# *Random consideration and choice: a case study of "default" options*[*]

Sean Horan[†]

This version: November 2018

**Abstract**

A growing number of stochastic choice models include a "default" option for situations where the decision maker selects none of the feasible alternatives. While this is a welcome development, these models also present an empirical challenge—since the situations where the decision-maker chooses nothing may be difficult to observe. Taking Manzini and Mariotti's (2014) *independent random consideration* model as a point of departure, I investigate what can be learned about models of choice with default when the no-choice behavior of the decision-maker is unobservable.

**JEL Codes:** C44, D01, D03. **Keywords:** random consideration; random utility; default; no-choice.

[*]I gratefully acknowledge the FRQSC for financial support; Victor Aguiar, Paola Manzini, Marco Mariotti and Yusufcan Masatlioglu for their helpful comments; and, Andrew Ellis for an insightful argument that helped lead to Theorem 3.

[†]Contact: Département de sciences économiques, Université de Montréal and CIREQ, C-6018 Pavillon Lionel-Groulx, 3150 rue Jean-Brillant, Montréal QC, Canada H3T 1N8. e-mail: sean.horan@umontreal.ca.

# 1 Introduction

When modelling discrete choice behavior, it is natural to include a "default" option to capture situations where the decision maker selects *none* of the feasible alternatives.[1] While the idea dates back to Corbin and Marley (1974),[2] the practice of including a no-choice option only took off with the recent work of Manzini and Mariotti (2014) and others (Aguiar, 2015 and 2017; Berbeglia, 2018; Brady and Rehbeck, 2016; Demirkany and Kimyaz, 2018; Echenique, Saito and Tserenjigmid, 2018; Zhang, 2016).[3] This is a welcome development. Indeed, the experimental evidence suggests that choice avoidance and deferral are rampant (Dhar, 1997). At the same time, the practice of including a default option presents an empirical challenge since, as Brady and Rehbeck (p. 1204) explain, "it may be difficult to observe a consumer "choosing" nothing" outside of an experimental setting.

The objective of the current paper is to investigate what can be learned about *models of choice with default* when non-choice probabilities are unobservable. My point of departure is Manzini and Mariotti's model of *independent random consideration*, where the decision maker independently notices each feasible alternative with some probability before selecting the most preferred of the alternatives she noticed. Since there is some chance that she notices nothing, this is a model of choice with default. Because choice data seldom contains information about non-choice however, their model is difficult to test directly. For this reason, I consider a *standardized* version of their model that "nets out" the probability of non-choice before renormalizing the choice probabilities of the feasible alternatives. This provides an indirect way to test the model against observable choice frequencies. Manzini and Mariotti briefly discuss this version of their model but leave its axiomatization as an open question.

I establish three results for the standardized independent random consideration model. Theorem 1 axiomatizes the model in terms of key differences from the *multinomial logit* (Luce, 1959; McFadden, 1974), the workhorse of the empirical literature. In turn, Theorem 2 shows that the model parameters can be identified from choice on small menus. Finally, Theorem 3 shows that the model is consistent not only with Block and Marschak's (1960) *random utility model* (RUM) but also Luce and Suppes' *independent RUM* (1965) where the utilities of the alternatives are independently distributed.

Together, these results suggest that the no-choice option does not play a significant role for independent random consideration. For one, the axioms that characterize the standardized model also serve, with minor modification, to characterize the original model (Proposition 1). What is more, the missing no-choice data does not present a problem for identification: the consideration function and preference, even the no-choice probabilities can be determined from observed choice frequencies (Proposition 2). Finally, like the standardized model, the original model is independent RUM (Proposition 3).

[1] For studies of default options in deterministic choice, see Clark (1995) and the recent work of Gerasimou (2018).

[2] See also Chapter 6 of Corbin (1973). However, the basic idea is much older than that. It boils down to *Hobson's choice*, named for a 16th century stable owner in Cambridge. When a customer visited his stable, Hobson would point to the horse in the stall closest to the door and say "chuse whether you will have this or none" (Fisher, 1660).

[3] A related strain of the literature considers discrete choice with a reference point (Kovach, 2018; Suleymanov, 2018).

I conclude by asking whether the lessons learned from the case study of independent random consideration have broader implications. For *all* models of choice with default, I show that the standardized model inherits some features of the original model. Perhaps the most striking connection is that the standardized model is RUM whenever the original model with default is RUM (Theorem 4).

At the same time, I emphasize that independent random consideration is somewhat unusual in terms of its close connection to the standardized model. For other models, the no-choice option plays a more integral role. As a result, the properties of the standardized model (in terms of axiomatics and identification) may be quite different from the original model with default. When this is the case, it is critical to ask whether the targeted applications are ones where non-choice probabilities are observable.

# 2 Preliminaries

## 2.1 Choice functions

For a finite universe of alternatives $X$, let $\mathcal{X} := 2^X \setminus \{\emptyset\}$ denote the collection of (non-empty) menus on $X$. A random **choice function** on $\mathcal{X}$ is a mapping $p : X \times \mathcal{X} \to [0,1]$ that satisfies two requirements for every menu $A \in \mathcal{X}$: (i) $\Sigma_{a \in A} p(a, A) \leq 1$; and (ii) $p(a, A) > 0$ if and only if $a \in A$.

A choice function $p$ is **standard** if $p(\,\cdot\,, A)$ defines a probability ($\Sigma_{a \in A} p(a, A) = 1$) for every $A \in \mathcal{X}$; and **non standard** if $p(\,\cdot\,, A)$ does *not* define a probability ($\Sigma_{a \in A} p(a, A) < 1$) for any $A \in \mathcal{X}$. Intuitively, $p$ is non standard if, for every menu, there is some chance that nothing is chosen.

To formalize, let $A^* := A \cup x^*$ denote the **non standard menu** that adds the (default or) **no-choice option** $x^* \notin X$ to the **standard menu** $A \in \mathcal{X}$. Let $\mathcal{X}^* := 2^{X^*} \setminus \{\emptyset\}$ denote the collection of menus on $X^* := X \cup \{x^*\}$, so that $\mathcal{X}^* \setminus \mathcal{X}$ is the collection of non standard menus.

Given a non standard choice function $p$ on $\mathcal{X}$, the corresponding **choice function with default** $p^* : X^* \times (\mathcal{X}^* \setminus \mathcal{X}) \to [0,1]$ is defined, for all $A^* \in \mathcal{X}^* \setminus \mathcal{X}$ and $x \in A^*$ by

$$p^*(x, A^*) := \begin{cases} p(x, A) & \text{for } x \in A\text{; and,} \\ 1 - \Sigma_{a \in A} p(a, A) & \text{for } x = x^*. \end{cases} \qquad (\mathbf{p}^*)$$

Note that $p^*$ is a *partial* standard choice function on $\mathcal{X}^*$ since $p^*(\,\cdot\,, A)$ is a probability measure for every non standard menu $A \in \mathcal{X}^* \setminus \mathcal{X}$ but is not defined for any standard menu $A \in \mathcal{X}$.

**Remark 1** *While there is a one-to-one correspondence between choice functions with default and non standard choice functions, it is more convenient to study the former. In the sequel, I follow the convention that a standard choice function on $\mathcal{X}$ is represented by an unstarred $p$ (or $q$) while a choice function with default on $\mathcal{X}^*$ is represented by a starred $p^*$ (or $q^*$).*

For any choice function with default $p^*$ on $\mathcal{X}^*$, the choice frequencies of the standard alternatives induce a **standardized choice function** $\widehat{p}^*$ on $\mathcal{X}$. For all menus $A \in \mathcal{X}$ and alternatives $x \in A$:

$$\widehat{p}^*(x, A) := \frac{p^*(x, A^*)}{\Sigma_{a \in A} p^*(a, A^*)} = \frac{p^*(x, A^*)}{1 - p^*(x^*, A^*)} \tag{$\widehat{\mathbf{p}}^*$}$$

In circumstances where no-choice is not directly observable, the standardized choice function $\widehat{p}^*$ reflects the empirical content of the decision maker's actual choice behavior $p^*$.

## 2.2 Random consideration

Ultimately, my goal is to say something about the standardized choice functions associated with various models of choice with default. In Section 3, I start with a special case of the **random consideration model** (RCM) (Cattaneo, Ma, Masatlioglu and Suleymanov, 2018). A random consideration rule $(\mathbf{\Gamma}, \succ)$ consists of: a collection $\mathbf{\Gamma} := \{\Gamma(\,\cdot\,; A)\}_{A \in \mathcal{X}}$ of probabilities $\Gamma(\,\cdot\,; A) : 2^A \to [0, 1]$ that, for each menu $A \in \mathcal{X}$, $\Gamma(\,\cdot\,; A)$, assign a probability to every subset of $A$; and a linear order $\succ$ on $X$.

Given a menu $A \in \mathcal{X}$, the decision maker first focuses on a **consideration set** of alternatives $Y \subseteq A$ with probability $\Gamma(Y; A)$ before choosing the most preferred alternative in $Y$ according to her **preference** $\succ$. For a menu $A \in \mathcal{X}$, let $\mathcal{A}^x_\succ := \{Y \subseteq A : x = \arg\max_\succ Y\}$ denote the collection of consideration sets where $x \in A$ is the most preferred alternative. Every consideration-preference pair $(\mathbf{\Gamma}, \succ)$ defines a choice function $p_{(\mathbf{\Gamma}, \succ)}$ such that, for all menus $A \in \mathcal{X}$,

$$p_{(\mathbf{\Gamma}, \succ)}(x, A) := \Sigma_{Y \in \mathcal{A}^x_\succ} \Gamma(Y; A) \quad \text{for all } x \in A. \tag{RCM}$$

While a variety of different restrictions on $\mathbf{\Gamma}$ are quite natural (see Cattaneo et al.), my main interest is the **independent random consideration model**, where the probability of considering each alternative is determined independently by a **random consideration function** $\gamma : X \to (0, 1)$.

Since $\gamma(x)$ describes the independent probability of "drawing" alternative $x$ when it is feasible, there are two natural ways to model the probability of considering the subset $Y \subseteq A$. One approach, pursued by Manzini and Mariotti,[4] is to draw "without replacement" from the menu $A$ so that

$$\Gamma^*(Y; A) := \prod_{y \in Y} \gamma(y) \cdot \left( \prod_{a \in A \setminus Y} [1 - \gamma(a)] \right).$$

[4]This procedure for selecting a subset of the feasible set originates with Regenwetter et al. (1988). The difference is that, unlike Manzini and Mariotti, Regenwetter et al. do not study the subsequent choice of an alternative from this set.

This approach entails a strictly positive probability that *no* alternative is drawn,[5] namely

$$\Gamma^*(\emptyset; A) := \prod_{a\in A}[1-\gamma(a)].$$

This approach defines a choice function with default $p^*_{(\gamma,\succ)} := p_{(\Gamma^*,\succ)}$ such that, for all $A \in \mathcal{X}$,

$$p^*_{(\gamma,\succ)}(x, A^*) = \Sigma_{Y\in \mathbf{C}^x_{\succ}(A)}\Gamma^*(Y; A) = \begin{cases}\gamma(x) \ \cdot \prod_{a\in A: a\succ x}[1-\gamma(a)] & \text{for all } x \in A\\ \prod_{a\in A}[1-\gamma(a)] & \text{for } x = x^*.\end{cases} \qquad \textbf{(I-RCM}^*\textbf{)}$$

An equally natural approach, which I pursue here, is to draw subsets from $A$ "with replacement" until a non-empty set is drawn. In that case, the probability of considering the subset $Y \subseteq A$ is

$$\Gamma(Y; A) := \frac{\Gamma^*(Y; A)}{1-\Gamma^*(\emptyset; A)}.$$

This approach defines a standard choice function $p_{(\gamma,\succ)} := p_{(\Gamma,\succ)}$ such that, for all $A \in \mathcal{X}$,

$$p_{(\gamma,\succ)}(x, A) = \Sigma_{Y\in \mathbf{C}^x_{\succ}(A)}\Gamma(Y; A) = \frac{\gamma(x)\cdot\prod_{a\in A: a\succ x}[1-\gamma(a)]}{1-\prod_{a\in A}[1-\gamma(a)]} \quad \text{for all } x \in A. \qquad \textbf{(I-RCM)}$$

This representation is the standardized version of Manzini and Mariotti's model since

$$p_{(\gamma,\succ)}(x, A) = \frac{\gamma(x)\cdot\prod_{a\in A: a\succ x}[1-\gamma(a)]}{1-\prod_{a\in A}[1-\gamma(a)]} = \frac{p^*_{(\gamma,\succ)}(x, A^*)}{1-p^*_{(\gamma,\succ)}(x^*, A^*)} = \widehat{p}^*_{(\gamma,\succ)}(x, A).$$

As such, the same consideration-preference pair $(\gamma, \succ)$ defines two *different* representations. However, it will be clear from context whether the intended representation is (I-RCM) and (I-RCM$^*$). In the sequel, a choice function that admits either kind of representation will be called **independent RCM**.

