p(A) \cap \sigma_p(A) = \sigma_p(A) = \tilde{A}$  and, in particular, this implies that  $p(a, A) = p(a, \tilde{A})$  for all  $a \in \tilde{A}$ . Let now  $A, B \in \mathcal{B}$  be such that  $A \subseteq B$ . Let also  $a \in A$ . We have three cases:

1.  $p(A, B) > 0$  and  $a \in \tilde{A}$ . This implies that  $\sigma_p(B) \cap A \neq \emptyset$ . Since  $\sigma_p$  is optimal, it follows that  $A \cap \tilde{B} = \sigma_p(B) \cap A = \sigma_p(A) = \tilde{A}$ . Since  $p$  is consistent on  $\mathcal{A}$ , we have that

$$p(a, A)p(A, B) = p(a, \tilde{A})p(A \cap \tilde{B}, \tilde{B}) = p(a, \tilde{A})p(\tilde{A}, \tilde{B}) = p(a, \tilde{B}) = p(a, B)$$

2.  $p(A, B) > 0$  and  $a \notin \tilde{A}$ . By contradiction, assume that  $a \in \sigma_p(B) = \tilde{B}$ . Since  $a \in A \subseteq B$ , we would have that  $A \subseteq B$  and  $\sigma_p(B) \cap A \neq \emptyset$ . Since  $\sigma_p$  is optimal, it would follow that  $a \in \sigma_p(B) \cap A = \sigma_p(A) = \tilde{A}$ , a contradiction. Thus,  $a \notin \tilde{B}$  and

$$p(a, B) = 0 = p(a, A)p(A, B)$$

3.  $p(A, B) = 0$ . It follows that  $a \notin \tilde{B}$ , yielding that

$$p(a, B) = 0 = p(a, A)p(A, B)$$

Points 1, 2, and 3 prove the implication. ■

**Proof of Proposition 2** In Proposition 10, define  $f : X \rightarrow \mathbb{R}$  by  $f(x) = q \cdot x$  for all  $x \in X$ . Given a stochastic demand  $d$ , let  $p$  be such that  $d(q, w)(\cdot) = p(\cdot, B(q, w))$ . Set also  $A = B(q, w)$  and  $B = B(q, w')$ . Clearly, we have that  $A \subseteq B$  and  $A, B \in \mathcal{B}$ . By Lemma 9, it follows  $A \cap \tilde{B} = \tilde{A} \cap \tilde{B}$ . Note that  $B_1 = \tilde{A} \cap \tilde{B}$  and  $B_2 = \tilde{B} \setminus \tilde{A}$  belong to  $\mathcal{A} \subseteq \mathcal{B}$  (provided they are non-empty), are pairwise disjoint, and  $\tilde{B} = B_1 \cup B_2 \in \mathcal{A}$ . Moreover,  $B_2 \subseteq A^c$ . Otherwise, there would exist  $x \in B_2 \cap A$ . Thus, we would have that  $x \in B_2$  and  $x \in A$ . By consistency, we could conclude that  $x \in \tilde{B}$ ,  $x \notin \tilde{A}$ , and

$$0 < p(x, B) = p(x, A)p(A, B) = 0$$

a contradiction. Observe also that

$$c(q, w') = \sum_{x \in B(q, w')} (q \cdot x) d(q, w')(x) = \sum_{x \in B(q, w')} (q \cdot x) p(x, B(q, w')) \quad (17)$$

$$= \sum_{x \in \widetilde{B(q, w')}} (q \cdot x) p(x, B(q, w')) = \phi(B(q, w')) = \phi(B) \quad (18)$$

