University of St Andrews

School of Economics and Finance Discussion Papers

# Dual Random Utility Maximisation

Paola Manzini and Marco Mariotti

School of Economics and Finance
Online Discussion Paper Series
issn 2055-303X
http://ideas.repec.org/s/san/wpecon.html
info: econ@st-andrews.ac.uk

School of Economics and Finance Discussion Paper No. 1605
23 Mar 2016 (revised 7 Nov 2016)
JEL Classification: D03
Keywords: Stochastic Choice, Attraction Effect, Similarity Effect

# Dual Random Utility Maximisation*

Paola Manzini[†] Marco Mariotti[‡]

This version: November 2016

**Abstract**

Dual Random Utility Maximisation (dRUM) is Random Utility Maximisation when utility depends on only two states. dRUM has many relevant behavioural interpretations and practical applications. We show that it is (generically) the only stochastic choice rule that satisfies Regularity and two new simple properties: Constant Expansion (if the choice probability of an alternative is the same across two menus, then it is the same in the combined menu), and Negative Expansion (if the choice probability of an alternative is less than one and differs across two menus, then it vanishes in the combined menu). By extending the theory to menu-dependent state probabilities we are able to accommodate prominent violations of Regularity such as the attraction, similarity and compromise effects.

**J.E.L. codes:** D03, D01

**Keywords:** Stochastic Choice, Attraction Effect, Similarity Effect.

*We are grateful to Sophie Bade, Mark Dean, Kfir Eliaz, Sean Horan, Jay Lu and Yusufcan Masatlioglu for insightful suggestions and advice, as well as to the audiences at BRIC 4 (Northwestern), the GSE Stochastic Choice workshop (Barcelona) and Royal Holloway University of London. Any errors are our own.

[†]School of Economics and Finance, University of St. Andrews, Castlecliffe, The Scores, St. Andrews KY16 9AL, Scotland, UK.

[‡]School of Economics and Finance, Queen Mary University of London, Mile End Road, London E1 4NS, UK.

# 1 Introduction

In Random Utility Maximisation (RUM) choices are determined by the maximisation of a utility that depends on random states. The state probabilities are independent of the menu faced by the decision maker. In this paper we develop the theory of RUM for the case in which the states are exactly two (dual RUM or dRUM in short) and the number of alternatives is finite. We characterise the behaviour of a dual random utility maximiser with a set of simple conditions. We also extend the theory to state probabilities that depend on the menu, a feature that -we argue- is compelling in many scenarios.

Dual RUM, while appearing peculiar, has in fact many interesting applications, both when stochastic choices are observed for an individual agent and when they describe instead random draws from a population. Here are a few examples:

- *Hidden conditions of choice.* The states correspond to on-off conditions that are relevant for the decision maker and are hidden to the observer, such as time pressure.[1] For instance, a store manager may wonder whether the variability in buying patterns fits a 'dual' form behaviour due to the presence or absence of time pressure at the time of purchase. Scanner data on the repeated purchases of each consumer are likely to be available to the manager, as well as which products were available at the time of each purchase, so that a stochastic choice function can be constructed. It is, however, unlikely that the manager can learn whether or not the consumer was in a hurry at the time of purchase. Our results allow a direct test of the dual behaviour hypothesis.

- *Dual-self processes.* An individual may be either in a 'cool' state, in which a long-run utility is maximised, or in a 'hot' state, in which a myopic self, subject to short-run impulses (such as temptation), takes control. Indeed, the menu-independent version of dRUM appears in the contracting model by Eliaz and Spiegler [11], and the menu-dependent version describes the implicit second-stage choices of the dual-self

[1]The effect of time pressure on decision making has recently attracted attention. See e.g. Reutskaya et al. [32] for a recent study of this effect on search behaviour in a shoppping environment.

model characterised by Chatterjee and Krishna [5] using preferences *over* menus. Correspondingly, our results can be seen as a direct characterisation of this model in terms of choice *from* menus.[2]

- *Hidden population heterogeneity.* dRUM may represent the variability of choice due to a binary form of hidden population heterogeneity. The states correspond to two types of individuals in the population. Interesting examples of this kind of binary heterogeneity are when the population is split, in unknown proportions, into high and low cognitive ability, strategic and non-strategic, 'fast and slow' or 'instinctive and contemplative' types (e.g. Caplin and Martin [4]; Gill and Prowse [17]; Proto, Rustichini and Sofianos [31]; Rubinstein [33],[34]). Such binary classifications are natural and widespread.[3]

- *Household decisions.* An important application to a multi-person situation concerns decision units composed of two agents, i.e. a household, for which the observer is uncertain about the exact identity of the decision maker. Household purchases can be observed through standard consumption data, but typically it is not known which of the two partners made the actual purchase decision on any given occasion. Dual RUM constitutes the basis for a 'random dictatorship' model of household decisions that could complement the 'collective' model (Chiappori [7]; Cherchye, De Rock and Vermeulen [6]).[4]

- *Normativity vs selfishness.* Many situations of choice present a conflict between a 'normative' mode and a 'selfish' mode of decision. This conflict is both introspectively clear and consistent with experimental observations, for example in dictator

[2] Eliaz and Spiegler [11] do not focus on characterisation but rather explore the rich implications of this probabilistic notion of naivete in a contract-theoretic framework, where a principal chooses the optimal menu of contracs to offer to partially naive agents.

[3] Caplin and Martin [4] are distinctive in that they consider agents who can *choose* whether to make automatic (fast) or considered (slow) choices, depending on an attentional cost.

[4] de Clippel and Eliaz [8] propose a dual-self model that can be thought of as a model of intra-household bargaining. Interestingly, this model exhibits the attraction and compromise effects that we study later in the paper.

games (Frolich, Oppenheimer and Moore [16]). This dichotomy can pertain both to a typology of individuals and to a single conflicted individual.

The models we study, in which only two alternatives receive positive probability, are of course a theoretical idealisation. In practice, an empirical distribution roughly conforming to the theory will be strongly *bimodal*, expressing broadly polarised preferences in society or within an individual. Evidence of bimodal distributions in individual and collective choice is found in disparate contexts.[5]

We now discuss the key analytical facts in the 'base case' of the theory. Suppose that choices are compatible with dRUM, but let us exclude for simplicity the special case in which the two states have exactly the same probabilities, which presents some peculiarities. Suppose you observe that $a$ is chosen with probability $\alpha$ both when the menu is $A$ and when the menu is $B$. If $0 < \alpha < 1$ then, while you do not know in which state the choices were made, you do know that $a$ has the highest utility in the *same* state, say $s$, in $A$ and $B$, and not in the other state $s'$. So, when the menu is $A \cup B$, $a$ continues to have the highest utility in state $s$, while in $s'$ some alternative in $A$ or $B$ has a higher utility than $a$. Therefore $a$ must be chosen with the same probability $\alpha$ also from $A \cup B$. A similar logic applies when $\alpha = 0$ or $\alpha = 1$. This *Constant Expansion* property is a first necessary property of dRUM.

Suppose instead that $a$ is chosen with different probabilities from $A$ and from $B$, and that these probabilities are less than one. Then it must be the case that $a$ has the highest utility in different states in $A$ and $B$ (this would not be the case if $a$ was chosen with probability one from $A$ or $B$, for then $a$ would have the highest utility in both states in at least one of the two menus). As a consequence, in each state there is some alternative that has a higher utility than $a$ in $A \cup B$, and therefore $a$ must be chosen with probability zero

[5] E.g. Frolich, Oppenheimer and Moore [16] and Dufwenberg and Muren [9] (choices in a dictator games concentrated on giving nothing or 50/50), Sura, Shmuelib, Bosec and Dubeyc [35] (bimodal distributions in ratings, such as Amazon), Plerou, Gopikrishnan, and Stanley [30] (phase transition to bimodal demand "bulls and bears"- in financial markets), Engelmann and Normann [13] (bimodality on maximum and minimum effort levels in minimum effort games), McClelland, Schulze and Coursey [26] (bimodal beliefs for unlikely events and willingness to insure).

from $A \cup B$. This *Negative Expansion property* is a second necessary property of dRUM.