## 2.3 Odds and odds ratios

My axiomatic characterization of the standardized model in Section 3 emphasizes its differences from the classical **Luce model** (1959). Recall that a (standard) Luce rule is a strictly positive utility function $u : X \to \mathbb{R}_{++}$, which defines a standard choice function $p_u$ such that, for all menus $A \in \mathcal{X}$,

$$p_u(x, A) := \frac{u(x)}{\Sigma_{a\in A}u(a)} \quad \text{for all } x \in A. \qquad \textbf{(Luce)}$$

[5]When $p^*_{(\gamma,\succ)}$ is viewed as a choice function with default, it is also consistent with the interpretation that the default alternative $x^*$ is *always* considered ($\Gamma^*(x^*; A^*) = 1$ for all $A^* \in \mathcal{X}^* \setminus \mathcal{X}$) but *never* desirable ($x \succ x^*$ for all $x \in X$).

The defining feature of behavior in this model is that the likelihood of selecting one alternative over another is independent of the other feasible alternatives. For a choice function $p$, let

$$L_{[a:b]}(A) := \frac{p(a, A \cup \{a, b\})}{p(b, A \cup \{a, b\})} \tag{$\mathbf{L_{[a:b]}}$}$$

denote the **odds** that $a$ is chosen over $b$ from $A \cup \{a, b\}$.[6] As Luce showed, a standard choice function $p$ is represented by a Luce rule if and only if for all menus $A, B \in \mathcal{X}$ and alternatives $a, b \notin A \cup B$:

$$L_{[a:b]}(A) = L_{[a:b]}(B).$$

Under the assumption of *full domain* (i.e., that $p(\,\cdot\,, A)$ is observable for all $A \in \mathcal{X}$), this condition can be stated more succinctly in terms of choice probabilities for menus that differ by a single alternative:

**L-Independence** *For all menus $A \in \mathcal{X}$ and alternatives $a, b, c \notin A$:*

$$L_{[a:b]}(A) = L_{[a:b]}(A \cup \{c\}).$$

This is the well-known "independence of irrelevant alternatives" condition from the literature. In economic terms, it states that the substitution pattern between any two alternatives $a$ and $b$ (as captured by their odds) does not change when another alternative $c$ is added to the menu.[7]

Put in another way, the Luce model imposes very stringent restrictions on the **odds ratio**

$$R_{[a:b]}(A; c) := \frac{L_{[a:b]}(A)}{L_{[a:b]}(A \cup \{c\})}, \tag{$\mathbf{R_{[a:b]}}$}$$

which measures the change in odds between $a$ and $b$ when a new alternative $c$ is added to $A \cup \{a, b\}$. Specifically, it requires $R_{[a:b]}(A; c) = 1$ for all menus $A \in \mathcal{X}$ and alternatives $a, b, c \notin A$. The literature suggests a variety of reasons for the odds of choosing $a$ over $b$ to be context-dependent. For instance, the well known *red bus/blue bus* example (Debreu, 1960; McFadden, 1974) suggests that the odds should change when the alternative $c$ added to the menu has more in common with $a$ than $b$.

## 3 Independent random consideration

In this section, I provide an axiomatic characterization of the standardized model of independent random consideration and show that its parameters are identified by choice from small menus. I also show that the model is consistent with a restrictive form of Block and Marschak's random utility model.

[6] Since it will always be clear from the context, I omit any explicit reference to the random choice function $p$.

[7] Equivalently, this axiom can be interpreted in terms of "removing" $c$ from the menu $A \cup \{a, b, c\}$. The same is true for the other axioms below which, for concordance, are all described in terms of "adding" an alternative.

## 3.1 Axioms

My axiomatization emphasizes the differences between the standardized model of independent random consideration and the Luce model. The first axiom mandates a certain amount of context-dependence:

**R-Asymmetry** *For all menus $A \in \mathcal{X}$ and distinct alternatives $a, b, c \notin A$, exactly one of the odds ratios $R_{[a:b]}(A; c)$, $R_{[a:c]}(A; b)$, and $R_{[b:c]}(A; a)$ is different from $1$.*

For any three alternatives, exactly one affects the substitution pattern between the other two. It follows that changes in odds cannot be "too" widespread when an alternative is added to the menu.

In turn, the second axiom requires a weak form of L-Independence where the odds ratios (rather than the odds themselves) must be independent of the irrelevant alternatives in the menu:

**R-Independence** *For all menus $A, B \in \mathcal{X}$ and alternatives $a, b, c \notin A \cup B$:*

$$R_{[a:b]}(A; c) = R_{[a:b]}(B; c).$$

While the odds between two alternatives $a$ and $b$ may change when some other alternative $c$ is added to the menu, the effect itself must be independent of the other feasible alternatives.

**Remark 2** *Marley (1991) proposed a more general condition that adds a parameter $\theta : \mathcal{X} \to \mathbb{R}_+$ to allow for some amount of context-dependence.*[8] *Assuming full domain, his condition is equivalent to the requirement that, for all menus $A, B \in \mathcal{X}$ and alternatives $a, b \notin A \cup B$,*

$$\frac{L_{[a:b]}(A \cup \{c\})^{\theta(A \cup \{c\})}}{L_{[a:b]}(A)^{\theta(A)}} = \frac{L_{[a:b]}(B \cup \{c\})^{\theta(B \cup \{c\})}}{L_{[a:b]}(B)^{\theta(B)}}.$$

*R-Independence is the special case of Marley's condition where $\theta(Y)$ is constant for all $Y \in \mathcal{X}$. In recent work, Ravid (2015) proposed a model that is entirely characterized by R-Independence.*

To provide some intuition for the third axiom, suppose that two different alternatives $c$ and $d$ affect the odds between $a$ and $b$. Then, R-Independence implies that the *relative* change

$$\frac{L_{[a:b]}(A) - L_{[a:b]}(A \cup \{c\})}{L_{[a:b]}(A) - L_{[a:b]}(A \cup \{d\})}$$

in the substitution pattern between $a$ and $b$ cannot depend on the "irrelevant" alternatives in $A$. The third axiom specifies that it must, in fact, depend on the odds of choosing $c$ and $d$ over $b$:

[8] For all $S, T \in \mathcal{X}$ and $x, y \notin S \cup T$, $L_{[x:y]}(S \cup T)^{\theta(S \cup T)} L_{[x:y]}(S \cap T)^{\theta(S \cap T)} = L_{[x:y]}(S)^{\theta(S)} L_{[x:y]}(T)^{\theta(T)}$.

**R-Proportionality** *For all menus $A \in \mathcal{X}$ and alternatives $a, b, c, d \notin A$:*

$$R_{[a:b]}(A; c), R_{[a:b]}(A; d) \neq 1 \implies \frac{L_{[a:b]}(A) - L_{[a:b]}(A \cup \{c\})}{L_{[a:b]}(A) - L_{[a:b]}(A \cup \{d\})} = \frac{L_{[c:b]}(A)}{L_{[d:b]}(A)}.$$

One implication is that two alternatives which affect the odds of $a$ over $b$ must do in the *same* direction: if adding $c$ to the menu decreases (increases) these odds, then adding $d$ must do the same.[9]

Finally, the fourth axiom states that the addition of an alternative to the menu must *decrease* the chance of choosing every other alternative. This is a strong form of the usual regularity requirement, which states that the addition of one alternative cannot *increase* the chance of choosing another:

**Strict Regularity** *For all menus $A \in \mathcal{X}$ and distinct alternatives $a, b \notin A$*

$$p(a, A \cup \{a\}) > p(a, A \cup \{a, b\}).$$

These four axioms characterize the standardized independent random consideration model.

**Theorem 1** *A standard choice function $p$ is independent RCM if and only if it satisfies R-Independence, R-Asymmetry, R-Proportionality, and Strict Regularity.*[10]

I leave the proof to the Mathematical Appendix.

**Remark 3** *Manzini and Mariotti used two axioms, called* I-Asymmetry *and* I-Independence, *to characterize the model with default. Conceptually, these axioms (which are restated in the Online Appendix) are similar to R-Asymmetry and R-Independence, except that they relate to the* ***impact***

$$I_a(A^*; c) := \frac{p^*(a, A^* \cup \{a\})}{p^*(a, A^* \cup \{a, c\})}$$

*of alternative $c \notin A^*$ on the probability of choosing alternative $a$ rather than the odds of choosing $a$ over $b$. Viewed from another perspective, my axioms relate to the* ***relative impact***

$$IR_{[a:b]}(A^*; c) := \frac{I_a(A^* \cup \{b\}; c)}{I_b(A^* \cup \{a\}; c)} = \frac{p^*(a, A^* \cup \{a, b\})}{p^*(a, A^* \cup \{a, b, c\})} \cdot \frac{p^*(b, A^* \cup \{a, b, c\})}{p^*(b, A^* \cup \{a, b\})} = R_{[a:b]}(A^*; c)$$

*of $c$ on the probability of choosing $a$ over $b$ rather than the impact of $c$ on probability of choosing $a$.*

## 3.2 Identification

Identification of the consideration and preference parameters is closely tied to changes in the odds. To see this, suppose that a standard choice function $p$ is represented by the pair $(\gamma, \succ)$; and, that the

[9] Since $L_{[c:b]}(A), L_{[d:b]}(A) > 0$, the right-hand side of the identity in R-Proportionality must be positive.

[10] R-Asymmetry is clearly independent of the other three axioms: every Luce rule *violates* R-Asymmetry but satisfies the other axioms.Table 1 of the Online Appendix illustrates the independence of the other three axioms.

odds ratio satisfies $R_{[a:b]}(A;c) < 1$ for alternatives $a, b, c \notin A$. Then:

(i) $a \succ c \succ b$; and, consequently, (ii) $R_{[a:b]}(A;c) = 1 - \gamma(c)$.

The intuition is that $c$ *only* affects the substitution pattern between $a$ and $b$ when it is between these two alternatives in terms of preference. Since $c$ increases the odds of choosing $a$, paying attention to $c$ must come at the expense of choosing $b$. Observations (i) and (ii) then follow directly.

**Remark 4** *Aguiar (2015 and 2017) partially extends observation (i) to a random consideration model where the consideration probabilities may exhibit some degree of menu-dependence. In his model,* $R_{[a:b]}(\emptyset;c) < 1$ *implies* $c \succ b$ *but* <u>*not*</u> *necessarily* $a \succ c$. *Echenique et al. (2018) propose a different model of random consideration where revealed preference tends to contradict observation (i). In their model,* $R_{[a:b]}(\emptyset;c) < 1$ *generally (but not always) implies* $b \succ c \succ a$.

Observation (i) may be leveraged to identify the relative preference between any two alternatives $x, y \notin A$. In particular, $x \succ y$ if and only if there is some "test" alternative $t \notin A$ such that

$$R_{[t:y]}(A;x) < 1, \tag{$\succ$-1}$$

$$R_{[x:t]}(A;y) < 1, \quad \text{or} \tag{$\succ$-2}$$

$$R_{[x:y]}(A;t) < 1. \tag{$\succ$-3}$$

Each of these conditions reflects a different possibility for the relative preference of the test alternative $t$ (i.e., it is either (1) above, (2) below, or (3) between $x$ and $y$ in terms of preference).

In turn, observation (ii) may be used to identify the consideration function. The idea is to select test alternatives $b, w \notin A$ that nest the alternative $x \notin A$. If $R_{[b,w]}(A;x) < 1$, then

$$\gamma(x) := 1 - R_{[b:w]}(A;x). \tag{$\gamma$-1}$$

This formula identifies $\gamma(x)$ for every alternative $x \in X$ except $\overline{x} := \arg\max_{\succ} X$ and $\underline{x} := \arg\min_{\succ} X$. For those alternatives, it is enough to fix a test alternative $t \neq \overline{x}, \underline{x}$. Then, the identities

$$\gamma(\underline{x}) := \frac{\gamma(t)}{L_{[t:\underline{x}]}(\emptyset)[1-\gamma(t)]} \qquad \text{and} \qquad \gamma(\overline{x}) := \frac{\gamma(t)}{L_{[t:\overline{x}]}(\emptyset)+\gamma(t)} \tag{$\gamma$-2}$$

determine $\gamma(\underline{x})$ and $\gamma(\overline{x})$ (given that $\gamma(t)$ was already determined by the identity ($\gamma$-1) above).

By R-Independence, $A$ may be taken to be the empty set in ($\gamma$-1) above. This shows that identification of the model parameters only requires choice from small menus (though, of course, the same identification strategies *also* work for choice from larger menus).