Similarly, we have that  $c(q, w) = \phi(A) = \phi(\tilde{A})$ . Finally, since  $B_2 \subseteq A^c$ , if  $B_2 \neq \emptyset$ , then we also have that  $\tilde{B}_2 \subseteq B_2 \subseteq A^c$

$$\begin{aligned} \phi(B_2) &= \sum_{x \in \tilde{B}_2} (q \cdot x) p(x, B_2) = \sum_{x \in \tilde{B}_2 \cap A^c} (q \cdot x) p(x, B_2) \\ &> w \sum_{x \in \tilde{B}_2 \cap A^c} p(x, B_2) = w \sum_{x \in \tilde{B}_2} p(x, B_2) = w \end{aligned}$$

By Proposition 10 and since  $B_1$  and  $B_2$  are disjoint, if  $B_1, B_2 \neq \emptyset$  then we conclude that

$$c(q, w') = \phi(B) = \phi(\tilde{B}) = \phi(B_1 \cup B_2) = p(B_1, \tilde{B})\phi(B_1) + p(B_2, \tilde{B})\phi(B_2) \quad (19)$$

We have two cases:

1.  $p(A, B) > 0$ . By Lemma 9, it follows that  $B_1 = \tilde{A} \cap \tilde{B} = \tilde{A} \neq \emptyset$ . On the one hand, by (19), if  $B_2 \neq \emptyset$ , then we have

$$\begin{aligned} c(q, w') &= \phi(B) = \phi(\tilde{B}) = \phi(B_1 \cup B_2) = p(B_1, \tilde{B})\phi(B_1) + p(B_2, \tilde{B})\phi(B_2) \\ &= p(B_1, \tilde{B})\phi(\tilde{A}) + p(B_2, \tilde{B})\phi(B_2) \geq p(B_1, \tilde{B})c(q, w) + p(B_2, \tilde{B})w \\ &\geq p(B_1, \tilde{B})c(q, w) + p(B_2, \tilde{B})c(q, w) = c(q, w) \end{aligned}$$

On the other hand, if  $B_2 = \emptyset$ , then  $\tilde{A} = \tilde{A} \cap \tilde{B} = B_1 = B_1 \cup B_2 = \tilde{B}$ . This implies that

$$c(q, w') = \phi(B) = \phi(\tilde{B}) = \phi(\tilde{A}) = \phi(A) = c(q, w)$$

2.  $p(A, B) = 0$ . By (14) and Lemma 9, it follows that

$$p(B_1, \tilde{B}) = p(\tilde{A} \cap \tilde{B}, \tilde{B}) = p(A \cap \tilde{B}, \tilde{B}) = p(A, B) = 0$$

that is,  $B_1 = \emptyset$  and  $\emptyset \neq \tilde{B} = B_2 \subseteq A^c$ , which immediately yields that

$$c(q, w') = \phi(B) = \phi(\tilde{B}) = \phi(B_2) > w \geq c(q, w)$$

Points 1 and 2 prove the statement. ■

**Proof of Lemma 3** Assume that  $p$  is stochastically monotone. We want to show that  $\sigma_p(B) \subseteq \partial^+ B$ . By contradiction, assume that there exists  $x \in \sigma_p(B)$  that does not belong to  $\partial^+ B$ . Then, there exists  $z \in B$  such that  $x \ll z$ . By stochastic monotonicity, we have that  $p(x, \{x, z\}) = 0$ , and so

$$0 < p(x, B) = p(x, \{x, z\})p(\{x, z\}, B) = 0$$

which is a contradiction. ■

**Proof of Proposition 4** Since  $d$  is stochastically monotone,  $p$  is stochastically monotone. Consider  $\partial^+ B(q, w)$ . Note that it is equal to  $\{x \in B(q, w) : q \cdot x = w\}$ . By Lemma 3,  $p(\partial^+ B(q, w), B(q, w)) = 1$ . So, the support of  $d(q, w)$  is contained in  $\partial^+ B(q, w)$ . ■