Theorem 1 says that, in the presence of the standard axiom of Regularity (adding new alternatives to a menu cannot increase the choice probability of an existing alternative), Constant and Negative Expansion are in fact *all* the behavioural implications of dRUM: the three axioms completely characterise dRUM when state probabilities are asymmetric. Observe that these axioms give us significantly different information from the representation: for example, they do not include any reference to the 'dual' aspect of the choice process.

The case that also allows for symmetric state probabilities is more complex and requires a separate treatment. The Constant Expansion axiom no longer holds, since now it could happen that $a$ is chosen in different states from $A$ and from $B$ and so it is not chosen from the union. Dropping this axiom, a *Contraction Consistency* property, which is implied by the other axioms in the generic case, must be assumed explicitly. It says that the *impact* of an alternative $a$ on another alternative $b$ (removing $a$ affects the choice probability of $b$) is preserved when moving from larger to smaller menus. Theorem 2 shows that Regularity, Negative Expansion and Contraction Consistency completely characterise general dRUMs.

The identification properties of the model are very different in the two cases. In the asymmetric case preferences are uniquely identified. Otherwise, uniqueness holds only up to a certain class of ordinal transformation of the rankings.

The symmetric state case is nongeneric, but its interest is not only technical since it serves as a tool to analyse the extension of the theory that we now describe.

While dRUM is a restriction of the classical RUM, in the final part of the paper we explore a type of random utility maximisation that explains behaviours outside the RUM family, by allowing the state probabilities to depend on the menu. As noted, this is an extremely natural feature in many contexts. This model is also interesting because it can - unlike dRUM - represent as (extended) random utility maximisation some prominent violations of Regularity, such as the attraction, similarity and compromise effects. At first sight, allowing total freedom in the way probabilities can depend on menus seems to spell a total lack of structure, but we show in Theorem 3 that in fact the menu-dependent probability model, too, is neatly disciplined by consistency properties, which mirror those characterising the fixed-probability model. The feature of these properties is that they are 'modal', in the sense that they make assertions on the relationship between the *certainty, impossibility and possibility*, rather than the numerical probability, of selection across menus.

The behaviour corresponding to a general RUM is not very well-understood: RUM is equivalent to the satisfaction of a fairly complex set of conditions that illuminate the representation rather than providing a behavioural interpretation (Block and Marschak [3]; Falmagne [14]; Barbera and Pattanaik [2]).[6] For this reason other scholars have focussed on interesting special cases where additional structure is added, leading to transparent behavioural characterisations. Beside the classical Luce [23] model and its modern variations (e.g. Gul, Natenzon and Pesendorfer [19]; Echenique and Saito [10]), among the recent contributions we recall in particular Gul and Pesendorfer [18], who assume alternatives to be lotteries and preferences to be von Neumann Morgenstern; Apesteguia, Ballester and Lu [1], who examine, in an abstract context, a family of utility functions that satisfies a single-crossing condition; and Lu and Saito [22], who characterise RUM in a intertemporal choice context. The two state case of RUM of this paper is another restriction that is useful from this perspective.

## 2 Preliminaries

Let $X$ be a finite set of $n \geq 2$ alternatives. The nonempty subsets of $X$ are called *menus*. Let $\mathcal{D}$ be the set of all nonempty menus.

A *stochastic choice rule* is a map $p : X \times \mathcal{D} \rightarrow [0,1]$ such that: $\sum_{a \in A} p(a, A) = 1$ for all

[6]To elaborate, the Block-Marschak-Falmagne (BMF) conditions require $\sum_{B: A \subseteq B \subseteq X} (-1)^{|B \setminus A|} p(a, B) \geq 0$ for all menus $A$ and alternative $a \in A$, where $p(a, B)$ is the probability of choosing $a$ from menu $B$. These conditions can be easily interpreted only in terms of the representation itself, i.e. as indicating certain features of the probabilities of the rankings. As noted by Fiorini [15], they come from a re-statement of the RUM model as $p(a, A) = \sum_{B: A \subseteq B \subseteq X} q(a, B)$, where $q$ is the probability that $a$ is ranked top in $B$ and not in any of the supersets of $B$. Solving this equation for $q$ in terms of $p$ via Moebius inversion yields the terms of the BMF conditions, which then simply assert the non-negativity of the ranking probabilities $q$.

$A \in \mathcal{D}$; $p(a, A) = 0$ for all $a \notin A$; and $p(a, A) \in [0, 1]$ for all $a \in A$, for all $A \in \mathcal{D}$. The value $p(a, A)$ may be interpreted either as the probability with which an individual agent (on whose behaviour we focus) chooses $a$ from the menu $A$, or as the fraction of a population choosing $a$ from $A$.

In Random Utility Maximisation (RUM) a stochastic choice rule is built by assuming that there is probabilistic state space and a state dependent utility which is maximised in the realised state. Provided that the event that two alternatives have the same realised utility has probability zero, RUM can be conveniently described in terms of random rankings (see e.g. Block and Marschak [3]), which we henceforth do. A *ranking* of $X$ is a bijection $r : \{1, ..., n\} \to X$. Let $\mathcal{R}$ be the set of all rankings. We interpret the image of $r \in \mathcal{R}$ as describing alternatives in decreasing order of preference: $r(1)$ is the most preferred alternative, $r(2)$ the second most preferred, and so on.

Let $\mathcal{R}(a, A)$ denote the set of rankings for which $a$ is the top alternative in $A$, that is,

$$\mathcal{R}(a, A) = \left\{ r \in \mathcal{R} : r^{-1}(b) < r^{-1}(a) \Rightarrow b \notin A \right\}.$$

Let $\mu$ be a probability distribution on $\mathcal{R}$. A *RUM* is a stochastic choice rule $p$ such that

$$p(a, A) = \sum_{r \in \mathcal{R}(a,A)} \mu(r)$$

A *dual RUM (dRUM)* is a RUM that uses only two rankings, that is one for which the following condition holds: $\mu(r)\mu(r') > 0, r \neq r' \Rightarrow \mu(r'') = 0$ for all rankings $r'' \neq r, r'$. Letting $\mu(r_1) = \alpha$, a dRUM $p$ is thus identified by a triple $(r_1, r_2, \alpha)$ of two rankings and a number $\alpha \in [0, 1]$. We say that $(r_1, r_2, \alpha)$ *generates* $p$.

To ease notation, from now on given rankings $r_i$, $i = 1, 2$, we shall write $a \succ_i b$ instead of $r_i^{-1}(a) < r_i^{-1}(b)$.

A dRUM is a RUM and thus satisfies, for all menus $A, B$:

**Regularity:** If $A \subset B$ then $p(a, A) \geq p(a, B)$.

The two key additional properties are the following, for all menus $A, B$:

**Constant Expansion**: If $p(a, A) = p(a, B) = \alpha$ then $p(a, A \cup B) = \alpha$.

**Negative Expansion**: If $p(a, A) < p(a, B) < 1$ then $p(a, A \cup B) = 0$.

Constant Expansion extends to stochastic choice the classical expansion property of rational deterministic choice: if an alternative is chosen (resp., rejected) from two menus, then it is chosen (resp., rejected) from the union of the menus. In the stochastic setting the axiom remains a menu-independence condition. For example, in the population interpretation, the same frequencies of choice for $a$ in two menus reveal that $a$ got support from the same group in both menus, and not from the other group; or from both groups in both menus; or from neither group in both menus. If the reasons for support are menu-independent they persist when the menus are merged, and the conclusion in the axiom must follow.