**Theorem 2** *For a standard choice function p represented by* $(\gamma, \succ)$, *the random consideration function* $\gamma$ *and preference* $\succ$ *are uniquely identified by choice from sets of two and three alternatives.*

## 3.3 Random utility

Independent random consideration is related to the **random utility model** (RUM). Formally, a (finite) random utility rule is a pair $(\mu, v)$ consisting of a probability measure $\mu$ on $\Omega \subseteq \mathbb{N}^{|X|}$ and a strict utility function $v : X \times \Omega \to \mathbb{R}$ that satisfies $v(x_i, \omega) = v(x_j, \omega')$ if and only if $i = j$ and $\omega_i = \omega'_i$. Letting $\Omega_v(x, A) := \{\omega \in \Omega : x \in \arg\max_{a \in A} v(a, \omega)\}$ denote the event that the alternative $x$ is $v$-maximal in $A$, the pair $(\mu, v)$ defines a choice function $p_{(\mu,v)}$ such that, for all menus $A \in \mathcal{X}$,

$$p_{(\mu,v)}(x, A) \; := \; \mu[\Omega_v(x, A)] \quad \text{for all } x \in A. \tag{\textbf{RUM}}$$

For the special **independent RUM**, the random utilities $v(x, \cdot)$ are independently distributed. Letting $\Omega_k$ denote the projection of the state space $\Omega$ onto its $k^{th}$ component, independence requires $\mu(A_i \times A_j) = \mu(A_i) \cdot \mu(A_j)$ for all $i, j \in \{1, ..., |X|\}$ and events $A_i \subseteq \Omega_i$ and $A_j \subseteq \Omega_j$.[11]

**Theorem 3** *If a standard choice function p is independent RCM, then it is independent RUM.*

I leave the proof to the Mathematical Appendix.[12]

**Remark 5** *Almost all random utility models explain violations of L-Independence via correlation among the utilities of different alternatives (see e.g., McFadden's [1978] generalized extreme values and Revelt and Train's [1998] mixed logit). Theorem 3 establishes that correlation is <u>not</u> necessary to accommodate a broader range of substitution patterns among alternatives. In this sense, independent random consideration is conceptually distinct from other random utility models.*

# 4 Implications for the model with default

Theorems 1 to 3 paint a clear picture of the standardized independent random consideration model. In this section, I consider the implications for the original model of choice with default.

I begin with the striking observation that Manzini and Mariotti's original model satisfies three of the four axioms from Theorem 1, namely R-Independence, R-Asymmetry, and R-Proportionality.[13,14] While their model does violate Strict Regularity, it still satisfies a strong form of regularity:

**R-Regularity** *For all menus $A^* \in \mathcal{X}^* \setminus \mathcal{X}$ and distinct alternatives $a, b \notin A$ such that $b \neq x^*$:*

**(i)** $p^*(a, A^* \cup \{a\}) \geq p^*(a, A^* \cup \{a, b\})$; *and,*

**(ii)** $p^*(a, A^* \cup \{a\}) \neq p^*(a, A^* \cup \{a, b\}) \iff R_{[a:x^*]}(A^*; b) = 1$.

[11] Besides the contributions of Tversky (1972) and Suck (2002), surprisingly little is known about independent RUM.
[12] The proof was inspired by a "re-sampling" argument that Andrew Ellis used to show that the model is RUM.
[13] In independent work, Ravid also observed that their model satisfies R-Independence.
[14] In the Mathematical Appendix, I state the implications of these axioms for the no-choice option $x^*$.

Part (i) of this axiom is the usual regularity requirement. In turn, part (ii) provides a direct link between instances of strict regularity and constant odds (relative to the no-choice option). In fact, the *entire* difference between the two models comes down to the difference in their regularity requirements.

**Proposition 1** *A choice function with default $p^*$ is independent RCM if and only if it satisfies R-Independence, R-Asymmetry, R-Proportionality, and R-Regularity.*

While the details are left to the Mathematical Appendix, it is worth noting that the proof of the "if direction" relies on Theorem 1. If $p^*$ satisfies the axioms of Proposition 1, then $\widehat{p}^*$ satisfies the axioms of Theorem 1 which, in turn, implies that $\widehat{p}^*$ is uniquely represented by a pair $(\gamma, \succ)$. What is more, by R-Regularity, the probability of choosing the no-choice option is consistent with independent random consideration (see Lemmas 7-8 of the Mathematical Appendix), i.e.,

$$p^*(x^*, A^*) = \prod_{a \in A} [1 - \gamma(a)] = p^*_{(\gamma,\succ)}(x^*, A^*). \qquad (*)$$

By using $(*)$ to "translate" the representation of the standardized choice function $\widehat{p}^* = p_{(\gamma,\succ)}$ back into the setting of choice with default, one obtains the desired representation $p^* = p^*_{(\gamma,\succ)}$.

Implicit in this line of reasoning is an important practical observation about the independent random consideration model. For any standard choice function $p$ consistent with the model, there is a continuum of choice functions with default $q^*$ such that $p = \widehat{q}^*$. However, only one of these choice functions $q^*$ is consistent with the independent random consideration model.[15]

**Proposition 2** ***(I)*** *A standard choice function $p$ is independent RCM if and only if $p = \widehat{q}^*$ for some choice function with default $q^*$ that is independent RCM.* ***(II)*** *In that case, the same preference-consideration pair $(\gamma, \succ)$ represents the two choice functions $p$ and $q^*$.*

Theorems 1 and 2 capture the behavioral content of the model with default when no-choice data is unobservable. By part (I), it is "as if" the choice frequencies $p$ consistent with the axioms of Theorem 1 are induced by a choice function with default $q^*$ that satisfies the axioms of Proposition 1.[16] By part (II), the parameter identification of $p$ (guaranteed by Theorem 2) is also sufficient to identify the parameters of $q^*$ and even *interpolate* the no-choice probabilities $q^*(x^*, A^*)$.

In combination, Theorem 1 and Proposition 1 show that behavior consistent with the standardized model closely resembles behavior consistent with the original model. Intuitively, they suggest that the no-choice option (which only affects the regularity properties of the model) plays a minor role in terms of behavior. The next result cements this view. Combined with Theorem 3, it shows that the no-choice option has *no* bearing on the fact that the model is independent RUM.

[15] For an illustration, see Table 2 of the Online Appendix.

[16] It is worth noting that every standard choice function $p$ that is independent RCM is also induced by a host of choice functions with default that are *not* independent RCM. See Table 2 of the Online Appendix for an example.

**Proposition 3** *If a choice function with default $p^*$ is independent RCM, then it is independent RUM.*

This strengthens Manzini and Mariotti's observation (pp. 1167-8) that their model is RUM.

# 5 Some general observations

Propositions 1 to 3 establish close connections between the standardized version of Manzini and Mariotti's model and the original model—in terms of axioms, identification and consistency with RUM. In this section, I show that, to some extent, all models of choice with default exhibit such connections.

## 5.1 Axioms

My analysis of independent random consideration suggests two general observations about the axiomatic connection between a model of choice with default and the standardized version of the model.

The first relates to "relative conditions" (like R-Asymmetry, R-Independence and R-Proportionality) that only depend on ratios involving choice probabilities of standard alternatives. In particular, a choice function with default $p^*$ satisfies a relative condition if and only if the standardized choice function $\widehat{p}^*$ does. Of course, some care must be taken in deciding whether an axiom is relative.

**Example 1 (Hazard rate conditions)** *Several recent models of choice with default (see e.g., Brady and Rehbeck, 2016; Demirkany and Kimyaz, 2018; Echenique, Saito and Tserenjigmid, 2018), have been characterized in terms of hazard rate conditions. Given a choice function with default $p^*$, the hazard rate $h_{(p^*,\succ)}$ associated with a linear order $\succ$ on $X$ is defined, for all $A \in \mathcal{X}$, by*

$$h_{(p^*,\succ)}(x, A^*) := \frac{p^*(x, A^*)}{1 - \Sigma_{a \in A: a \succ x} p^*(a, A^*)} \quad \textit{for all } x \in A.$$

*The hazard rate $h_{(p^*,\succ)}$ is* *<u>not</u>* *a ratio of choice probabilities involving standard alternatives since*

$$h_{(p^*,\succ)}(x, A^*) = \frac{p^*(x, A^*)}{1 - \Sigma_{a \in A: a \succ x} p^*(a, A^*)} = \frac{p^*(x, A^*)}{\Sigma_{a \in A: a \preceq x} p^*(a, A^*) + p^*(x^*, A^*)}.$$

This example leads directly to the second observation. For some axioms in the setting of choice with default, it is easy to define a "relativized" version of the axiom for the standard choice setting. The *impact* requirements in I-Asymmetry and I-Independence, for instance, have straightforward implications for *impact ratios* (as discussed in Remark 3 above). For other axioms, like the hazard rate conditions in Example 1, one cannot simply "translate" to the standard setting by taking ratios.

## 5.2 Identification

In general, the setting of choice with default facilitates identification. If a model parameter can be identified from standardized choice data, then it can also be identified when no-choice probabilities are observable. In other words, nothing stands to be gained by standardizing the choice data.

The independent random consideration model is surprising precisely because nothing is lost by standardizing the choice data. At the very least, one might expect standardization to "destroy" information about the default alternative. To illustrate, consider the following generalization of the Luce model discussed by Corbin (1973) and Echenique et al. (2018).

**Example 2 (Non-standard Luce model)** *A non-standard Luce rule is a strictly positive utility function $u : X^* \to \mathbb{R}_{++}$. It defines a choice function with default $p_u^*$ such that, for all $A \in \mathcal{X}$,*

$$p_u^*(x, A^*) := \frac{u(x)}{\Sigma_{a\in A} u(a) + u(x^*)} \quad \text{for all } x \in A. \qquad (\textbf{Luce}^*)$$

*From the observed choice frequencies $\widehat{p}_u^*$, one can recover the utility $u(x)$ of any alternative $x \in X$ (up to a positive multiplicative constant). However, the choice frequencies reveal nothing about the utility of the no-choice option $x^*$. This is evident from the fact that*

$$\widehat{p}_u^*(x, A) = \frac{p_u^*(x, A^*)}{1 - p_u^*(x^*, A^*)} = \frac{u(x)}{\Sigma_{a\in A} u(a)}.$$

The key difference with the independent random consideration model is that choice frequencies are related to the no-choice probabilities (via the consideration function $\gamma$). Since $\gamma$ can be identified from $\widehat{p}^*_{(\gamma,\succ)}$, the formula $(*)$ can then be used to interpolate the no-choice probabilities $p^*_{(\gamma,\succ)}(x^*, \cdot)$.

## 5.3 Random utility

For independent random consideration, standardization preserves consistency with *independent* RUM (see Theorem 3 and Proposition 3). A weaker result extends to all models of choice with default.

**Theorem 4** *If a choice function $p^*$ is RUM, then so is the standardized choice function $\widehat{p}^*$.*

This result implies that, for any default model that is consistent with RUM, the standardized model is also consistent with RUM. From a practical standpoint, this observation is quite useful since it means that one can rely on conventional tools of discrete choice to analyze the consistency of choice frequency data with the standardized model. Consider, for instance, Aguiar's proposed generalization of Manzini and Mariotti's model. Since this *random categorization* model is RUM (see Lemma 1 of Aguiar, 2017), Theorem 4 implies that the standardized version of the model is also RUM.

In the other direction, a standardized choice function may be consistent with RUM even when the original choice function with default is not. Another proposed generalization of Manzini and Mariotti's model, Brady and Rehbeck's *random feasibility* model, serves to illustrate.

**Example 3 (Random feasibility)** *Brady and Rehbeck (p. 1211) showed that the choice function with default $p^*$ defined below is consistent with their random feasibility model.*

| $p^*(j,i)$ | $a$ | $b$ | $c$ | $x^*$ |
|---|---|---|---|---|
| $\{a,x^*\}$ | $1/2$ | – | – | $1/2$ |
| $\{b,x^*\}$ | – | $3/4$ | – | $1/4$ |
| $\{c,x^*\}$ | – | – | $1/2$ | $1/2$ |
| $\{a,b,x^*\}$ | $1/3$ | $1/2$ | – | $1/6$ |
| $\{a,c,x^*\}$ | $1/2$ | – | $1/4$ | $1/4$ |
| $\{b,c,x^*\}$ | – | $11/13$ | $1/13$ | $1/13$ |
| $\{a,b,c,x^*\}$ | $7/20$ | $11/20$ | $1/20$ | $1/20$ |

$\xrightarrow{\textit{standardization}}$

| $\widehat{p}^*(j,i)$ | $a$ | $b$ | $c$ |
|---|---|---|---|
| $\{a\}$ | 1 | – | – |
| $\{b\}$ | – | 1 | – |
| $\{c\}$ | – | – | 1 |
| $\{a,b\}$ | $2/5$ | $3/5$ | – |
| $\{a,c\}$ | $2/3$ | – | $1/3$ |
| $\{b,c\}$ | – | $11/12$ | $1/12$ |
| $\{a,b,c\}$ | $7/19$ | $11/19$ | $1/19$ |

*As Brady and Rehbeck pointed out, $p^*$ violates Regularity.*[17] *So, it is inconsistent with RUM and, thus, independent random consideration. While the standardized version of $p^*$ is also inconsistent with independent random consideration,*[18] *it is consistent with RUM. In particular, $\widehat{p}^*$ has a random utility representation $(\mu, v)$ with probability measure $\mu = (77/228, 7/228, 17/57, 16/57, 3/95, 2/95)$ and utility function $v$ that is consistent with the vector of orderings $(abc, acb, bac, bca, cab, cba)$.*

# 6 References


**Aguiar, V.** 2015. "Stochastic Choice and Attention Capacities: Inferring Preferences from Psychological Biases."

**Aguiar, V.** 2017. "Random Categorization and Bounded Rationality." *Economics Letters* 159 46–52.

**Berbeglia, G.** 2018. "The Generalized Stochastic Preference Choice Model."

**Brady, R. and J. Rehbeck.** 2016. "Menu-Dependent Stochastic Feasibility." *ECMA* 84 1203–1223.

**Block, H., and J. Marschak.** 1960. "Random Orderings and Stochastic Theories of Responses" in *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*.