**Proof of Lemma 5** Clearly, since  $q < q'$ ,  $B(q', w) \subseteq B(q, w)$ . Set  $A = B(q', w)$  and  $B = B(q, w)$ . In Proposition 10, let  $f(x) = x$ . Given a stochastic demand  $d$ , let  $p$  be the consistent random choice rule such that  $d(q, w)(\cdot) = p(\cdot, B(q, w))$ . Clearly,



we have that  $A \subseteq B$  and  $A, B \in \mathcal{B}$ . By Lemma 9, it follows  $A \cap \tilde{B} = \tilde{A} \cap \tilde{B}$ . Note that  $B_1 = \tilde{A} \cap \tilde{B}$  and  $B_2 = \tilde{B} \setminus \tilde{A}$  belong to  $\mathcal{A} \subseteq \mathcal{B}$  (provided they are not empty), are pairwise disjoint, and  $\tilde{B} = B_1 \cup B_2$ . Moreover,  $B_2 \subseteq A^c$ . Otherwise, there would exist  $x \in B_2 \cap A$ . Thus, we would have that  $x \in B_2$  and  $x \in A$ . By consistency, we could conclude that  $x \in \tilde{B}$ ,  $x \notin \tilde{A}$ , and  $x \in A \subseteq B$

$$0 < p(x, B) = p(x, A) p(A, B) = 0$$

a contradiction. Observe that

$$\begin{aligned} \bar{d}(q, w) &= \sum_{x \in B(q, w)} xd(q, w)(x) = \sum_{x \in B(q, w)} xp(x, B(q, w)) \\ &= \sum_{x \in \widetilde{B(q, w)}} xp(x, B(q, w)) = \phi(B(q, w)) = \phi(B) \end{aligned}$$

Similarly, we have that  $\bar{d}(q', w) = \phi(A) = \phi(\tilde{A})$ . Finally, since  $B_2 \subseteq A^c$ , if  $B_2 \neq \emptyset$ , then we also have that  $\tilde{B}_2 \subseteq B_2 \subseteq A^c$

$$\phi(B_2) = \sum_{x \in \tilde{B}_2} xp(x, B_2) = \sum_{x \in \tilde{B}_2 \cap A^c} xp(x, B_2) \text{ and } q' \cdot \phi(B_2) = \sum_{x \in \tilde{B}_2 \cap A^c} (q' \cdot x) p(x, B_2) > w$$

By Proposition 10 and since  $B_1$  and  $B_2$  are disjoint, if  $B_1, B_2 \neq \emptyset$ , then we have that

$$\bar{d}(q, w) = \phi(B) = \phi(\tilde{B}) = \phi(B_1 \cup B_2) = p(B_1, \tilde{B})\phi(B_1) + p(B_2, \tilde{B})\phi(B_2) \quad (20)$$

By (14) and Lemma 9 and since  $A \subseteq B$ , recall that

$$p(A, B) = p(A \cap \tilde{B}, \tilde{B}) = p(\tilde{A} \cap \tilde{B}, \tilde{B}) = p(B_1, \tilde{B}) \quad (21)$$

By contradiction, assume that  $\bar{d}(q', w) > \bar{d}(q, w)$ . We have three cases:

1.  $p(B_2, \tilde{B}) = 0$ . This implies that  $p(B_1, \tilde{B}) = 1$ . By Lemma 9 and (21), it follows that  $B_1 = \tilde{A} \cap \tilde{B} = A \cap \tilde{B} = \tilde{A}$  as well as  $B_2 = \emptyset$ , that is,  $\tilde{B} = B_1 = \tilde{A}$ . We can conclude that

$$\bar{d}(q, w) = \phi(B) = \phi(\tilde{B}) = \phi(\tilde{A}) = \bar{d}(q', w)$$

a contradiction.