Negative Expansion is instead peculiar to stochastic choice since the antecedent of the axiom is impossible in deterministic choice. It is a 'dual' form of menu-independence condition. For example, different frequencies of choice for $a$ in two menus reveal (given that they are less than unity) that in one menu $a$ was not supported by one group, and in the other menu $a$ was not supported by the other group. Therefore, if the reasons for support are menu-independent, no group supports $a$ when the menus are merged.

# 3 Asymmetric dRUMs

The case in which all non-degenerate choice probabilities are exactly equal to $\frac{1}{2}$ requires a separate method of proof and a different axiomatisation from all other cases $\alpha \in [0, 1]$. Thus we begin our analysis by excluding for the moment this special case. The characterisation of the general case will be based on the results in this section.

A stochastic choice rule is *symmetric* if $p(a, A) \in (0, 1)$ for some $A$ and $a \in A$ and $p(a, A) = \frac{1}{2}$ for all $A$ and $a \in A$ for which this is the case. A stochastic choice rule is *asymmetric* if it is not symmetric.

**Remark 1** Asymmetric dRUMs, but not symmetric ones, satisfy Weak Stochastic Transitivity (WST), defined as follows:[7]

$$p(a,\{a,b\}) \geq \frac{1}{2}, p(b,\{b,c\}) \geq \frac{1}{2} \Rightarrow \ p(a,\{a,c\}) \geq \frac{1}{2}$$

On the other hand, consider the symmetric dRUM generated by $c \succ_1 a \succ_1 b$, $b \succ_2 c \succ_2 a$ and $\alpha = \frac{1}{2}$. Then $p(a,\{a,b\}) = \frac{1}{2} = p(b,\{b,c\})$ but $p(a,\{a,c\}) = 0$, violating WST. Asymmetric dRUMs however do satisfy the following slightly weaker version of WST: $p(a,\{a,b\}) \geq \frac{1}{2}$, $p(b,\{b,c\}) \geq \frac{1}{2}$, $p(a,\{a,b\})\, p(b,\{b,c\}) > \frac{1}{4} \Rightarrow p(a,\{a,c\}) \geq \frac{1}{2}$. General RUMs can violate both properties.

Our main result in this section is the following characterisation:

**Theorem 1** *An asymmetric stochastic choice rule $p$ is a dRUM if and only if it satisfies Regularity, Constant Expansion and Negative Expansion.*

In order to prove the theorem we establish two preliminary results. The first lemma shows that in any menu there are at most two alternatives receiving positive choice probability:

**Lemma 1** *Let $p$ be a stochastic choice rule that satisfies Regularity, Constant Expansion and Negative Expansion. Then for any menu $A$, if $p(a,A) \in (0,1)$ for some $a \in A$ there exists $b \in A$ for which $p(b,A) = 1 - p(a,A)$.*

**Proof:** Suppose by contradiction that for some menu $A$ there exist $b_1, ..., b_n \in A$ such that $n > 2$ and $p(b_i, A) > 0$ for all $i$. Since $n > 2$, we have $p(b_i, A) + p(b_j, A) < 1$ for all distinct $i, j$, and therefore $p(b_i, \{b_i, b_j\}) > p(b_i, A)$ for some $i, j$. Fixing such a pair $(i, j)$, if it were $p(b_i, \{b_i, b_j\}) = p(b_i, \{b_i, b_k\})$ for all $k \neq i, j$, then by Constant Expansion $p(b_i, \{b_i, b_j\}) = p(b_i, \{b_1, ..., b_n\}) = p(b_i, A)$ (the last equality holding by Regularity), a contradiction. Therefore $p(b_i, \{b_i, b_j\}) \neq p(b_i, \{b_i, b_k\})$ for some $k \neq i, j$. Clearly $p(b_i, \{b_i, b_j\}) < 1$ and $p(b_i, \{b_i, b_k\}) < 1$ (otherwise by Regularity we have the contradiction $p(b_i, A) = 1$). But then by Negative Expansion $p(b_i, \{b_i, b_j, b_k\}) = 0$, contradicting Regularity and $p(b_i, A) > 0$. $\square$

[7] To see this, let w.l.o.g. $\alpha > \frac{1}{2}$ and suppose $p(a,\{a,b\}) \geq \frac{1}{2}$ and $p(b,\{b,c\}) \geq \frac{1}{2}$. Therefore $a \succ_1 b$ and $b \succ_1 c$, so that $a \succ_1 c$ and then $p(a,\{a,c\}) > \frac{1}{2}$.

The next lemma shows that the two alternatives receiving positive probability do not change across menus.

**Lemma 2** *Let $p$ be a stochastic choice rule that satisfies Regularity, Constant Expansion and Negative Expansion. Then there exists $\alpha \in [0,1]$ such that, for any menu $A$ and all $a \in A$, $p(a, A) \in \{0, \alpha, 1-\alpha, 1\}$.*

**Proof:** If for all $A$ we have $p(a, A) = 1$ for some $a$ there is nothing to prove. Then take $A$ for which $p(a, A) = \alpha \in (0,1)$ for some $a$. By Lemma 1 there exists exactly one alternative $b \in A$ for which $p(b, A) = 1-\alpha$. Note that by Regularity $p(a, \{a, c\}) \geq \alpha$ for all $c \in A \setminus \{a\}$ and $p(b, \{b, c\}) \geq 1-\alpha$ for all $c \in A \setminus \{b\}$.

Now suppose by contradiction that there exists a menu $B$ and a $c \in B$ for which $p(c, B) = \beta \notin \{0, \alpha, 1-\alpha, 1\}$. By Lemma 1 there exists exactly one alternative $d \in B$ for which $p(d, B) = 1-\beta$.

Consider the menu $A \cup B$. By Regularity $p(e, A \cup B) = 0$ for all $e \in (A \cup B) \setminus \{a, b, c, d\}$. Moreover, by Lemma 1 $p(e, A \cup B) = 0$ for some $e \in \{a, b, c, d\}$, and w.l.o.g., let $p(a, A \cup B) = 0$.

If $p(b, A \cup B) > 0$ then by Regularity and Negative Expansion $p(b, A \cup B) = 1-\alpha$ (if it were $p(b, A \cup B) < 1-\alpha$, then $p(b, A \cup B) \neq p(b, A)$ with $p(b, A \cup B), p(b, A) < 1$, so that Negative Expansion would imply the contradiction $p(b, A \cup B) = 0$). It follows from Lemma 1 that either $p(c, A \cup B) = \alpha$ or $p(d, A \cup B) = \alpha$. In the former case, since $\alpha > 0$ a reasoning analogous to the one followed so far implies that $p(c, A \cup B) = p(c, B) = \beta$, contradicting $\alpha \neq \beta$; and in the latter case we obtain the contradiction $p(d, A \cup B) = p(d, B) = 1-\beta$.

The proof is concluded by arriving at analogous contradictions starting from $p(c, A \cup B) > 0$ or $p(d, A \cup B) > 0$. $\square$

**Proof of Theorem 1:** Necessity, described informally in the introduction, is easily checked. For sufficiency, we proceed by induction on the cardinality of $X$. A preliminary observation is that if $p$ satisfies the axioms on $X$, then its restriction to $X \setminus \{a\}$ also satisfies the axioms on $X \setminus \{a\}$.[8]

[8] If $A \subset B \subseteq X \setminus \{a\}$ then by Regularity (defined on the entire $X$) $p(b, A) \geq p(b, B)$. If $A, B \subseteq X \setminus \{a\}$

Moving to the inductive argument, if $|X| = 2$ the result is easily shown. For $|X| = n > 2$, we consider two cases.