**Cattaneo, M.D., X. Ma, Y. Masatlioglu and E. Suleymanov.** 2018. "A Random Attention Model."

**Clark, S.A.** 1995. "Indecisive Choice Theory." *MSS* 30 155–170.

**Corbin, R.** 1973. *New Directions in Theories of Preferential Choice.* MSc. Thesis, McGill University.

**Corbin, R., and A.A.J. Marley.** 1974. "Random Utility Models with Equality: An Apparent, but not Actual, Generalization of Random Utility Models." *J Math Psych* 11 274–293.


[17] In particular, $p^*(a,\{a,b,c,x^*\}) = 7/20 > 1/3 = p^*(a,\{a,b,x^*\})$.

[18] Since $R_{[a:b]}(\emptyset, c) = 22/21$ and $R_{[a:c]}(\emptyset, b) = 2/7$, $p^*$ and $\widehat{p}^*$ both violate R-Asymmetry.

**Dhar, R.** 1997. "Consumer Preference for a No-Choice Option." *J Cons Res* 24(2) 215–231.

**Debreu, G.** 1960. "Review of *Individual Choice Behavior: A Theoretical Analysis.*" *AER* 50 186–188.

**Demirkany, Y. and M. Kimyaz.** 2018. "Hazard Rate, Stochastic Choice and Consideration Sets."

**Echenique, F., K. Saito, and G. Tserenjigmid.** 2018. "The Perception-Adjusted Luce Model." *MSS* 93 67–76.

**Fishburn, P.C.** 1998. "Stochastic Utility," in *Handbook of Utility Theory: Principles (Vol. 1)*, Dordrecht: Kluwer.

**Fisher, S.** 1660. "The Rustick's Alarm to the Rabbies."

**Gerasimou, G.** 2018. "Indecisiveness, Undesirability and Overload Revealed Through Rational Choice Deferral." *Economic Journal* 128(614) 2450–2479.

**Kovach, M.** 2018. "Thinking Inside the Box: Status Quo Bias and Stochastic Consideration."

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

**Luce, R.D. and P. Suppes.** 1965. "Preference, Utility, and Subjective Probability" in *Handbook of Mathematical Psychology (Vol. III)*.

**Manzini, P. and M. Mariotti.** 2014. "Stochastic Choice and Consideration Sets." *ECMA* 82 1153–1176.

**Marley, A.A.J.** 1991. "Context Dependent Probabilistic Choice Models Based on Measures of Binary Advantage." *MSS* 21 201–231.

**Masatlioglu, Y., D. Nakajima, and E. Ozbay.** 2012. "Revealed Attention." *AER* 102 2183–2205.

**McFadden, D.** 1974. "Conditional Logit Analysis of Qualitative Choice Behavior" in *Frontiers in Econometrics*.

**McFadden, D.** 1978. "Modelling the Choice of Residential Location" in *Spatial Interaction Theory and Planning Models*.

**Ravid D.** 2015. "Focus then Compare." University of Chicago working paper.

**Regenwetter M., A.A.J. Marley, and H. Joe.** 1998. "Random Utility Threshold Models of Subset Choice." *Aus J Psych* 50 175–185.

**Revelt, D. and K. Train.** 1998. "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level." *RESTAT* 80 647–657.

**Suck, R.** 2002. "Independent Random Utility Representations." *MSS* 43 371–389.

**Suleymanov, E.** 2018. "Stochastic Attention and Search."

**Tversky, A.** 1972a. "Choice by Elimination." *J Math Psych* 9 341–367.

**Zhang, J.** 2016. "Stochastic Choice with Categorization."

# 7 Mathematical Appendix

**Note 1** *In all of the proofs below, I use the following abbreviated notation:* $L_{ab}(A) := L_{[a:b]}(A \setminus \{a, b\})$.

## 7.1 Proof of Theorems 1-2

**Lemma 1** *Given a function* $\gamma : X \to \mathbb{R}$, *the following identity holds for all* $A := \{a_1, ..., a_n\} \in \mathcal{X}$*:*

$$\Sigma_{i=1}^{n}\gamma(a_i)\prod_{j=1}^{i-1}[1-\gamma(a_j)] = 1 - \prod_{i=1}^{n}[1-\gamma(a_i)] \tag{1}$$

*where* $\prod_{j=i}^{i-1}[1-\gamma(a_j)] := 1$ *(so that* $\Sigma_{i=1}^{1}\gamma(a_i)\prod_{j=1}^{0}[1-\gamma(a_j)] = \gamma(a_1) = 1 - \prod_{i=1}^{1}[1-\gamma(a_i)]$ *by definition).*

**Proof.** The proof is by induction on $n$.

**Base case** $n = 2$**:** Direct computation gives $\gamma(a_1) + \gamma(a_2) \cdot [1 - \gamma(a_1)] = 1 - [1 - \gamma(a_1)] \cdot [1 - \gamma(a_2)]$.
**Induction Step:** Suppose that the formula (1) holds for $n = N$ and consider the case $n = N + 1$. Observe that

$$\Sigma_{i=1}^{N+1}\gamma(x_i)\prod_{j=1}^{i-1}[1-\gamma(a_j)] = \gamma(a_1) + [1-\gamma(a_1)]\left(\Sigma_{i=2}^{N+1}\gamma(a_i)\prod_{j=2}^{i-1}[1-\gamma(a_j)]\right). \tag{2}$$

Combining (2) with the induction hypothesis gives

$$\Sigma_{i=1}^{N+1}\gamma(a_i)\prod_{j=1}^{i-1}[1-\gamma(a_j)] = \gamma(a_1) + [1-\gamma(a_1)]\left(1 - \prod_{i=2}^{N+1}[1-\gamma(a_i)]\right) = 1 - \prod_{i=1}^{N+1}[1-\gamma(a_i)]. \tag{3}$$

This completes the induction step. ■

**Note 2** *For the remainder of section, let* $\succ$ *denote the binary relation defined by (*$\succ$*-1)-(*$\succ$*-3) from Section 3.*

**Lemma 2** *If* $p$ *satisfies the axioms of Theorem 1, then* $\succ$ *is a linear order.*

**Proof.** To establish the result, I show that $\succ$ is (i) total, (ii) asymmetric, and (iii) transitive.

**(i)** Fix distinct $x, y, t \in X$. By R-Asymmetry, exactly one of the following holds: $L_{ty}(X) \neq L_{ty}(X \setminus \{x\})$; $L_{ty}(X) \neq L_{ty}(X \setminus \{x\})$; or, $L_{xy}(X) \neq L_{xy}(X \setminus \{t\})$. In every case, either $x \succ y$ or $y \succ x$ by definition of $\succ$. So, $\succ$ is total.

**(ii)** Towards a contradiction, suppose that $x \succ y$ and $y \succ x$. Denote by $\succ_1$, $\succ_2$, and $\succ_3$ the three revealed preferences defined by ($\succ$-1)-($\succ$-3) in Section 3. By definition, there are test alternatives $t, t' \in X$ such that $t$ reveals $x \succ_i y$ and $t'$ reveals $y \succ_j x$ for some $i, j \in \{1, 2, 3\}$. By R-Asymmetry and R-Independence, $t$ and $t'$ must be distinct. By R-Independence, it is without loss of generality that $x \succ y$ and $y \succ x$ are revealed on $A := \{x, y, t, t'\}$.

In the following tables, let $(i, j)$ denote the case where (a) $x \succ_i y$ and (b) $y \succ_j x$.

| $(i,j)$ | 1 | 2 | 3 |
|---|---|---|---|
| 1 | (a) $L_{ty}(A) > L_{ty}(A \setminus \{x\})$<br>(b) $L_{t'x}(A) > L_{t'x}(A \setminus \{y\})$ | (a) $L_{ty}(A) > L_{ty}(A \setminus \{x\})$<br>(b) $L_{yt'}(A) > L_{yt'}(A \setminus \{x\})$ | (a) $L_{yt}(A) > L_{yt}(A \setminus \{x\})$<br>(b) $L_{yx}(A) > L_{yx}(A \setminus \{t'\})$ |
| 2 | $---$ | (a) $L_{xt}(A) > L_{xt}(A \setminus \{y\})$<br>(b) $L_{yt'}(A) > L_{yt'}(A \setminus \{x\})$ | (a) $L_{xt}(A) > L_{xt}(A \setminus \{y\})$<br>(b) $L_{yx}(A) > L_{yx}(A \setminus \{t'\})$ |
| 3 | $---$ | $---$ | (a) $L_{xy}(A) > L_{xy}(A \setminus \{t\})$<br>(b) $L_{yx}(A) > L_{yx}(A \setminus \{t'\})$ |

In case (1,1), (b) implies $L_{yt'}(A) = L_{yt'}(A \setminus \{x\})$ by R-Asymmetry. Dividing this inequality by (a) gives (a') $L_{tt'}(A) > L_{tt'}(A \setminus \{x\})$. Similarly, (a) implies $L_{tx}(A) = L_{tx}(A \setminus \{y\})$ by R-Asymmetry. Dividing this inequality by (b) gives (b') $L_{tt'}(A) > L_{tt'}(A \setminus \{y\})$. Applying the same kind of reasoning in the remaining cases gives the following:

| $(i,j)$ | 1 | 2 | 3 |
|---|---|---|---|
| 1 | (a') $L_{tt'}(A) > L_{tt}(A \setminus \{x\})$<br>(b') $L_{tt'}(A) < L_{tt'}(A \setminus \{y\})$ | (a') $L_{tt'}(A) > L_{tt'}(A \setminus \{x\})$<br>(b') $L_{tt'}(A) = L_{tt'}(A \setminus \{y\})$ | (a') $L_{tt'}(A) > L_{tt'}(A \setminus \{x\})$<br>(b') $L_{tt'}(A) = L_{tt'}(A \setminus \{y\})$ |
| 2 | $---$ | (a') $L_{tt'}(A) > L_{tt'}(A \setminus \{x\})$<br>(b') $L_{tt'}(A) < L_{tt'}(A \setminus \{y\})$ | (a') $L_{tt'}(A) = L_{tt'}(A \setminus \{x\})$<br>(b') $L_{tt'}(A) < L_{tt'}(A \setminus \{y\})$ |
| 3 | $---$ | $---$ | (a') $L_{tt'}(A) = L_{tt'}(A \setminus \{x\})$<br>(b') $L_{tt'}(A) = L_{tt'}(A \setminus \{y\})$ |

These two tables are sufficient to show that each of the six $(i,j)$ cases such that $i \leq j$ entails a contradiction:

**(1,1)** Since $L_{xt}(\{x,t\})/L_{yt}(\{y,t\}) > 0$, the combination of (a') and (b') contradicts R-Proportionality.

**(2,2)** By similar reasoning to case $(1,1)$, the combination of (a') and (b') contradicts R-Proportionality.

**(3,3)** By similar reasoning to case $(1,1)$, the combination of (a) and (b) contradicts R-Proportionality.

**(1,2)** By R-Asymmetry, (b') implies (b'.i) $L_{ty}(A) \neq L_{ty}(A \setminus \{t'\})$ or (b'.ii) $L_{yt'}(A) \neq L_{yt'}(A \setminus \{t\})$. **(b'.i)** By R-Asymmetry, (b) implies $L_{xy}(A) = L_{xy}(A \setminus \{t'\})$. So, $L_{tx}(A) \neq L_{tx}(A \setminus \{t'\})$. But this contradicts (a') by R-Asymmetry. **(b'.ii)** By R-Asymmetry, (a) implies $L_{xy}(A) = L_{xy}(A \setminus \{t\})$. So, $L_{xt'}(A) \neq L_{xt'}(A \setminus \{t\})$ which again contradicts (a').