2.  $1 > p(B_2, \tilde{B}) > 0$ . This implies that  $1 > p(B_1, \tilde{B}) > 0$ . In particular, we have that  $B_1, B_2 \neq \emptyset$ . By Lemma 9 and (21), it follows that  $B_1 = \tilde{A} \cap \tilde{B} = \tilde{A}$ . By (20), we have that

$$\begin{aligned} \bar{d}(q, w) &= p(B_1, \tilde{B})\phi(B_1) + p(B_2, \tilde{B})\phi(B_2) \\ &= p(B_1, \tilde{B})\phi(\tilde{A}) + p(B_2, \tilde{B})\phi(B_2) \\ &= p(B_1, \tilde{B})\bar{d}(q', w) + p(B_2, \tilde{B})\phi(B_2) \end{aligned}$$

This yields that

$$0 > \bar{d}(q, w) - \bar{d}(q', w) = p(B_2, \tilde{B}) \left[ \sum_{x \in \tilde{B}_2} xp(x, B_2) - \bar{d}(q', w) \right]$$

that is,

$$0 > \sum_{x \in \tilde{B}_2} xp(x, B_2) - \bar{d}(q', w) \implies \bar{d}(q', w) > \sum_{x \in \tilde{B}_2} xp(x, B_2)$$

In turn, since  $\tilde{B}_2 \subseteq B_2 \subseteq A^c$ , this yields that

$$w \geq q' \cdot \bar{d}(q', w) \geq q' \cdot \left( \sum_{x \in \tilde{B}_2} xp(x, B_2) \right) = \sum_{x \in \tilde{B}_2 \cap A^c} (q' \cdot x) p(x, B_2) > w$$

a contradiction.

3.  $p(B_2, \tilde{B}) = 1$ . This implies that  $p(B_1, \tilde{B}) = 0$ . In particular, we have that  $B_1 = \emptyset$  and  $\emptyset \neq \tilde{B} = B_2$ . This implies that

$$\bar{d}(q, w) = \phi(B) = \phi(\tilde{B}) = \phi(B_2)$$

yielding that

$$w \geq q' \cdot \bar{d}(q', w) \geq q' \cdot \bar{d}(q, w) = q' \cdot \phi(B_2) > w$$

a contradiction.

Points 1, 2, and 3 prove the statement. ■

**Proof of Lemma 6** We first prove an ancillary claim:

*Claim* Let  $\bar{B} = B(\bar{q}, \bar{w})$  and  $\hat{B} = B(\hat{q}, \hat{w})$ . If  $p(\bar{B} \cap \hat{B}, \hat{B}) = 0$ , then

$$\bar{q} \cdot \bar{d}(\hat{q}, \hat{w}) > \bar{w}$$

*Proof of the Claim* By assumption, the support of  $p(\cdot, \hat{B})$  is contained in  $\hat{B} \cap \bar{B}^c$ , in particular,  $\tilde{\hat{B}} \subseteq \bar{B}^c$ . It follows that

$$\bar{q} \cdot \bar{d}(\hat{q}, \hat{w}) = \sum_{x \in \tilde{\hat{B}}} (\bar{q} \cdot x) p(x, \hat{B}) > \sum_{x \in \tilde{\hat{B}}} \bar{w} p(x, \hat{B}) = \bar{w}$$

proving the claim. □

Consider  $(q, w)$  and  $(q', w')$  in  $\mathbb{R}_{++}^n \times \mathbb{R}_{++}$ . Define  $B = B(q, w)$  and  $B' = B(q', w')$ . By the previous claim and setting  $\bar{B} = B'$  and  $\hat{B} = B$ , it follows that  $p(B' \cap B, B) > 0$ . Otherwise, we would have that  $w' < q' \cdot \bar{d}(q, w) = c(q', w') \leq w'$ , a contradiction. Since  $p(B' \cap B, B) > 0$ , denote by  $\bar{x} \in \sigma_p(B) \cap B'$ . By Proposition 11 and since  $\bar{x} \in \sigma_p(B)$ , we have that  $B \ni \bar{x} \succsim y$  for all  $y \in B$ . We have two cases:

1.  $p(B' \cap B, B') = 0$ . By the previous claim and setting  $\bar{B} = B$  and  $\hat{B} = B'$ , it follows that  $q \cdot \bar{d}(q', w') > w$ .
2.  $p(B' \cap B, B') > 0$ . This implies that

$$q \cdot \bar{d}(q', w') = \sum_{x \in \widetilde{B}'} (q \cdot x) p(x, B') = \sum_{x \in \widetilde{B}' \cap B} (q \cdot x) p(x, B') + \sum_{x \in \widetilde{B}' \cap B^c} (q \cdot x) p(x, B') \quad (22)$$

Given  $x \in \widetilde{B}'$ , we have two subcases:

- (a)  $x \in \widetilde{B}' \cap B^c$ . In this case,  $q \cdot x > w$ .
- (b)  $x \in \widetilde{B}' \cap B = \sigma_p(B') \cap B$ . By Proposition 11, it follows that  $x \succsim y$  for all  $y \in B'$ . In particular, since  $\bar{x} \in B'$ , this implies that  $x \succsim \bar{x}$ . At the same time, since  $\bar{x} \in B$  is such that  $\bar{x} \succsim y$  for all  $y \in B$ , we have that  $x \succsim y$  for all  $y \in B$ . By Proposition 11 and since  $x \in B$ , this yields that  $x \in \sigma_p(B)$ . By stochastic monotonicity, we can conclude that  $q \cdot x = w$ .

To sum up, by (22) and points a and b, we can conclude that

$$q \cdot \bar{d}(q', w') \geq \sum_{x \in \widetilde{B}' \cap B} wp(x, B') + \sum_{x \in \widetilde{B}' \cap B^c} wp(x, B') = w \quad (23)$$

By (23) as well as points 1 and 2, we have that

$$\begin{aligned} (q' - q) \cdot (\bar{d}(q', w') - \bar{d}(q, w)) &= q' \cdot (\bar{d}(q', w') - \bar{d}(q, w)) - q \cdot (\bar{d}(q', w') - \bar{d}(q, w)) \\ &= q' \cdot \bar{d}(q', w') - q' \cdot \bar{d}(q, w) - q \cdot (\bar{d}(q', w') - \bar{d}(q, w)) \\ &= w' - w' + w - q \cdot \bar{d}(q', w') = w - q \cdot \bar{d}(q', w') \end{aligned}$$

proving the main statement. ■

**Proof of the Law of Demand** Consider an initial price and wealth pair  $(q, w)$ . Let  $q' \in \mathbb{R}_{++}^n$  be such that  $q'_k > q_k$  and  $q'_i = q_i$  for all  $i \neq k$ . Let  $w' = w'(q')$  be such that  $w' = q' \cdot \bar{d}(q, w) \geq q \cdot \bar{d}(q, w) = w$ , since  $p$  is stochastically monotone. By Lemma 4 and since  $p$  is stochastically monotone, it follows that  $q' \cdot \bar{d}(q, w) = w' = c(q', w')$ . In view of Lemma 6, the difference  $\bar{d}(q', w') - \bar{d}(q, w)$  quantifies a substitution effect on the goods' average demand due only to the price change  $q' - q$ . This suggests the following decomposition:

$$\bar{d}(q', w) - \bar{d}(q, w) = \underbrace{\bar{d}(q', w) - \bar{d}(q', w')}_{\text{wealth effect}} + \underbrace{\bar{d}(q', w') - \bar{d}(q, w)}_{\text{substitution effect}} \quad (24)$$

in which the r.h.s. accounts for, respectively, the wealth and substitution effects on the goods' demand. Note that the elements in (24) are vectors. Thus, the equality holds componentwise. Since good  $k$  is normal and  $w \leq w'$ , we have

$$\bar{d}_k(q', w) \leq \bar{d}_k(q', w') \quad (25)$$

By Lemma 6 and the choice of  $q$  and  $q'$  and since  $q' \cdot \bar{d}(q, w) = w' = c(q', w')$ , we have that

$$(q'_k - q_k) (\bar{d}_k(q', w') - \bar{d}_k(q, w)) = (q' - q) \cdot (\bar{d}(q', w') - \bar{d}(q, w)) \leq 0$$