<u>CASE A</u>: $p(a, X) = 1$ for some $a \in X$.

Denote $p'$ the restriction of $p$ to $X \setminus \{a\}$. By the preliminary observation and the inductive hypothesis there exist two rankings $r_1'$ and $r_2'$ on $X \setminus \{a\}$ and $\alpha \in [0,1]$ such that $(r_1', r_2', \alpha)$ generates $p'$ on $X \setminus \{a\}$. Extend $r_1'$ and $r_2'$ to $r_1$ and $r_2$ on $X$ by letting $r_1(1) = r_2(1) = a$ and letting $r_i$ agree with $r_i'$, $i = 1, 2$, for all the other alternatives, that is $c \succ_i d \Leftrightarrow c \succ_i' d$ for all $c, d \neq a$ (where for all $x, y$ we write $x \succ_i' y$ to denote $(r_i')^{-1}(x) < (r_i')^{-1}(y)$). Let $\tilde{p}$ be the dRUM generated by $(r_1, r_2, \alpha)$. Take any menu $A$ such that $a \in A$. Then $\tilde{p}(a, A) = 1$. Since by Regularity $p(a, A) = 1$, we have $p = \tilde{p}$ as desired.

<u>CASE B</u>: $p(a, X) = \alpha \neq \frac{1}{2}$ and $p(b, X) = 1 - \alpha$ for distinct $a, b \in X$ and $\alpha \in (0,1)$.

Note that by Lemma 1 and 2 this is the only remaining possibility (in particular, if it were $p(a, X) = \frac{1}{2}$ then $p(c, A) \in \left\{0, \frac{1}{2}, 1\right\}$ for all menus $A$ and $c \in A$, in violation of $p$ being asymmetric). Assume w.l.o.g. that $\alpha > \frac{1}{2}$. As in Case A, by the inductive step there are two rankings $r_1'$ and $r_2'$ on $X \setminus \{a\}$ and $\beta \in [0,1]$ such that $(r_1', r_2', \beta)$ generates the restriction $p'$ of $p$ to $X \setminus \{a\}$. If for all $A \subseteq X \setminus \{a\}$ there is some $c_A \in A$ such that $p(c_A, A) = 1$, then the two rankings must agree on $X \setminus \{a\}$. In this case, for any $\gamma \in [0,1]$, $(r_1', r_2', \gamma)$ also generates $p'$ on $X \setminus \{a\}$, and in particular we are free to choose $\gamma = \alpha$. If instead there exists $A \subseteq X \setminus \{a\}$ for which $p(c, A) \neq 1$ for all $c \in A$, we know from Lemmas 1 and 2 that then there exist exactly two alternatives $c, d \in A$ such that $p(d, A) = \alpha$ and $p(c, A) = 1 - \alpha$. Therefore in this case it must be $\beta = \alpha$ (otherwise $(r_1', r_2', \beta)$ would not generate the choice probabilities $p(d, A) = \alpha$ and $p(c, A) = 1 - \alpha$). Note that we must have $r_2'(1) = b$ since by Regularity and Lemma 2 either $p(b, X \setminus \{a\}) = 1$ (in which case also $r_1'(1) = b$) or $p(b, X \setminus \{a\}) = 1 - \alpha$.

Now extend $r_1'$ to $r_1$ on $X$ by setting $r_1(1) = a$ and letting the rest of $r_1$ agree with $r_1'$

---

and $p(b, A) = p(b, B) = \alpha$ then $A \cup B \subseteq X \setminus \{a\}$ and by Constant Expansion (on $X$) $p(b, A \cup B) = \alpha$. If $A, B \subseteq X \setminus \{a\}$, $p(b, A) \neq p(b, B)$ and $p(b, A), p(b, B) < 1$ then $A \cup B \subseteq X \setminus \{a\}$ and by Negative Expansion (on $X$) $p(b, A \cup B) = 0$.

on $X \backslash \{a\}$; and extend $r_2'$ to $r_2$ on $X$ by setting

$$\begin{aligned} a &\succ {}_2 c \text{ for all } c \text{ such that } p(a, \{a, c\}) = 1 \\ c &\succ {}_2 a \text{ for all } c \text{ such that } p(a, \{a, c\}) = \alpha. \end{aligned}$$

and letting the rest of $r_2$ agree with $r_2'$ on $X \backslash \{a\}$. Note that by Regularity and Lemma 2 the two possibilities in the display are exhaustive (recall the assumption $\alpha > \frac{1}{2}$).

We need to check that this extension is consistent, that is, that $r_2$ does not rank $a$ above some $c$ but below some $d$ that is itself below $c$. Suppose to the contrary that there existed $c$ and $d$ for which $c \succ_2 d$ but $a \succ_2 c$ and $d \succ_2 a$. These three inequalities imply, respectively, that $p(c, \{c, d\}) \geq 1 - \alpha$ (since $r_1'$ and $r_2'$ generate $p'$ on $X \backslash \{a\}$), $p(a, \{a, c\}) = 1$ and $p(a, \{a, d\}) = \alpha$ (by the construction of $r_2$).

If $c \succ_1' d$ then $p(d, \{c, d\}) = 0$, hence $p(d, \{a, c, d\}) = 0$ by Regularity. Since $p(c, \{a, c\}) = 0$ it is also $p(c, \{a, c, d\}) = 0$ by Regularity, and thus $p(a, \{a, c, d\}) = 1$ contradicting Regularity and $p(a, \{a, d\}) = \alpha < 1$.

If instead $c \succ_1' d$ then $p(d, \{c, d\}) = \alpha$. Since $p(d, \{a, d\}) = 1 - \alpha$, by Negative Expansion $p(d, \{a, c, d\}) = 0$ and we can argue as above. This shows that the extension $r_2$ is consistent.

Let $\tilde{p}$ the dRUM generated by $(r_1, r_2, \alpha)$. We now show that $p = \tilde{p}$.

Take any menu $A$ for which $a \in A$. By Regularity (recalling the assumption $\alpha > \frac{1}{2}$) either (1) $p(a, A) = 1$ or (2) $p(a, A) = \alpha$ and $p(c, A) = 1 - \alpha$ for some $c \in A \backslash \{a\}$. In case (1), $p(a, \{a, d\}) = 1$ for all $d \in A \backslash \{a\}$ by Regularity. Thus $a \succ_1 d$ and $a \succ_2 d$ for all $d \in A \backslash \{a\}$, and therefore $\tilde{p}(a, A) = 1 = p(a, A)$ as desired.

Consider then case (2). Since $p(c, A) = 1 - \alpha$ we have $p(c, \{a, c\}) \geq 1 - \alpha$ by Regularity. Then by Regularity again we have $p(a, \{a, c\}) = \alpha$ and therefore $c \succ_2 a$, so that $\tilde{p}(a, A) = \alpha$ as desired. It remains to be checked that $\tilde{p}(c, A) = 1 - \alpha$. This means showing that $c \succ_2 d$ also for all $d \in A \backslash \{a, c\}$, which is equivalent to showing that $c \succ_2' d$ for all $d \in A \backslash \{a\}$. If $p(c, A \backslash \{a\}) = 1$, this is certainly the case since $(r_1', r_2', \alpha)$ generates the restriction of $p$ to $A \backslash \{a\}$. If instead $p(c, A \backslash \{a\}) = 1 - \alpha$ (the only other possibility by Regularity and Lemmas 1 and 2), noting that $\alpha \neq 1 - \alpha$, again this must be the case given that $(r_1', r_2', \alpha)$ generates the restriction of $p$ to $A \backslash \{a\}$ (observe in passing that this conclusion might not hold in the case $\alpha = \frac{1}{2}$), and this concludes that proof. ■

This result is of interest first because, like all 'revealed preference' style axiomatisations, it provides us with a way to test the theory *directly* through the observation of choice data, without the need for any parametric specification or for non-behaviour data.[9] An empirical test using our axioms would of course need to treat statistically the assertions that probabilities of choice are equal or different. Exactly the same need arises in any other characterisation of stochastic choice models: for example, in the case of the classical and popular Luce/logit model [23], which is characterised by a single axiom asserting the equality between two probability ratios.[10]

Aside from testing purposes, the characterisation is also useful because it gives a *complete* yet succinct picture of the theory in terms of assertions on behaviour only (as opposed to the cognitive process that underlies behaviour), in particular about the 'comparative statics' of merging and expansion of menus. For example, if market research shows that an existing model $a$ would turn out to have the same share against a new model $b$ alone that it has against an existing model $c$, then $a$ is neither harmed nor damaged by the introduction of $b$. On the other hand, if $a$ had different shares against $b$ and $c$, then it would be significantly harmed by the introduction of $b$.