**(1,3)** By similar reasoning to case $(1,2)$, the combination of (a') and (b') contradicts R-Asymmetry.

**(2,3)** By similar reasoning to case $(1,2)$, the combination of (a') and (b') contradicts R-Asymmetry.

**(iii)** Let $x \succ y \succ z$. Letting $A := \{x,y,z\}$, R-Asymmetry implies exactly one of: (a) $L_{xy}(A) \neq L_{xy}(A \setminus \{z\})$; (b) $L_{yz}(A) \neq L_{yz}(A \setminus \{x\})$; or, (c) $L_{xz}(A) \neq L_{xz}(A \setminus \{y\})$. I show that identities (a) and (b) lead to contradictions while (c) delivers the desired result. **(a)** If $L_{xy}(A) > L_{xy}(A \setminus \{z\})$, then $z \succ_1 y$ which contradicts $y \succ z$ by (ii) above. If $L_{xy}(A) < L_{xy}(A \setminus \{z\})$, then $y \succ_3 x$ which contradicts $x \succ y$ by (ii) above. **(b)** This identity entails similar contradictions. **(c)** If $L_{zx}(A) > L_{zx}(A \setminus \{y\})$, then $z \succ_2 y$ (and also $y \succ_1 x$) which contradicts $y \succ z$ (and also $x \succ y$) by (ii) above. So, $L_{xz}(A) > L_{xz}(A \setminus \{y\})$ which, by definition, gives $x \succ_3 z$. Consequently, $x \succ z$. ■

**Note 3** *If $p$ satisfies the axioms of Theorem 1, then, by Lemma 2, one can label the $n := |X|$ elements of $X := \{x_1, ..., x_n\}$ so that $x_i \succ x_j$ iff $i < j$. Using this convention, one can then define $\gamma(x_i)$ for $i = 1$, $i = 2$, and $3 \leq i \leq n$ as follows:*

$$\gamma(x_1) := \frac{\gamma(x_2)}{L_{x_2x_1} + \gamma(x_2)}, \quad \gamma(x_2) := 1 - \frac{L_{x_1x_3}}{L_{x_1x_2} \cdot L_{x_2x_3}}, \quad \text{and} \quad \gamma(x_i) := \frac{\gamma(x_2)}{L_{x_1x_i} \cdot L_{x_2x_1}}. \tag{4}$$

**Note 4** *For any set $A := \{a_1, ..., a_m\}$, I also abbreviate $L_{a_ia_j}$ by $L^A_{ij}$ in the next three results.*

**Lemma 3** *If $p$ satisfies satisfies the axioms of Theorem 1, then $\gamma(x) \in (0,1)$ for all $x \in X$.*

**Proof.** By Lemma 2, $L^X_{12} \cdot L^X_{23} = L_{x_1x_3}(\{x_1,x_2,x_3\}) > L^X_{13} > 0$. By definition of $\gamma(x_2)$, it then follows that $\gamma(x_2) \in (0,1)$. Similarly, $\gamma(x_1) \in (0,1)$ since $L^X_{21} > 0$ and $\gamma(x_2) \in (0,1)$. What is more, for $i \geq 3$, $\gamma(x_i) > 0$ since $L^X_{1i}, L^X_{21} > 0$ and $\gamma(x_2) \in (0,1)$. To complete the proof, I show that $\gamma(x_i) < 1$. First, observe that Lemma 2 and R-Independence imply

$$L^X_{1i} \cdot L^X_{i2} = \frac{L_{x_1x_i}(\{x_1,x_i\}) \cdot L^X_{12}}{L_{x_1x_i}(\{x_1,x_2,x_i\})} = \frac{L_{x_1x_i}(\{x_1,x_3,x_i\}) \cdot L^X_{12}}{L_{x_1x_i}(\{x_1,x_2,x_3,x_i\})} = \frac{L_{x_1x_3}(\{x_1,x_3\}) \cdot L_{x_3x_i}(\{x_3,x_i\}) \cdot L^X_{12}}{L_{x_1x_3}(\{x_1,x_2,x_3\}) \cdot L_{x_3x_i}(\{x_3,x_i\})} = L^X_{13} \cdot L^X_{32}. \tag{5}$$

Combining (5) with the definitions of $\gamma(x_2)$ and $\gamma(x_i)$ in (4) gives

$$\gamma(x_i) := \frac{\gamma(x_2)}{L^X_{1i} \cdot L^X_{21}} = \frac{1 - (L^X_{13} \cdot L^X_{32}) \cdot L^X_{21}}{L^X_{1i} \cdot L^X_{21}} = \frac{1 - (L^X_{1i} \cdot L^X_{i2}) \cdot L^X_{21}}{L^X_{1i} \cdot L^X_{21}}. \tag{6}$$

By way of contradiction, suppose that $\gamma(x_i) \geq 1$. Then, (6) implies $1 \geq L^X_{1i} \cdot L^X_{21} \cdot (1 + L^X_{i2})$. Letting $A = \{x_1, x_2, x_i\}$:

$$1 \geq L^X_{1i} \cdot L^X_{21} \cdot (1 + L^X_{i2}) = L^X_{1i} \cdot \left(\frac{p(x_2, A)}{p(x_1, A)}\right) \cdot \left(\frac{p(x_2, A) + p(x_i, A)}{p(x_2, A)}\right) = L_{1i} \cdot \left(\frac{1 - p(x_1, A)}{p(x_1, A)}\right) \tag{7}$$

since $L^X_{12} = L_{x_1x_2}(\{x_1, x_2, x_i\})$ and $L^X_{2i} = L_{x_2x_i}(\{x_1, x_2, x_i\})$. Re-arranging (7) gives $p(x_1, \{x_1, x_2, x_i\}) \cdot (1 + L^X_{1i}) \geq L^X_{1i}$ or, equivalently, $p(x_1, \{x_1, x_2, x_i\}) \geq p(x_1, \{x_1, x_i\})$. But, this contradicts Strict Regularity. ■

**Note 5** *For the next two results, I label the elements of any set $A := \{a_1, ..., a_m\} \in \mathcal{X}$ so that $a_i \succ a_j$ iff $i < j$.*

**Lemma 4** *If $p$ satisfies the axioms of Theorem 1, then $p(\cdot, A) = p_{(\gamma,\succ)}(\cdot, A)$ for $A \in \mathcal{X}$ if*

$$L^A_{jj+1} = \frac{\gamma(a_j)}{\gamma(a_{j+1})[1 - \gamma(a_j)]} \text{ for all } j \in \{1, ..., m-1\}. \tag{8}$$

**Proof.** Since $L_{a_ia_{i+1}}(A) = L^A_{ii+1}$ by Lemma 2, it follows that

$$L_{a_ia_m}(A) = \frac{p(a_i, A)}{p(a_m, A)} \cdot \frac{p(a_{i+1}, A)}{p(a_{i+1}, A)} = L_{a_ia_{i+1}}(A) \cdot L_{a_{i+1}a_m}(A) = L^A_{ii+1} \cdot L_{a_{i+1}a_m}(A) = ... = \prod_{i \leq j \leq m} L^A_{jj+1}. \tag{9}$$

Combining (9) with the formula $(\Sigma_{1\leq i\leq m} L_{a_i,a_m}(A)) \cdot p(a_m, A) = \Sigma_{1\leq i\leq m} p(a_i, A) = 1$ gives

$$\left(\Sigma_{1\leq i\leq m} \prod_{i\leq j\leq m} L^A_{jj+1}\right) \cdot p(a_m, A) = 1 \tag{10}$$

Replacing the formulas from (8) into (10) and applying the identity in (1) then gives

$$p(a_m, A) = \frac{\gamma(a_m) \cdot \prod_{1\leq i\leq m-1}[1 - \gamma(a_i)]}{1 - \prod_{1\leq i\leq m-1}[1 - \gamma(a_i)]} = p_{(\gamma,\succ)}(a_m, A). \tag{11}$$

So, $p(\cdot, A) = p_{(a_i,\succ)}(\cdot, A)$ obtains by combining $p(a_i, A) = L^A_{ii+1} \cdot p(a_{i+1}, A)$ with (11) and the formulas from (8). ■

**Proof of Theorem 1.** (**Necessity**) It is clear that the axioms are necessary to represent $p$ by an independent RCM.

**(Sufficiency)** By Lemma 2, $\succ$ is a linear order; and, by Lemma 3, $\gamma$ is a consideration function. To complete the proof, I show, by induction on $|X| := N$ for $N \geq 4$, that $p_{(\gamma,\succ)}(\cdot, A) = p(\cdot, A)$ for all $A \in \mathcal{X}$.

**Base cases $N = 3$ and $N = 4$:** To apply Lemma 4, it suffices to establish the identities in (8) for each $A \in \mathcal{X}$:

<u>$A = \{x_1, x_2\}$</u>: The definition of $\gamma(x_1)$ in (4) may be re-arranged to obtain the required identity

$$L^X_{12} = \frac{\gamma(x_1)}{\gamma(x_2)[1 - \gamma(x_1)]}. \tag{12}$$

<u>$A = \{x_1, x_i\}$ for $i = 3, 4$</u>: Combining the definitions of $\gamma(x_1)$ and $\gamma(x_i)$ in (4) gives the required identity

$$L^X_{1i} = \frac{\gamma(x_2)}{\gamma(x_i) \cdot L^X_{21}} = \frac{\gamma(x_2)}{\gamma(x_i) \cdot \frac{\gamma(x_2)[1-\gamma(x_1)]}{\gamma(x_1)}} = \frac{\gamma(x_1)}{\gamma(x_i) \cdot [1 - \gamma(x_1)]}. \tag{13}$$

<u>$A = \{x_2, x_3\}$</u>: Combined with (12) and (13), the definition of $\gamma(x_2)$ in (4) gives the required identity

$$L^X_{23} = \frac{L^X_{13}}{L^X_{12} \cdot [1 - \gamma(x_2)]} = \frac{\frac{\gamma(x_1)}{\gamma(x_3)\cdot[1-\gamma(x_1)]}}{\frac{\gamma(x_1)}{\gamma(x_2)[1-\gamma(x_1)]} \cdot [1 - \gamma(x_2)]} = \frac{\gamma(x_2)}{\gamma(x_3) \cdot [1 - \gamma(x_2)]}. \tag{14}$$

Identities (12) and (14) are sufficient to establish the case $N = 3$. The case $N = 4$ requires two additional steps:

$\underline{A = \{x_2, x_4\}}$: Combined with (13) and (14), formula (5) (from the proof of Lemma 3) gives the required identity

$$L^X_{24} = \frac{L^X_{14} \cdot L^X_{23}}{L^X_{13}} = \frac{\frac{\gamma(x_1)}{\gamma(x_4)\cdot[1-\gamma(x_1)]} \cdot \frac{\gamma(x_2)}{\gamma(x_3)\cdot[1-\gamma(x_2)]}}{\frac{\gamma(x_1)}{\gamma(x_3)[1-\gamma(x_1)]}} = \frac{\gamma(x_2)}{\gamma(x_4)\cdot[1-\gamma(x_2)]}. \tag{15}$$

$\underline{A = \{x_3, x_4\}}$: First, observe that R-Proportionality implies

$$\frac{L^X_{14} - L^X_{12} \cdot L^X_{24}}{L^X_{14} - L^X_{13} \cdot L^X_{34}} = \frac{L^X_{14} - L^X_{x_1 x_4}(\{x_1, x_2, x_4\})}{L^X_{14} - L^X_{x_1 x_4}(\{x_1, x_3, x_4\})} = \frac{L^X_{24}}{L^X_{34}} \tag{16}$$

or $L^X_{34} \cdot (L^X_{14} + L^X_{13} \cdot L^X_{24} - L^X_{12} \cdot L^X_{24}) = L^X_{14} \cdot L^X_{24}$. Combined with (12), (13) and (15), this gives the required identity

$$L^X_{34} = \frac{L^X_{14} \cdot L^X_{24}}{L^X_{14} + L^X_{13} \cdot L^X_{24} - L^X_{12} \cdot L^X_{24}} = \frac{1 \cdot \frac{\gamma(x_2)}{\gamma(x_4)\cdot[1-\gamma(x_2)]}}{1 + \frac{\gamma(x_4)}{\gamma(x_3)} \cdot \frac{\gamma(x_2)}{\gamma(x_4)\cdot[1-\gamma(x_2)]} - \frac{\gamma(x_4)}{\gamma(x_2)} \cdot \frac{\gamma(x_2)}{\gamma(x_4)\cdot[1-\gamma(x_2)]}} = \frac{\gamma(x_3)}{\gamma(x_4)[1-\gamma(x_3)]}. \tag{17}$$

Identities (12)-(17) are sufficient to establish the case $N = 4$.

**Induction Step:** Suppose that the result holds for $N = n-1 \geq 4$. For $N = n$, let $X_{-n} := X \setminus \{x_n\}$. By the induction hypothesis, the identities in (8) hold for all $A \in \mathcal{X}_{-n}$. To show that the identities in (8) hold for all $A \in \mathcal{X} \setminus \mathcal{X}_{-n}$, it suffices to establish these identities for all sets $A = \{x_k, x_n\}$ such that $k \neq n$. The result then follows by Lemma 4.