Since  $q'_k - q_k > 0$ , it follows that  $\bar{d}_k(q', w') - \bar{d}_k(q, w) \leq 0$ . By (24) and (25), this implies that

$$\bar{d}_k(q', w) - \bar{d}_k(q, w) = [\bar{d}_k(q', w) - \bar{d}_k(q', w')] + [\bar{d}_k(q', w') - \bar{d}_k(q, w)] \leq 0$$

proving the statement. For,  $\bar{d}_k(q', w) \leq \bar{d}_k(q, w)$  where in  $q'$  only the price of  $k$  increased, while the other prices did not change and wealth remained constant. A similar argument holds if we consider a decrease in price, that is  $q'_k < q_k$ , rather than an increase. In that case, we would have that  $\bar{d}_k(q', w) \geq \bar{d}_k(q, w)$ . ■

## References

- [1] K. J. Arrow, Rational choice functions and orderings, *Economica*, 26, 121-127, 1959.
- [2] T. Bandyopadhyay, I. Dasgupta, P. K. Pattanaik, Stochastic revealed preference and the theory of demand, *Journal of Economic Theory*, 84, 95-110, 1999.
- [3] G. S. Becker, Irrational behavior and economic theory, *Journal of Political Economy*, 70, 1-13, 1962.
- [4] S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci, and A. Rustichini, Multinomial logit processes and preference discovery: inside and outside the black box, IGER WP 615, 2017.
- [5] F. Echenique and K. Saito, General Luce model, mimeo, 2015.
- [6] X. Gabaix, A sparsity-based model of bounded rationality, *Quarterly Journal of Economics*, 129, 1661-1710, 2014.
- [7] C. Halldin, The choice axiom, revealed preference, and the theory of demand, *Theory and Decision*, 5, 139-160, 1974.
- [8] M. Jerison and J. K.-H. Quah, Law of demand, *The New Palgrave Dictionary of Economics*, 2<sup>nd</sup> Ed., 2006.
- [9] I. Krajbich and M. Dean, How can neuroscience inform economics?, *Current Opinion in Behavioral Sciences*, 5, 51-57, 2015

- [10] R. D. Luce, *Individual choice behavior: A theoretical analysis*, Wiley, New York, 1959.
- [11] A. Mas-Colell, M. D. Whinston, and J. R. Green, *Microeconomic theory*, Oxford University Press, Oxford, 1995.
- [12] F. Matejka and A. McKay, Rational inattention to discrete choices: A new foundation for the multinomial logit model, *American Economic Review*, 105, 272-298, 2015.
- [13] A. Mossin, Elements of a stochastic theory of consumption, *Swedish Journal of Economics*, 70, 200-220, 1968.
- [14] V. Pareto, Sunto di alcuni capitoli di un nuovo trattato di economia pura, *Giornale degli Economisti*, 20, 216-235, 1900 (trans. in *Giornale degli Economisti*, 2008).
- [15] A. Renyi, On a new axiomatic theory of probability, *Acta Mathematica Hungarica*, 6, 285-335, 1955.
- [16] P. A. Samuelson, A note on the pure theory of consumer's behaviour, *Economica*, 5, 61-71, 1938.
- [17] E. E. Slutsky, Sulla teoria del bilancio del consumatore, *Giornale degli Economisti*, 51, 1-26, 1915 (trans. in *Giornale degli Economisti*, 2012).
- [18] M. Woodford, Stochastic choice: An optimizing neuroeconomic model, *American Economic Review P&P*, 104, 495-500, 2014.