We derived from more fundamental properties the fact that in an asymmetric dRUM alternatives can be chosen from any menu with only two fixed non-degenerate choice probabilities $\alpha \neq \frac{1}{2}$ and $1-\alpha$. One may wonder whether a characterisation can be obtained by simply postulating directly that there exists $\alpha \in (0,1)$, $\alpha \neq \frac{1}{2}$, such that for all $A \subseteq X$ and $a \in A$, $p(a, A) \in \{0, \alpha, 1-\alpha, 1\}$. This "range" property would certainly not be very satisfactory as an axiom at the *theoretical* level, since it would be very close to the representation, with the same binary flavour. Nevertheless, it could serve as a *practical* test of the theory since its validity should be easily detectable from the data. As it turns out, the range property and Regularity are anyway not sufficient to ensure that the data are

[9] Needless to say, this latter type of data may be useful when available . But they may also be hard to come by in some situations (consider the examples in the Introduction: e.g. household budget surveys are the typical data used in the empirical analysis of family consumption). In addition, the information they provide on the theory may not always be clear-cut.

[10] See McCausland and Marley [27] and McCausland, Davis-Stober, Marley, Park and Brown [28] for a thorough approach to the Bayesian testing of stochastic choice conditions.

generated by a dRUM: they must still be disciplined by further across-menu consistency properties. Table 1 presents a stochastic choice rule $p$ that satisfies the range property and Regularity (and even Constant Expansion) but is not a dRUM.

| | $\{a,b,c,d\}$ | $\{a,b,c\}$ | $\{a,b,d\}$ | $\{a,c,d\}$ | $\{b,c,d\}$ | $\{a,b\}$ | $\{a,c\}$ | $\{a,d\}$ | $\{b,c\}$ | $\{b,d\}$ | $\{c,d\}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $a$ | 0 | $\frac{2}{3}$ | $\frac{1}{3}$ | 0 | – | $\frac{2}{3}$ | $\frac{2}{3}$ | $\frac{1}{3}$ | – | – | – |
| $b$ | 0 | 0 | 0 | – | 0 | $\frac{1}{2}$ | – | – | $\frac{2}{3}$ | $\frac{1}{3}$ | – |
| $c$ | $\frac{1}{3}$ | $\frac{1}{3}$ | – | $\frac{1}{3}$ | $\frac{1}{3}$ | – | $\frac{1}{3}$ | – | $\frac{1}{3}$ | – | $\frac{1}{3}$ |
| $d$ | $\frac{2}{3}$ | – | $\frac{2}{3}$ | $\frac{2}{3}$ | $\frac{2}{3}$ | – | – | $\frac{2}{3}$ | – | $\frac{2}{3}$ | $\frac{2}{3}$ |

Table 1 A Stochastic Choice Rule that fails Negative Expansion

That $p$ is not a dRUM follows from Theorem 1 and the fact that $p$ fails Negative Expansion (e.g. $p(a, \{a, b\}) = \frac{2}{3} \neq \frac{1}{3} = p(a, \{a, d\})$ while $p(a, \{a, b, d\}) > 0$).

# 4 General dRUMs

The asymmetric case we have dealt with so far is generic. It would perhaps not merit devoting much attention to extending the characterisation to the special symmetric case, were it not for the fact that precisely this case is the key tool to treat the menu-dependence of the state probabilities, which is definitely of conceptual and practical interest.

The easy inductive method we used in the proof of Theorem 1 breaks down for general dRUMs. We have to engage more directly with reconstructing the rankings from choice probabilities. The difficulty here is that, while it is clear that an alternative $a$ being ranked above another alternative $b$ in *some* ranking is revealed by the fact that removing $a$ from a menu increases the choice probability of $b$, it is not immediately evident in *which* ranking $a$ is above $b$ (whereas in the asymmetric case the two rankings are distinguished by the distinct probabilities $\alpha$ and $1 - \alpha$).

If the probabilities of the rankings that generate a dRUM are allowed to take on the value $\frac{1}{2}$, then Constant Expansion is no longer a necessary property. In fact, if $p(a, A) = p(a, B) = \frac{1}{2}$ it could happen that $a$ is top in $A$ according to ranking $r_1$ (but not according to $r_2$) and top in $B$ according to $r_2$ (but not according to $r_1$). Then in $A \cup B$ there will be alternatives that are above $a$ in both rankings, so that $p(a, A \cup B) = 0$.

Towards a characterisation, we note that the axioms in the statement of Theorem 1 imply a strong contraction consistency property that is at the heart of general dRUM. The property uses the concept of *impact* (Manzini and Mariotti [25]): $b$ impacts $a$ in $A$ whenever $p(a, A) > p(a, A \cup \{b\})$. If $b$ impacts $a$ in $A$ and $a \in B \subseteq A$ then $b$ impacts $a$ in $B$. Formally:

**Contraction Consistency**: For all $a \in B \subseteq A$ and $b \in X$: If $p(a, A) > p(a, A \cup \{b\})$ then $p(a, B) > p(a, B \cup \{b\})$.

Contraction Consistency simply says that impact is inherited from menus to submenus. It is evidently a stochastic generalisation of a property that holds for deterministic choice functions that maximise a partial ordering: in that case, if $b$ impacts $a$ (i.e. turns $a$ from chosen to unchosen) in a menu $A$ then (1) $b$ is higher than $a$ in the ordering and (2) $a$ is a maximal alternative in $A$ (otherwise it would not be chosen from $A$ and it could not be impacted by $b$). It thus follows that $a$ continues to be chosen in a subset $B$ of $A$, and that $b$ continues to impact $a$ in $B$. In Appendix D we show that Contraction Consistency is implied by the other properties in the statement of Theorem 1.

We are now ready for the main result of this section.

**Theorem 2** *A stochastic choice rule is a dRUM if and only if it satisfies Regularity, Negative Expansion and Contraction Consistency.*

The proof is long and thus confined to an Appendix.

We note a subtle relation between the symmetric case of the theory and the *deterministic* "top-and-the-top" (TAT) model studied in Eliaz, Richter and Rubinstein [12]. A TAT is a choice procedure in which the agent uses two ordering and deterministically picks all the alternatives in a menu that are top in at least one of the two orderings. While we are dealing with a stochastic procedure, when the orderings have the same probabilities the resulting probabilities of choice are uninformative in distinguishing the orderings: the only information they give concerns their *support* (whereas in the asymmetric case the orderings can be told apart as the $\alpha$-ordering and the $(1-\alpha)$-ordering). So, in this special case the stochastic choice function contains in fact the same information on the rankings as a TAT. Note, however, the important point that of course our axioms do not only characterise the special case. The generic case they characterise is the asymmetric one, which has no correspondence with the deterministic procedure.[11]

# 5 Menu-dependent state probability and violations of Regularity

So far we have assumed that the probabilities with which the two rankings are applied are fixed across menus. This is quite natural, for example, in the population interpretation. Yet, in other scenarios the dependence of the ranking probabilities on the menu seems a compelling feature. We highlight some leading such cases.