$\underline{A = \{x_1, x_n\}}$: Reasoning as in the case $A = \{x_1, x_i\}$ for $i = 3, 4$ above, it follows that

$$L^X_{1n} = \frac{\gamma(x_2)}{\gamma(x_n) \cdot L^X_{21}} = \ldots = \frac{\gamma(x_1)}{\gamma(x_n)\cdot[1-\gamma(x_1)]}.$$

$\underline{A = \{x_k, x_n\} \text{ for } k \neq 1, n-1}$: Reasoning as in the case $A = \{x_2, x_4\}$ above, it follows that

$$L^X_{kn} = \frac{L^X_{1n} \cdot L^X_{kn-1}}{L^X_{1n-1}} = \ldots = \frac{\gamma(x_k)}{\gamma(x_n)\cdot[1-\gamma(x_k)]}.$$

$\underline{A = \{x_{n-1}, x_n\}}$: Reasoning as in the case $A = \{x_3, x_4\}$ above, it follows that

$$L^X_{n-1n} = \frac{L^X_{1n} \cdot L^X_{2n}}{L^X_{1n} + L^X_{1n-1} \cdot L^X_{2n} - L^X_{12} \cdot L^X_{2n}} = \ldots = \frac{\gamma(x_{n-1})}{\gamma(x_n)[1-\gamma(x_{n-1})]}.$$

Combined with the induction hypothesis, this shows that the identities in (8) hold for all $A \in \mathcal{X}$. ■

**Note 6** *The definition of $\gamma$ given in (4) is equivalent to the definition of $\gamma$ in ($\gamma$-1)-($\gamma$-2). To see this, let $t := x_2$, $\overline{x} := x_1$ and $\underline{x} := x_n$. Then, using the identities (5) and (16), the definitions in (4) may be re-written as follows:*

$$\gamma(x_1) := \frac{\gamma(x_2)}{L_{x_2x_1} + \gamma(x_2)} = \frac{\gamma(t)}{L_{[t:\overline{x}]}(\emptyset) + \gamma(t)};$$

$$\gamma(x_2) := 1 - \frac{L_{x_1x_3}}{L_{x_1x_2} \cdot L_{x_2x_3}} = 1 - \frac{L_{x_1x_n}}{L_{x_1x_2} \cdot L_{x_2x_n}} = 1 - R_{[\overline{x}:\underline{x}]}(A; x_2);$$

$$\gamma(x_n) := \frac{\gamma(x_2)}{L_{x_1x_n} \cdot L_{x_2x_1}} = \frac{\gamma(x_2) \cdot L_{x_1x_2} \cdot L_{x_2x_3}}{L_{x_2x_n} \cdot L_{x_1x_3}} = \frac{\gamma(x_2)}{L_{x_2x_n}[1 - \gamma(x_2)]} = \frac{\gamma(t)}{L_{[t:\underline{x}]}(\emptyset)[1 - \gamma(t)]};\ \textit{and, for all } i \in \{3, ..., n\},$$

$$\gamma(x_i) := \frac{\gamma(x_2)}{L_{x_1x_i} \cdot L_{x_2x_1}} = \frac{1 - \frac{L_{x_1x_3}}{L_{x_1x_2} \cdot L_{x_2x_3}}}{L_{x_1x_i} \cdot L_{x_2x_1}} = \frac{L_{x_1x_2} \cdot L_{x_ix_n} - L_{x_1x_n} \cdot L_{x_nx_2} \cdot L_{x_1x_n}}{L_{x_1x_i} \cdot L_{x_ix_n}} = \frac{L_{x_1x_i} \cdot L_{x_ix_n} - L_{x_1x_n}}{L_{x_1x_i} \cdot L_{x_ix_n}} = 1 - R_{[\overline{x}:\underline{x}]}(A; x_i).$$

**Proof of Theorem 2.** Suppose that $(\gamma, \succ)$ and $(\gamma', \succ')$ both represent $p$. By way of contradiction, first suppose that $\succ$ and $\succ'$ differ on $A = \{a, b, c\}$. Without loss of generality, suppose that $a \succ b \succ c$. If $c \succ' b \succ' a$, then

$$1 - \gamma(b) = \frac{L_{ac}(A \setminus \{b\})}{L_{ac}(A)} = \frac{1}{1 - \gamma'(b)}$$

by equation ($\gamma$-1). Since $\gamma(b) \in (0, 1)$ by construction, it follows that $1 - \gamma'(b) > 1$, which contradicts $\gamma'(b) > 0$. In any other case where $\succ$ and $\succ'$ differ on $A$, the middle alternative in $\succ'$ is either $a$ or $c$. Since $(\gamma, \succ)$ and $(\gamma', \succ')$ both represent $p$, it follows that $L_{ab}(A) = L_{ab}(A \setminus \{c\})$, $L_{bc}(A) = L_{bc}(A \setminus \{a\})$, and $L_{ac}(A) = L_{ac}(A \setminus \{b\})$ [where the first two equations are based on $(\gamma, \succ)$ and last one on $(\gamma', \succ')$]. But this contradicts R-Asymmetry. So, $\succ = \succ'$.

To show that $\gamma = \gamma'$, fix any alternative $x \in X \setminus \{\overline{x}, \underline{x}\}$. From equation ($\gamma$-1) and R-Independence, it follows that

$$1 - \gamma(x) = \frac{L_{\overline{x}\underline{x}}(\{\overline{x}, \underline{x}\})}{L_{\overline{x}\underline{x}}(\{\overline{x}, x, \underline{x}\})} = 1 - \gamma'(x)$$

so that $\gamma(x) = \gamma'(x)$. Using equations ($\gamma$-2), it then follows that $\gamma(\overline{x}) = \gamma'(\overline{x})$ and $\gamma(\underline{x}) = \gamma'(\underline{x})$ as well. ■

## 7.2 Proof of Propositions 1-2

For the reader's convenience, I first state the implications of my R-Asymmetry, R-Independence, and R-Proportionality axioms for the no-choice option $x^*$. For R-Asymmetry, I follow the notational convention that $R_{[a:b]}(A; x^*) := 1$.

**R-Asymmetry for** $x^*$ For all standard menus $A \in \mathcal{X}$ and distinct alternatives $a, b \notin A$:

$$R_{[a:x^*]}(A; b) \neq 1 \iff R_{[b:x^*]}(A; a) = 1.$$

**R-Independence for** $x^*$ For all standard menus $A, B \in \mathcal{X}$ and distinct alternatives $a, b \notin A \cup B$:

$$R_{[a:x^*]}(A; b) = R_{[a:x^*]}(B; b).$$

**R-Proportionality for** $x^*$ For all standard menus $A \in \mathcal{X}$ and distinct alternatives $a, c, d \notin A$:

$$R_{[a:x^*]}(A; c), R_{[a:x^*]}(A; d) \neq 1 \implies \frac{L_{[a:x^*]}(A) - L_{[a:x^*]}(A \cup \{c\})}{L_{[a:x^*]}(A) - L_{[a:x^*]}(A \cup \{d\})} = \frac{L_{[c:x^*]}(A)}{L_{[d:x^*]}(A)}.$$

**Lemma 5** *If $p^*$ satisfies the axioms of Proposition 1, then $\widehat{p}^*$ satisfies the axioms of Theorem 1.*

**Proof.** It is clear that $\widehat{p}^*$ inherits L-Asymmetry, L-Independence and L-Proportionality from $p^*$.

To see that $\widehat{p}^*$ satisfies Strict Regularity, fix $A^* \in \mathcal{X}^* \setminus \mathcal{X}$ and $b \in A$. Then, $p^*(x^*, A^* \setminus \{b\}) \geq p^*(x^*, A^*)$ by R-Regularity (i); and, by R-Regularity (ii), $p^*(x^*, A^* \setminus \{b\}) \neq p^*(x^*, A^*)$ since $L_{x^*x^*}(A^* \setminus \{b\}) = 1 = L_{x^*x^*}(A^*)$. So, $p^*(x^*, A^* \setminus \{b\}) > p^*(x^*, A^*)$. Similarly, by R-Regularity (i), $p^*(a, A^* \setminus \{b\}) \geq p^*(a, A^*)$ for all $a \in A \setminus \{b\}$. Combining these last two identities gives $\widehat{p}^*(a, A \setminus \{b\}) = \frac{p^*(a,A^* \setminus \{b\})}{1-p^*(x^*,A^* \setminus \{b\})} > \frac{p^*(a,A^*)}{1-p^*(x^*,A^*)} = \widehat{p}^*(a, A)$ for all distinct $a, b \in A$. ■

**Lemma 6** *If $p^*$ satisfies the axioms of Proposition 1 and $\succ$ is the revealed preference for $\widehat{p}^*$, then:*

$$x \succ y \text{ if and only if } p^*(x, A^*) = p^*(x, A^* \setminus \{y\}) \text{ for all } A^* \in \mathcal{X}^* \setminus \mathcal{X} \text{ such that } x, y \in A.$$

**Proof.** Consider the relation $\succ^*$ defined by $x \succ^* y$ if $p^*(x, A^*) = p^*(x, A^* \setminus \{y\})$ for all $A^* \in \mathcal{X}^* \setminus \mathcal{X}$ such that $x, y \in A$. I show that: (i) $\succ^*$ is a linear order; and (ii) $\succ^*$ coincides with $\succ$.

**(i)** To see that $\succ^*$ is complete and asymmetric, fix $A^*, B^* \in \mathcal{X}^* \setminus \mathcal{X}$ such that $x, y \in A^* \cap B^*$. Then:

$$\begin{aligned} p^*(x, A^* \setminus \{y\}) = p^*(x, A^*) \quad &\underset{\text{R-Reg.}}{\iff} L_{yx^*}(A^* \setminus \{x\}) \neq L_{yx^*}(A^*) \\ &\underset{\text{R-Ind.}}{\iff} L_{yx^*}(B^* \setminus \{x\}) \neq L_{yx^*}(B^*) \\ &\underset{\text{R-Asy.}}{\iff} L_{xx^*}(B^* \setminus \{y\}) = L_{xx^*}(B^*) \quad \underset{\text{R-Reg.}}{\iff} p^*(y, B^* \setminus \{x\}) \neq p^*(y, B^*). \end{aligned}$$

For transitivity, suppose that $x \succ^* y \succ^* z \succ^* x$ and fix a menu $A^* \subseteq X^*$ such that $x, y, z \in A$. By definition of $\succ^*$ and the same kind of reasoning used above: $L_{xx^*}(A^* \setminus \{z\}) = L_{xx^*}(A^*)$; $L_{yx^*}(A^* \setminus \{x\}) = L_{yx^*}(A^*)$; and, $L_{zx^*}(A^* \setminus \{y\}) = L_{zx^*}(A^*)$. Without loss of generality, suppose that $L_{yz}(A^* \setminus \{x\}) = L_{yz}(A^*)$. Since $L_{yx^*}(A^* \setminus \{x\}) = L_{yx^*}(A^*)$, $L_{zx^*}(A^* \setminus \{x\}) = L_{zx^*}(A^*)$. But, this contradicts $L_{xx^*}(A^* \setminus \{z\}) = L_{xx^*}(A^*)$ by R-Asymmetry.

**(ii)** Suppose that $x \succ^* y \succ^* z$ and fix $A^* = \{x, y, z\}$. By definition, $p^*(x, A^*) = p^*(x, A^* \setminus \{y\})$ and $p^*(z, A^*) \neq p^*(z, A^* \setminus \{y\})$. By R-Regularity, this last inequality gives $p^*(z, A^*) < p^*(z, A^* \setminus \{y\})$. Dividing by $p^*(x, A^*) = p^*(x, A^* \setminus \{y\})$ gives $L_{xz}(A^*) > L_{xz}(A^* \setminus \{y\})$. So, by definition of $\succ$ and $\widehat{p}^*$, it follows that $x \succ y \succ z$. ■

**Lemma 7** *If $p^*$ satisfies the axioms of Proposition 1, $p^*(x^*, A^*) = \prod_{x \in A} p^*(x^*, \{x, x^*\})$ for all $A^* \in \mathcal{X}^* \setminus \mathcal{X}$.*

**Proof.** The proof is by induction on $|A|$. By Lemma 6, one can label the elements of $A$ so that $a_i \succ a_j$ iff $i < j$.

**Base Case:** For $A = \{a_1, a_2\}$, R-Regularity and R-Asymmetry imply

$$\frac{p^*(a_2, A^* \setminus \{a_1\})}{p^*(x^*, A^* \setminus \{a_1\})} = \frac{p^*(a_2, A^*)}{p^*(x^*, A^*)} \tag{18}$$

What is more, $p^*(a_1, A^*) = p^*(a_1, A^* \setminus \{a_2\})$ by Lemma 6. Combining this identity with the requirement that $p^*(x^*, A^*) + p^*(a_1, A^*) + p^*(a_2, A^*) = 1$ and replacing into equation (18) gives

$$\frac{p^*(a_2, A^* \setminus \{a_1\})}{p^*(x^*, A^* \setminus \{a_1\})} = \frac{1 - p^*(x^*, A^*) - p^*(a_1, A^* \setminus \{a_2\})}{p^*(x^*, A^*)}. \tag{19}$$

Re-arranging (19) gives $p^*(x^*, A^*) = p^*(x^*, \{a_1, x^*\}) \cdot p^*(x^*, \{a_2, x^*\})$, as required.