- In the dual-self interpretation, if the duality of the self is due to temptation, then the presence of tempting alternatives may increase the probability that the short-term self is in control, and possibly the more so the more numerous the tempting alternatives are.
- In the household interpretation, husband and wife may have different 'spheres of control', so that menus containing certain items are more likely to be under the control of one of the two partners.
- If choices are subject to unobserved time pressure, the effect of time pressure is likely to be different in large and small menus. Since the latter are arguably easier to analyse, time pressure may be activated with a lower probability in simple menus.
- The logic of Luce and Raiffa's [24] well-known 'frog legs example' is that the presence of a specific item $a^*$ (frog legs in the example) in a menu triggers the maximisation of a different preference order because $a^*$ conveys information on the nature of

[11]The key property used by Eliaz, Richter and Rubinstein [12] to characterise TAT says that if an $x$ is chosen from two menus $A$ and $B$ and also from $A \cap B$, and if the choice from $A \cap B$ consists of exactly two elements, then $x$ is chosen from $A \cup B$. In addition their axiomatisation also includes a direct assumption on the number of chosen alternatives.

the available alternatives, so that the choice is 'dual': in all menus containing $a^*$ one preference order is maximised while in all those not containing $a^*$ a different order is maximised. Making the probabilities in a dRUM menu-dependent further generalises this idea to a probabilistic context, to allow for example $p\,(\text{steak}, \{\text{steak, chicken}\}) < p\,(\text{steak}, \{\text{steak, chicken, frog legs}\})$, while at the same time avoiding the extreme assumption that steak is chosen for sure when frog legs are also available.

- The 'similarity' and the 'attraction' effects (discussed below) are very prominent in the psychology and behavioural economics literature. According to these effects, the choice probabilities of two alternatives that are chosen approximately with equal frequency in a binary context can be shifted in favour of one or the other through the addition of an appropriate 'decoy' alternative. This decoy is in a relation of similarity or dominance (in a space of characteristics) with either one of the 'target' alternatives. Such effects can be described as dual RUM by making the probabilities of the two rankings dependent on the presence and nature of the decoy. Analogous considerations apply for the 'compromise effect', according to which the probability of choosing an alternative increases when it is located at an intermediate position in the space of characteristics compared to other more extreme alternatives.

A *Menu-dependent dRUM (mdRUM)* is a stochastic choice rule $p$ for which the following is true: there exists a triple $(r_1, r_2, \tilde{\alpha})$ where $r_1$ and $r_2$ are rankings on $X$ and $\tilde{\alpha} : 2^X \backslash \varnothing \to (0,1)$ is a function that for each menu $A$ assigns a probability $\tilde{\alpha}\,(A)$ to $r_1$ (and $1 - \tilde{\alpha}\,(A)$ to $r_2$), such that, for all $A$, $p\,(a, A) = p'\,(a, A)$ where $p'$ is the dRUM generated by $(r_1, r_2, \tilde{\alpha}\,(A))$.

Compared to a dRUM, an mdRUM loses all the properties that pertain to the specific magnitudes of choice probabilities, which depend on the menu in an unrestricted way. For example, an mdRUM fails Regularity. But it is precisely this feature that gives the model its descriptive power, while maintaining consistency properties that concern the *possibility*, the *impossibility* and the *certainty* of the event in which an alternative is chosen: in this sense such properties are about the 'mode' of choice.[12]

[12]'Mode' and 'modal' are meant here and elsewhere in the logical and not statistical meaning.

Let's say that $b$ *modally impacts* $a$ in $A$ if

$$\begin{aligned} p(a,A) &> 0 \text{ and } p(a, A\cup\{b\}) = 0 \\ \text{or } p(a,A) &= 1 \text{ and } p(a, A\cup\{b\}) \in (0,1) \end{aligned}$$

That is, $b$ modally impacts $a$ if adding $b$ transforms the choice of $a$ from possible (including certain) to impossible or from certain to merely possible.

The following properties are clearly necessary for an mdRUM:

**Modal Regularity**: Let $a \in B \subset A$. (i) If $p(a,A) > 0$ then $p(a,B) > 0$. (ii) If $p(a,A) = 1$ then $p(a,B) = 1$.

**Modal Impact Consistency**: Let $b \notin A$. If $b$ does not modally impact $a$ in $A$ for all $a \in A$ then $p(b, A\cup\{b\}) = 0$.

**Modal Contraction Consistency**: Suppose that $b$ modally impacts $a$ in $A$. Then $b$ modally impacts $a$ in any $B \subset A$.

Modal Regularity says that, after alternatives are removed from a menu, any alternative that was possibly chosen and is still feasible remains possibly chosen (including certainly chosen), and any alternative that was certainly chosen and is still feasible remains certainly chosen. Thus, Modal Regularity only excludes certain types of 'extreme' menu dependence, whereby the composition of the menu does not only affect the numerical values of the probabilities of choice, but the very possibility of choice.

The analog of Modal Impact Consistency holds in the menu-independent model, and indeed in *any* RUM: if $b$ is chosen with positive probability in a menu, then it impacts some alternative in any sub-menu. Observe that (unlike for Regularity) the modal version is stronger, as it states is conclusion from a weaker premise.

Modal Contraction Consistency is a straightforward modal analog of the corresponding axiom seen previously.

To prove quickly that the properties are sufficient as well as necessary we exploit a relationship that exists between mdRUMs and dRUMs. One can associate with each $p$ another stochastic choice rule $\hat{p}$ that is a symmetric dRUM if and only if $p$ is an mdRUM.

This trick allows us to rely on the previous results, as we then just need to 'translate' the properties that make $\hat{p}$ an mdRUM into properties of $p$, which is an easier task.

As in the previous results, we avoid any assumption explicitly relating to the cardinality of the support of $p$. So we first prove a lemma analogous to Lemma 1. Say that a stochastic choice rule $p$ is *binary* whenever it assigns positive probability to at most two elements in any menu (i.e. for any menu $A$, if $p(a, A) \in (0,1)$ for some $a \in A$ then there exists $b \in A$ for which $p(b, A) = 1 - p(a, A)$).

**Lemma 3** *Let $p$ be a stochastic choice rule that satisfies Modal Regularity and Modal Impact Consistency. Then $p$ is binary.*

**Proof:** Suppose by contradiction that for some menu $A$ there exist $b_1, ..., b_n \in A$ such that $n > 2$, $p(b_i, A) > 0$ for all $i = 1, ...n$ and $\sum_{i=1}^{n} p(b_i) = 1$. By Modal Regularity both $p(b_i, \{b_1, ..., b_{n-1}\}) > 0$ for all $i = 1, ..., n-1$ and $p(b_i, \{b_1, ..., b_n\}) > 0$ for all $i = 1, ..., n$. Then $b_n$ does not modally impact any alternative in $\{b_1, ..., b_{n-1}\}$, contradicting Modal Impact Consistency. $\square$

**Theorem 3** *A stochastic choice rule is an mdRUM if and only if it satisfies Modal Regularity, Modal Impact Consistency and Modal Contraction Consistency.*

**Proof.** Necessity is obvious. For sufficiency, given a stochastic choice rule $p$, let us say that $\hat{p}$ is the *conjugate of* $p$ if for all menus $A$ and $a \in A$:

$$\begin{aligned} \hat{p}(a, A) &= 1 \Leftrightarrow p(a, A) = 1 \\ \hat{p}(a, A) &= \frac{1}{2} \Leftrightarrow 1 > p(a, A) > 0 \\ \hat{p}(a, A) &= 0 \Leftrightarrow p(a, A) = 0 \end{aligned}$$

If $p$ satisfies Modal Regularity and Modal Impact Consistency, by Lemma 3 it is binary and so the conjugate of a $p$ always exists and is defined uniquely, but it is not necessarily a dRUM. When it is a dRUM, it is a symmetric dRUM. We now investigate when this is the case. Suppose that $p$ is an mdRUM generated by $(r_1, r_2, \tilde{\alpha})$. Then, in view of the relationship $p(a, A) = p'(a, A)$ where $p'$ is the dRUM generated by $(r_1, r_2, \tilde{\alpha}(A))$, its conjugate $\hat{p}$ is a dRUM generated by $\left(r_1, r_2, \frac{1}{2}\right)$, and therefore by Theorem 2 $\hat{p}$ satisfies Regularity, Negative Expansion and Contraction Consistency. Conversely, if the conjugate $\hat{p}$ of a $p$ satisfies these axioms then $\hat{p}$ is a dRUM generated by some $\left(r_1, r_2, \frac{1}{2}\right)$ and therefore $p$ is an mdRUM generated by $(r_1, r_2, \tilde{\alpha})$ where $\tilde{\alpha}$ is defined so that for any $A$ such that $p(a, A) > 0$ and $p(b, A) > 0$ for distinct $a$ and $b$, $\tilde{\alpha}(A) = p(a, A)$ and $1 - \tilde{\alpha}(A) = p(b, B)$ (or vice-versa). This reasoning shows that $p$ is an mdRUM if and only if it is binary and its conjugate satisfies Regularity, Negative Expansion and Contraction Consistency. Therefore we just need to verify that $\hat{p}$ satisfies these axioms if $p$ satisfies the axioms in the statement.

Step 1. $\hat{p}$ satisfies Regularity.

Let $A \subset B$. If $\hat{p}(a, B) = 0$ then Regularity cannot be violated. If $0 < \hat{p}(a, B) < 1$ then it must be $\hat{p}(a, B) = \frac{1}{2}$, which implies $p(a, B) > 0$ and hence by Modal Regularity part (i) $p(a, A) > 0$. Then $\hat{p}(a, A) \in \left\{\frac{1}{2}, 1\right\}$, satisfying Regularity. Finally if $\hat{p}(a, B) = 1$ then $p(a, B) = 1$, so that by part (ii) of Modal Regularity $p(a, A) = 1$, and then $\hat{p}(a, A) = 1$ as desired.

Step 2. $\hat{p}$ satisfies Negative Expansion.

If $\hat{p}(a, A) < \hat{p}(a, B) < 1$ it must be $\hat{p}(a, A) = 0$ and $\hat{p}(a, B) = \frac{1}{2}$, and therefore $p(a, A) = 0$. If $p(a, A \cup B) > 0$ then Modal Regularity part (i) is violated. Therefore $p(a, A \cup B) = 0$ and thus $\hat{p}(a, A \cup B) = 0$ as desired.

Step 3. $\hat{p}$ satisfies Contraction Consistency.

Suppose that $b$ impacts $a$ in $A$, that is $\hat{p}(a, A) > \hat{p}(a, A \cup \{b\})$. This means that either $\hat{p}(a, A) \in \left\{\frac{1}{2}, 1\right\}$ and $\hat{p}(a, A \cup \{b\}) = 0$, or $\hat{p}(a, A) = 1$ and $\hat{p}(a, A \cup \{b\}) = \frac{1}{2}$. Therefore either $p(a, A) \in (0, 1]$ and $p(a, A \cup \{b\}) = 0$, or $p(a, A) = 1$ and $p(a, A \cup \{b\}) \in (0, 1)$. Let $B \subset A$. Then by Modal Contraction Consistency either $p(a, B) > 0$ and $p(a, B \cup \{b\}) = 0$, or $p(a, B) = 1$ and $p(a, B \cup \{b\}) \in (0, 1)$. This means that either $\hat{p}(a, B) \in \left\{\frac{1}{2}, 1\right\}$ and $\hat{p}(a, B \cup \{b\}) = 0$, or $\hat{p}(a, B) = 1$ and $\hat{p}(a, B \cup \{b\}) = \frac{1}{2}$, so that in either case $b$ impacts $a$ in $B$. ■

As discussed, we consider it as an important feature of a stochastic theory of choice that it offers a pathway to representing anomalies related to violations of Regularity. In particular, marketers use a number of strategies to manipulate the attractiveness or otherwise of alternatives. The *attraction effect* (also known as the 'asymmetric dominance' effect, see Huber, Payne and Puto [20], Huber and Puto [21]) refers to the fact that the choice frequency of a target alternative $t$ increases when a new decoy alternative $d$ is introduced in a menu, with the property that the $d$ is markedly worse than the target $t$, while incomparable to a third ('other') alternative, $o$. This ranking is generally induced by presenting alternatives as described in two desirable attributes/dimensions: while the ranking between $t$ and $o$ in one dimension is reversed in the other, $d$ is Pareto dominated by $t$ but Pareto incomparable to $o$. The *compromise effect* instead refers to the introduction of a different type of decoy, which has the highest degree of one attribute and the lowest of another in such a way that $t$ is now 'middle ranking'.

However, subsequent research has identified various other strategies to increase $t$'s choice probability: the decoy may be Pareto dominated by both $t$ and $o$, or may Pareto dominate the target but be unavailable for choice (i.e. 'phantom' decoy), and so on. The compromise and attraction effects are just components of a family.

All these effects are easy to accommodate in our setup, by supposing that there are two rankings $\succ_i$ such that $t \succ_1 o \succ_1 d$ and $o \succ_2 t \succ_2 d$, and that the introduction of $d$ increases the probability of ranking $\succ_1$. The advantage of this way of modelling the phenomenon is that the mechanism holds regardless of the type of decoy that is introduced, whether it is a phantom alternative, or one that induces compromise, or a symmetrically dominated alternative, and so on. What conforms to the intentions of the manipulator and accords with the structure of our model is that the target alternative is made more appealing not by improvements to it, but simply by *framing*. For example, in the compromise effect it is hard to tell whether what is *behaviourally* a compromise-seeking attitude really reflects a compromise seeking *psychology*. Indeed, Mochon and Frederick [29] find experimentally that an *order* effect could be a more plausible explanation for the 'compromise effect': the alternative presented as second seems more salient in the choice between three items, regardless of its attributes. Menu-dependence of random orderings is a catch-all concept that gathers the factors, psychological or of other nature, that affect choice.

We conclude this section by noting yet another dimension, beside violations of Regularity, in which mdRUMs are behaviourally rich:

**Remark 2** mdRUM is consistent with violations of Weak Stochastic Transitivity. For example consider $p$ given by $p(a, \{a, b\}) = p(a, \{a, b, c\}) = p(b, \{b, c\}) = \frac{2}{3}$, $p(a, \{a, c\}) =$

$\frac{1}{3} = p(b, \{a, b, c\})$. Then $p$ violates Weak Stochastic Transitivity but it is an mdRUM generated by the rankings $a \succ_1 c \succ_1 b$ and $b \succ_2 c \succ_2 a$ with $\tilde{\alpha}(\{a, b\}) = \tilde{\alpha}(\{a, b, c\}) = \frac{2}{3}$ and $\tilde{\alpha}(\{b, c\}) = \tilde{\alpha}(\{a, c\}) = \frac{1}{3}$.