**Induction Step:** Suppose that the result holds for $|A| = n - 1$ and consider the case $|A| = n$. As in the base case,

$$\frac{p^*(a_i, A^* \setminus \{a_1\})}{p^*(x^*, A^* \setminus \{a_1\})} = \frac{p^*(a_i, A^*)}{p^*(x^*, A^*)} \tag{20}$$

for all $a_i \in A \setminus \{a_1\}$. Summing the equations of the form (20) over the alternatives in $A \setminus \{a_1\}$ gives

$$\frac{1-p^*(x^*, A^* \setminus \{a_1\})}{p^*(x^*, A^* \setminus \{a_1\})} = \frac{1-p^*(x^*, A^*) - q(a_1, A^*)}{p^*(x^*, A^*)} \quad \text{or} \quad p^*(x^*, A^*) = [1-p^*(a_1, A^*)] \cdot p^*(x^*, A \setminus \{a_1\}). \tag{21}$$

By Lemma 6, $p^*(a_1, A^*) = ... = p^*(a_1, \{a_1, x^*\})$. Applying this identity and the induction hypothesis to (21) gives

$$p^*(x^*, A^*) = [1-p^*(a_1, \{a_1, x^*\})] \cdot p^*(x^*, A \setminus \{a_1\}) = \prod_{x \in A} p^*(x^*, \{x, x^*\}), \tag{22}$$

which completes the proof. ■

**Lemma 8** *If $p^*$ satisfies the axioms of Proposition 1, then $p^*(x^*, \{x, x^*\}) = 1 - \gamma(x)$ for all $x \in X$.*

**Proof.** For $x \in X \setminus \{\overline{x}, \underline{x}\}$, pick test alternatives $b, w \in X$ such that $b \succ x \succ w$. Then, by Lemma 6 and R-Regularity, it follows that $L_{wx^*}(\{b, x, w, x^*\}) = L_{wx^*}(\{b, w, x^*\})$ and $p^*(b, \{b, x, w, x^*\}) = p^*(b, \{b, w, x^*\})$ so that

$$\frac{L_{bw}(\{b, w\})}{L_{bw}(\{b, x, w\})} = \frac{\frac{L_{bx^*}(\{b,w,x^*\})}{L_{wx^*}(\{b,w,x^*\})}}{\frac{L_{bx^*}(\{b,x,w,x^*\})}{L_{wx^*}(\{b,x,w,x^*\})}} = \frac{L_{bx^*}(\{b, w, x^*\})}{L_{bx^*}(\{b, x, w, x^*\})} = \frac{\frac{p^*(b,\{b,w,x^*\})}{p^*(x^*,\{b,w,x^*\})}}{\frac{p^*(b,\{b,x,w,x^*\})}{p^*(x^*,\{b,x,w,x^*\})}} = \frac{p^*(x^*, \{b, w, x^*\})}{p^*(x^*, \{b, x, w, x^*\})}. \tag{23}$$

Since $p^*(x^*, A^*) = \prod_{a \in A} p^*(x^*, \{a, x^*\})$ by Lemmas 6-7,

$$\frac{p^*(x^*, \{b, w, x^*\})}{p^*(x^*, \{b, x, w, x^*\})} = \frac{p^*(x^*, \{b, x^*\}) \cdot p^*(x^*, \{w, x^*\})}{p^*(x^*, \{b, x^*\}) \cdot p^*(x^*, \{x, x^*\}) \cdot p^*(x^*, \{w, x^*\})} = p^*(x^*, \{x, x^*\}). \tag{24}$$

By Lemma 5 and Theorem 1, $\widehat{p}^*$ is represented by $(\gamma, \succ)$. Combined with identity ($\gamma$-1) of the text, (23)-(24) give

$$1 - \gamma(x) = \frac{L_{bw}(\{b, w\})}{L_{bw}(\{b, x, w\})} = \frac{p^*(x^*, \{b, w, x^*\})}{p^*(x^*, \{b, x, w, x^*\})} = p^*(x^*, \{x, x^*\}). \tag{25}$$

For $\overline{x}$ and $\underline{x}$, pick a test alternative $t \in X \setminus \{\overline{x}, \underline{x}\}$. From the identities in ($\gamma$-2) of the text,

$$\frac{\gamma(\overline{x})}{1-\gamma(\overline{x})} = \gamma(t) \cdot L_{\overline{x}t}(t, \overline{x}) \quad \text{and} \quad \gamma(\underline{x}) = \frac{\gamma(t)}{L_{t\underline{x}}(t, \underline{x}) \cdot [1-\gamma(t)]}. \tag{26}$$

Since $1 - \gamma(t) = p^*(x^*, \{t, x^*\})$ by (25), the same kind of reasoning as above can be used to show that

$$\gamma(t) \cdot L_{\overline{x}t}(\{t, \overline{x}\}) = ... = \frac{p^*(\overline{x}, \{\overline{x}, x^*\})}{p^*(x^*, \{\overline{x}, x^*\})} \quad \text{and} \quad \frac{\gamma(t)}{L_{t\underline{x}}(\{t, \underline{x}\}) \cdot [1-\gamma(t)]} = ... = p^*(\underline{x}, \{\underline{x}, x^*\}). \tag{27}$$

Combining the identities in (26) and (27) gives

$$\gamma(\overline{x}) \cdot [1 - p^*(\overline{x}, \{\overline{x}, x^*\})] = [1-\gamma(\overline{x})] \cdot p^*(\overline{x}, \{\overline{x}, x^*\}) \quad \text{and} \quad \gamma(\underline{x}) = p^*(\underline{x}, \{\underline{x}, x^*\}). \tag{28}$$

By simplifying the identities in (28), one obtains $p^*(x^*, \{x, x^*\}) = 1 - \gamma(x)$ for $x \in \{\underline{x}, \overline{x}\}$, as required. ■

**Proof of Proposition 1.** (**Necessity**) It is clear that any independent RCM with default $p^*$ satisfies the axioms.

(**Sufficiency**) By Lemma 5 and Theorem 1, $\widehat{p}^*$ is independent RCM and uniquely represented by some pair $(\gamma, \succ)$. So:

$$p^*(a, A^*) = \gamma(a) \cdot \prod_{x \in A: x \succ a} [1-\gamma(x)] \cdot \left[\frac{1-p^*(x^*, A^*)}{1-\prod_{x \in A}[1-\gamma(x)]}\right]. \tag{29}$$

for all $A^* \in \mathcal{X}^* \setminus \mathcal{X}$ and $a \in A$. Since $p^*(x^*, A^*) = \prod_{x \in A} p^*(x^*, \{x^*, x\})$ by Lemma 7 and $p^*(x^*, \{x, x^*\}) = 1 - \gamma(x)$ by Lemma 8, the bracketed term in the identity (29) is 1. This establishes that $p^*$ is independent RCM. ■

**Proof of Proposition 2.** **(I)** The 'if' direction is obvious. For the 'only if' direction, suppose that $(\gamma, \succ)$ represents $p$. Define the choice function with default $p^*_\gamma$ by

$$p^*_\gamma(a, A^*) := \begin{cases} p(a, A) \cdot (1 - \prod_{x \in A}[1 - \gamma(x)]) & \text{for } a \neq x^* \\ \prod_{x \in A}[1 - \gamma(x)] & \text{for } a = x^*. \end{cases}$$

It is easy to check that $p$ is the standardization of $p^*_\gamma$. What is more, the pair $(\gamma, \succ)$ represents $p^*_\gamma$.

**(II)** Suppose that $q^* \neq p^*_\gamma$ is an independent RCM with default whose standardization is $p$. Since $q$ is represented by $(\gamma', \succ') \neq (\gamma, \succ)$, the pair $(\gamma', \succ')$ must also represent $p$, which contradicts the uniqueness established in Theorem 1. ■

## 7.3 Proof of Proposition 3 and Theorems 3-4

**Proof of Proposition 3.** Suppose that $p^*$ is represented by $(\gamma, \succ)$.

Label the alternatives of $X = \{x_1, ..., x_m\}$ so that $x_i \succ x_j$ iff $i < j$; and let $x_{m+1} := x^*$. For each $x_i \in X^*$, define a random variable $\omega_i$ on $\{0, 1\}$ with probability $\mu_i$ given by $\mu_i(\omega_i = 0) := \gamma(x_i)$ and $\mu_i(\omega_i = 1) := 1 - \gamma(x_i)$. Denote the state space by $\Omega^* := \{0, 1\}^{m+1}$ and a realization of the random variables by $\omega$. Let $\mu := \prod_{i=1}^{m+1} \mu_i$ denote the product measure on $\Omega^*$. Fix some $\epsilon \in (0, \frac{1}{2})$ and define the utility function $v : X^* \times \Omega^* \to [0, 1]$ by

$$v(x_i, \omega) := \begin{cases} (1 - \frac{i}{m}\epsilon)/2^{\omega_i} & \text{if } i \neq m+1 \\ 1/2 & \text{if } i = m+1. \end{cases}$$

By construction, $(\mu, v)$ defines an independent random utility rule.

I show that $p^*_{(\gamma,\succ)} = p^*_{(\mu,v)}$. First observe that:

(i) for $i < j$, $v(x_i, \omega) > v(x_j, \omega)$ iff $\omega_i = 0$; and,

(ii) for $j < i$, $v(x_i, \omega) > v(x_j, \omega)$ iff $\omega_i = 1$ and ($\omega_j = 0$ or $j = m + 1$).

By applying this observation to the definition $p^*_{(\mu,v)}(x_i, A) := \mu(\{\omega : \omega \in \Omega^*_v(x_i, A)\}) = \Sigma_{\omega \in \Omega^*_v(x_i,A)} \mu(\omega)$ one obtains

$$p^*_{(\mu,v)}(x_i, A) = \mu(\omega : \omega_i = 0) \cdot \left( \prod_{x_j \in A : j < i} \mu(\omega : \omega_j = 1) \right) \cdot \left( \prod_{x_j \in A : j \geq i} \mu(\omega : \omega_j \geq 0) \right) = \gamma(x_i) \cdot \prod_{y \in A : y \succ x_i} [1 - \gamma(y)].$$

Since $x_i$ and $A$ were chosen arbitrarily, this establishes the desired result. ■

**Proof of Theorem 3.** Suppose that $p$ is represented by $(\gamma, \succ)$.

Label the alternatives of $X = \{x_1, ..., x_m\}$ as in Proposition 3; and define a random variable $\omega_i$ on $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ with probability $\mu_i$ given by $\mu_i(\omega_i = k) := \gamma(x_i)[1 - \gamma(x_i)]^k$. Denote the state space by $\Omega := (\mathbb{N}_0)^m$ and a realization of the random variables by $\omega$. Let $\mu := \prod_{i=1}^m \mu_i$ denote the product measure on $\Omega$. Define the utility function $v : X \times \Omega \to [0, 1]$ by $v(x_i, \omega) := (1 - \frac{i}{m}\epsilon)/2^{\omega_i}$ (as in Proposition 3). By construction, $(\mu, v)$ defines an independent random utility rule.

I show that $p_{(\gamma,\succ)} = p_{(\mu,v)}$. First observe that:

(i) for $i < j$, $v(x_i, \omega) > v(x_j, \omega)$ iff $\omega_i \leq \omega_j$; and,

(ii) for $j < i$, $v(x_i, \omega) > v(x_j, \omega)$ iff $\omega_i < \omega_j$.

By applying this observation to the definition $p_{(\mu,v)}(x_i, A) := \mu(\{\omega : \omega \in \Omega_v(x_i, A)\}) \ = \Sigma_{\omega \in \Omega_v(x_i,A)} \mu(\omega)$ one obtains

$$p_{(\mu,v)}(x_i, A) = \Sigma_{k=0}^{\infty} \mu(\omega : \omega_i = k) \cdot \left( \prod_{x_j \in A : j < i} \mu(\omega : \omega_j > k) \right) \cdot \left( \prod_{x_j \in A : j \geq i} \mu(\omega : \omega_j \geq k) \right)$$

$$= \gamma(x_i) \cdot \prod_{x_j \in A : j < i} [1 - \gamma(x_j)] \cdot \left[ \Sigma_{k=0}^{\infty} [1 - \gamma(x_i)]^k \cdot \left( \prod_{x_j \in A \setminus \{x_i\}} \mu(\omega : \omega_j \geq k) \right) \right].$$

Since $\mu(\omega : \omega_j \geq k) = [1 - \gamma(x_j)]^k \Sigma_{l=0}^{\infty} \mu(\omega_j = l) = [1 - \gamma(x_j)]^k$, the preceding identity implies

$$p_{(\mu,v)}(x_i, A) = \gamma(x_i) \cdot \prod_{y \in A : y \succ x_i} [1 - \gamma(y)] \cdot \left[ \Sigma_{k=0}^{\infty} \prod_{y \in A} [1 - \gamma(y)]^k \right].$$

Moreover, since $\Sigma_{k=0}^{\infty} \prod_{y \in A} [1 - \gamma(y)]^k = (1 - \prod_{y \in A} [1 - \gamma(y)])^{-1}$, it follows that

$$p_{(\mu,v)}(x_i, A) = \frac{\gamma(x_i) \cdot \prod_{y \in A : y \succ x_i} [1 - \gamma(y)]}{1 - \prod_{y \in A} [1 - \gamma(y)]} = p_{(\gamma,\succ)}(x_i, A).$$

Since $x_i$ and $A$ were chosen arbitrarily, this establishes the desired result. ■

**Proof of Theorem 4.** Suppose that $p^*$ has a random utility representation $(\mu, v)$. Label the alternatives of $X = \{x_1, ..., x_m\}$ with $x_{m+1} := x^*$. Without loss of generality, let the state space be the set of permutations $\mathcal{P}$ on $\{1, ..., m+1\}$; and $v : X^* \times \mathcal{P} \to \mathbb{R}$ to be the canonical utility function (on $X^*$) defined by $v(x_i, \pi) := \pi(i)$ for all $\pi \in \mathcal{P}$.

Let $\succ_\pi$ denote the linear order defined by $v(\cdot, \pi)$ and $\succ^*$ the *initial segment* that restricts $\succ$ to the upper-contour set $U(x^*, \succ^*) := \{x \in X : x \succ x^*\}$. Then, define the induced distribution $\mu^*$ on initial segments where $\mu^*(\succ^*) := \mu(\{\pi \in \mathcal{P} : \succ_\pi^* = \succ^*\})$ (i.e., $\mu^*(\succ^*)$ is the probability of drawing a linear order with initial segment $\succ^*$). Finally, for initial segments $\succ_j^*$ and $\succ_k^*$, let $\succ_j^* \oplus \succ_k^*$ denote the initial segment that appends $\succ_k^* |_{X \setminus U(x^*, \succ_j^*)}$ to the end of $\succ_j^*$.

Relying on this notation, consider the following (infinitely repeated) choice experiment:

(i) In stage $i = 0$, draw an initial segment $\succ^*$ according to $\mu^*$ and set $\succ_0^* := \succ^*$.

(ii) In any subsequent stage $i \geq 1$, draw an initial segment $\succ^*$ according to $\mu^*$ and set $\succ_i^* := \succ_{i-1}^* \oplus \succ^*$.

Let $\widehat{\mathcal{P}}$ denote the set of permutations on $\{1, ..., m\}$. For any $\hat{\pi} \in \widehat{\mathcal{P}}$, let $\hat{\mu}(\hat{\pi})$ denote the probability that the outcome of the preceding experiment is $\succ_{\hat{\pi}}$. Let $\hat{v}$ denote the canonical utility function on $X$. By construction, $(\hat{\mu}, \hat{v})$ defines a random utility rule for $X$ on the state space $\widehat{\mathcal{P}}$. I show that $\widehat{p}_{(\mu,v)}^* = p_{(\hat{\mu},\hat{v})}$. By definition, it follows that

$$p_{(\hat{\mu},\hat{v})}(x_i, A) := \hat{\mu}(\{\hat{\pi} : \hat{\pi} \in \widehat{\mathcal{P}}_{\hat{v}}(x_i, A)\}) = \Sigma_{\hat{\pi} \in \widehat{\mathcal{P}}_{\hat{v}}(x_i,A)} \hat{\mu}(\hat{\pi}) = \Sigma_{\pi \in \mathcal{P}_v(x_i, A^*)} \mu(\pi) + \left( \Sigma_{\pi \in \mathcal{P}_v(x_{m+1}, A^*)} \mu(\pi) \right) \cdot p_{(\hat{\mu},\hat{v})}(x_i, A).$$

Simplifying the preceding expression gives

$$p_{(\hat{\mu},\hat{v})}(x_i, A) = \frac{\Sigma_{\pi \in \mathcal{P}_v(x_i, A^*)} \mu(\pi)}{1 - \Sigma_{\pi \in \mathcal{P}_v(x_{m+1}, A^*)} \mu(\pi)} = \frac{\mu(\{\pi : \pi \in \mathcal{P}_v(x_i, A^*)\})}{1 - \mu(\{\pi : \pi \in \mathcal{P}_v(x_{m+1}, A^*)\})} = \frac{p_{(\mu,v)}^*(x_i, A^*)}{1 - p_{(\mu,v)}^*(x_{m+1}, A^*)} = \widehat{p}_{(\mu,v)}^*(x_i, A).$$

Since $x_i$ and $A$ were chosen arbitrarily, this establishes the desired result. ■

# 8 Online Appendix

**(a) Manzini and Mariotti's axioms:** For the reader's convenience, I restate their axioms in terms of my notation:

**I-Asymmetry**[19] For all non standard menus $A^* \in \mathcal{X}^*$ and distinct alternatives $a, b \notin A$:

$$I_a(A^*; b) \neq 1 \implies I_b(A^*; a) = 1.$$

**I-Independence** For all non standard menus $A^*, B^* \in \mathcal{X}^*$ and distinct alternatives $a, b \notin A \cup B$:

$$I_a(A^*; b) = I_a(B^*; b) \quad \text{and} \quad I_{x^*}(A^*; b) = I_{x^*}(B^*; b).$$

**(b) Independence of the axioms in Theorem 1:** Table 1 establishes the independence of the axioms in this result.

The choice function $p_1$ violates Strict Regularity because $p_1(b, \{a, b, d\}) = p_1(b, \{a, b, c, d\}) = 6/25$. In addition:

$$p_1(a, \{a, d\}) = p_1(a, \{a, b, d\}) = p_1(a, \{a, b, c, d\}) = \frac{2}{5} \quad \text{and} \quad p_1(b, \{b, d\}) = p_1(b, \{b, c, d\}) = \frac{2}{5}.$$

However, it is straightforward to check that $p_1$ satisfies the other three axioms of Theorem 1.

The choice function $p_2$ violates R-Proportionality because

$$\frac{L_{ad}(\{a, d\}) - L_{ad}(\{a, c, d)}{L_{ad}(\{a, d\}) - L_{ad}(\{a, b, d\})} = 1 \neq \frac{2}{3} = \frac{L_{ba}(\{a, b\})}{L_{ca}(\{a, c\})}.$$

However, it is straightforward to check that $p_2$ satisfies the other three axioms of Theorem 1.

The choice function $p_3$ violates R-Independence because

$$\frac{L_{ad}(\{a, d\})}{L_{ad}(\{a, b, d\})} = \frac{5}{12} \neq \frac{1}{2} = \frac{L_{ad}(\{a, c, d)\}}{L_{ad}(\{a, b, c, d\})} \quad \text{and} \quad \frac{L_{ad}(\{a, d\})}{L_{ad}(\{a, c, d\})} = \frac{5}{12} \neq \frac{1}{2} = \frac{L_{ad}(\{a, b, d\})}{L_{ad}(\{a, b, c, d\})}.$$

However, it is straightforward to check that $p_3$ satisfies the other three axioms of Theorem 1. (Indeed, it is straightforward to construct a similar example starting from any standard choice function that is independent RCM. To do so, simply re-adjust the odds $L_{ad}(\{a, d\})$ in a way that preserves Strict Regularity.)

**(b) Illustration of Proposition 2:** Table 2 illustrates the two models covered by this result.

The choice function with default $q^*$ in the left panel satisfies the axioms of Theorem 1 but violates R-Regularity:

$$L_{bx^*}(\{a, b, c, x^*\}) = \frac{2}{3} \neq 1 = L_{bx^*}(\{b, c, x^*\}) \quad \text{and} \quad q^*(b, \{a, b, c, x^*\}) = \frac{6}{28} \neq \frac{3}{8} = q^*(b, \{b, c, x^*\}).$$

From Proposition 1, it follows that $q^*$ is not independent RCM.

The standard choice function $p$ in the center panel is independent RCM with consideration-preference pair $(\gamma, \succ)$ defined by $\gamma(a) = \gamma(b) = \gamma(c) = \frac{1}{3}$ and $a \succ b \succ c$. What is more, $p$ is also the standardization of $q^*$.

The choice function with default $p^*_\gamma$ in the right panel (defined in the proof of Proposition 2) is the unique independent RCM whose standardization is $p$. It adjusts the weights of $q^*(x^*, \cdot)$ to make $q^*$ consistent with R-Regularity.

[19]Given I-Independence, this is equivalent to $I_a(A^*; b) \neq 1 \iff I_b(A^*; a) = 1$ (see Manzini and Mariotti, p. 1159).

**Table 1: Independence of the axioms in Theorem 1**

| $p_1(j,i)$ | $a$ | $b$ | $c$ | $d$ |
|---|---|---|---|---|
| $\{a\}$ | 1 | – | – | – |
| $\{b\}$ | – | 1 | – | – |
| $\{c\}$ | – | – | 1 | – |
| $\{d\}$ | – | – | – | 1 |
| $\{a,b\}$ | 5/8 | 3/8 | – | – |
| $\{a,c\}$ | 1/2 | – | 1/2 | – |
| $\{a,d\}$ | 2/5 | – | – | 3/5 |
| $\{b,c\}$ | – | 1/2 | 1/2 | – |
| $\{b,d\}$ | – | 2/5 | – | 3/5 |
| $\{c,d\}$ | – | – | 2/3 | 1/3 |
| $\{a,b,c\}$ | 5/11 | 3/11 | 3/11 | – |
| $\{a,b,d\}$ | 10/25 | 6/25 | – | 9/25 |
| $\{a,c,d\}$ | 2/5 | – | 2/5 | 1/5 |
| $\{b,c,d\}$ | – | 2/5 | 2/5 | 1/5 |
| $\{a,b,c,d\}$ | 10/25 | 6/25 | 6/25 | 3/25 |

| $p_2(j,i)$ | $a$ | $b$ | $c$ | $d$ |
|---|---|---|---|---|
| $\{a\}$ | 1 | – | – | – |
| $\{b\}$ | – | 1 | – | – |
| $\{c\}$ | – | – | 1 | – |
| $\{d\}$ | – | – | – | 1 |
| $\{a,b\}$ | 3/5 | 2/5 | – | – |
| $\{a,c\}$ | 1/2 | – | 1/2 | – |
| $\{a,d\}$ | 1/2 | – | – | 1/2 |
| $\{b,c\}$ | – | 1/2 | 1/2 | – |
| $\{b,d\}$ | – | 1/2 | – | 1/2 |
| $\{c,d\}$ | – | – | 3/5 | 2/5 |
| $\{a,b,c\}$ | 3/7 | 2/7 | 2/7 | – |
| $\{a,b,d\}$ | 3/7 | 2/7 | – | 2/7 |
| $\{a,c,d\}$ | 3/8 | – | 3/8 | 2/8 |
| $\{b,c,d\}$ | – | 3/8 | 3/8 | 2/8 |
| $\{a,b,c,d\}$ | 9/25 | 6/25 | 6/25 | 4/25 |

| $p_3(j,i)$ | $a$ | $b$ | $c$ | $d$ |
|---|---|---|---|---|
| $\{a\}$ | 1 | – | – | – |
| $\{b\}$ | – | 1 | – | – |
| $\{c\}$ | – | – | 1 | – |
| $\{d\}$ | – | – | – | 1 |
| $\{a,b\}$ | 3/4 | 1/4 | – | – |
| $\{a,c\}$ | 3/4 | – | 1/4 | – |
| $\{a,d\}$ | 5/7 | – | – | 2/7 |
| $\{b,c\}$ | – | 2/3 | 1/3 | – |
| $\{b,d\}$ | – | 2/3 | – | 1/3 |
| $\{c,d\}$ | – | – | 2/3 | 1/3 |
| $\{a,b,c\}$ | 6/9 | 2/9 | 1/9 | – |
| $\{a,b,d\}$ | 6/9 | 2/9 | – | 1/9 |
| $\{a,c,d\}$ | 6/9 | – | 2/9 | 1/9 |
| $\{b,c,d\}$ | – | 4/7 | 2/7 | 1/7 |
| $\{a,b,c,d\}$ | 12/19 | 4/19 | 2/19 | 1/19 |

**Table 2: Choice functions related as in Proposition 2**

| $q^*(j,i)$ | $a$ | $b$ | $c$ | $x^*$ |
|---|---|---|---|---|
| $\{a, x^*\}$ | $1/2$ | – | – | $1/2$ |
| $\{b, x^*\}$ | – | $1/2$ | – | $1/2$ |
| $\{c, x^*\}$ | – | – | $1/2$ | $1/2$ |
| $\{a, b, x^*\}$ | $3/8$ | $2/8$ | – | $3/8$ |
| $\{a, c, x^*\}$ | $3/8$ | – | $2/8$ | $3/8$ |
| $\{b, c, x^*\}$ | – | $3/8$ | $2/8$ | $3/8$ |
| $\{a, b, c, x^*\}$ | $9/28$ | $6/28$ | $4/28$ | $9/28$ |

$\xrightarrow[\text{standardization}]{}$

| $p(j,i)$ | $a$ | $b$ | $c$ |
|---|---|---|---|
| $\{a\}$ | 1 | – | – |
| $\{b\}$ | – | 1 | – |
| $\{c\}$ | – | – | 1 |
| $\{a, b\}$ | $3/5$ | $2/5$ | – |
| $\{a, c\}$ | $3/5$ | – | $2/5$ |
| $\{b, c\}$ | – | $3/5$ | $2/5$ |
| $\{a, b, c\}$ | $9/19$ | $6/19$ | $4/19$ |

$\underset{\text{Proposition 2}}{\Longleftrightarrow}$

$\xleftarrow[\text{standardization}]{}$

| $p^*_\gamma(j,i)$ | $a$ | $b$ | $c$ | $x^*$ |
|---|---|---|---|---|
| $\{a, x^*\}$ | $1/3$ | – | – | $2/3$ |
| $\{b, x^*\}$ | – | $1/3$ | – | $2/3$ |
| $\{c, x^*\}$ | – | – | $1/3$ | $2/3$ |
| $\{a, b, x^*\}$ | $3/9$ | $2/9$ | – | $4/9$ |
| $\{a, c, x^*\}$ | $3/9$ | – | $2/9$ | $4/9$ |
| $\{b, c, x^*\}$ | – | $3/9$ | $2/9$ | $4/9$ |
| $\{a, b, c, x^*\}$ | $9/27$ | $6/27$ | $4/27$ | $8/27$ |