# 6 Identification

Identification is straightforward and unique in the generic asymmetric case. If $p$ is a dRUM generated by some $(r_1, r_2, \alpha)$, it is easily checked that $r_1$ and $r_2$ must satisfy:

$$
\begin{aligned}
a &\succ {}_1 b \Leftrightarrow p(a, \{a, b\}) \in \{\alpha, 1\} \\
a &\succ {}_2 b \Leftrightarrow p(a, \{a, b\}) \in \{1 - \alpha, 1\}
\end{aligned}
$$

Thus if $\alpha \neq 1 - \alpha$, the rankings can be *uniquely* (up to a relabelling of the rankings) inferred from any dRUM (this definition can be used to provide an alternative, constructive proof of Theorem 1). The uniqueness feature is lost when we drop the requirement of asymmetry. For example consider the dRUM $p$ generated by $\left(r_1, r_2, \frac{1}{2}\right)$ as follows

$$
\begin{gathered}
a \succ_1 b \succ_1 c \succ_1 d \\
b \succ_2 a \succ_2 d \succ_2 c
\end{gathered}
$$

The same $p$ could be alternatively generated by the two different rankings

$$
\begin{gathered}
a \succ_1 b \succ_1 d \succ_1 c \\
b \succ_2 a \succ_2 c \succ_2 d
\end{gathered}
$$

On the other hand, if these rankings had probabilities $\alpha$ and $1 - \alpha$ with $\alpha \neq \frac{1}{2}$, one could tell the two possibilities apart by observing whether $p(d, \{d, c\}) = \alpha$ or $p(d, \{d, c\}) = 1 - \alpha$.

More in general, the source of non-uniqueness when rankings have symmetric probabilities is the following. Suppose that, given the rankings $r_1$and $r_2$, the set of alternatives can be partitioned as $X = A_1 \cup ... \cup A_n$ such that if $a \in A_k$, $b \in A_l$ and $k < l$, then $a \succ_i b$ for $i = 1, 2$. Now construct two new rankings $r'_1$ and $r'_2$ in the following manner. Fix $A_k$, and for all $a, b \in A_k$ set $a \succ_i b$ if and only if $a \succ_j b$, $i \neq j$, while for all other cases $r'_1$

and $r'_2$ coincide with $r_1$ and $r_2$ (i.e. $r'_1$ and $r'_2$ swap the subrankings within $A_k$). Then it is clear that the probabilities generated by $\left(r'_1, r'_{2,}\frac{1}{2}\right)$ are the same as those generated by $\left(r_1, r_{2,}\frac{1}{2}\right)$. Apart from these "swaps", the identification is unique.

# 7 Concluding remarks

We have argued that dual random utility maximisation constitutes a parsimonious framework to represent the variability of choice in many contexts of interest, both in the fixed state probability version and in the new menu-dependent state probability version introduced in this paper.

The properties at the core of our analysis are simple and informative about how choices react to the merging, expansion and contraction of menus. As such, they provide the theoretical basis for comparative statics analysis and for inferring behaviour in larger menus from behaviour in smaller ones, or vice-versa.

The menu-dependent version of the model significantly extends the range of phenomena that the model can encompass, because it does not need to satisfy the full Regularity property but only a 'modal' version of it. As a consequence, for example the attraction and the compromise effects and Luce and Raiffa's 'frog legs' types of anomalies can be described as menu-dependent dual random utility maximisation, even if though they are incompatible with any standard RUM.

The following two issues, one theoretical and the other empirical, seem interesting to us for future research:

- Can we obtain appealing behavioural characterisations of random utility maximisation with at most $k$ states, with $2 < k \leq n!$?

- As noted in the introduction, bimodal choice distributions (of which the theory in this paper is a theoretical idealisation) are relatively common empirically. What type of econometric adaptation is best suited to treat this type of distributions when generated by the model?[13]

[13]For example, one could postulate a random, rather than deterministic, utility $u_i + \varepsilon$ conditional on each

# References


[1] Apesteguia, Jose, Miguel Angel Ballester and Jay Lu (2016) “Single-crossing Random Utility Models", mimeo

[2] Salvador Barberá and Prasanta K. Pattanaik (1986) “Falmagne and the Rationalizability of Stochastic Choices in Terms of Random Orderings”, *Econometrica* 54: 707-15

[3] Block, H.D. and J. Marschak (1960) “Random Orderings and Stochastic Theories of Responses”, in Olkin, I., S. G. Gurye, W. Hoeffding, W. G. Madow, and H. B. Mann (eds.) *Contributions to Probability and Statistics*, Stanford University Press, Stanford, CA.

[4] Caplin, Andrew, and Daniel Martin (2016), “The Dual-Process Drift Diffusion Model: Evidence from Response Times,” *Economic Inquiry*, 54 (2): 1274-1282.

[5] Chatterjee, Kalyan, and R. Vijay Krishna (2009) “A “Dual Self" Representation for Stochastic Temptation", *American Economic Journal: Microeconomics*, 1(2):148-67.

[6] Cherchye, Laurens, Bram De Rock and Frederic Vermeulen (2007) “The collective model of household consumption: a nonparametric characterization”, *Econometrica*, 75: 553-574.

[7] Chiappori, Pierre-Andre. (1988), “Rational household labor supply”, *Econometrica*, 56: 63-89.

[8] de Clippel, Geoffroy and Kfir Eliaz (2012) “Reason-based choice: a bargaining rationale for the attraction and compromise effects”, *Theoretical Economics* 7: 125–162

[9] Dufwenberg, Martin, and Astri Muren (2006) "Generosity, anonymity, gender", *Journal of Economic Behavior & Organization*, 61(1): 42-49.

[10] Echenique, Federico and Kota Saito (2015) “General Luce Model”, mimeo.


state $i$, with an appropriate specification of the errors $\varepsilon$ and of the dependence of the state utilities $u_i$ on observable factors.

[11] Eliaz, Kfir and Ran Spiegler (2006) “Contracting with Diversely Naive Agents”, *The Review of Economic Studies* 73(3): 689-714

[12] Kfir Eliaz, Michael Richter, and Ariel Rubinstein (2011) “Choosing the two finalists", *Economic Theory*, 46(2):211–219.

[13] Engelmann, Dirk, and Hans-Theo Normann (2010) “Maximum effort in the minimum-effort game,” *Experimental Economics*, 13(3): 249-259.

[14] Falmagne, Jean-Claude (1978). “A Representation Theorem for Finite Random Scale Systems", *Journal of Mathematical Psychology*, 18 (1): 52–72.

[15] Fiorini, Samuel (2004) “A short proof of a theorem of Falmagne”, *Journal of Mathematical Psychology* 48 (1): 80-82.

[16] Frolich, Norman, Joe Oppenheimer and J.Bertrand Moore (2001) “Some doubts about measuring self-interest using dictator experiments: the costs of anonymity”, *Journal of Economic Behavior and Organization*, 46: 271-90.

[17] Gill, David and Victoria Prowse (2016) “Cognitive Ability, Character Skills, and Learning to Play Equilibrium: A Level-k Analysis”, forthcoming in *Journal of Political Economy*.

[18] Gul, Faruk and Wolfgang Pesendorfer (2006) “Random Expected Utility”, *Econometrica*, 74: 121–146.

[19] Gul, Faruk, Paulo Natenzon and Wolfgang Pesendorfer (2014) “Random Choice as Behavioral Optimization”, *Econometrica*, 82(5): 1873-1912.

[20] Huber, J., J. W. Payne and C. Puto, 1982 “Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis", *Journal of Consumer Research*, 9: 90–98.

[21] Huber, J., and C. Puto, 1983 “Market boundaries and product choice: Illustrating attraction and substitution effects", *Journal of Consumer Research*, 10: 31–44.

[22] Lu, Jay and Kota Saito (2016) “Random Intertemporal Choice”, mimeo.

[23] Luce, R. Duncan (1959) Individual Choice Behavior; a theoretical analysis. Wiley: New York.

[24] Luce, R. Duncan and Howard Raiffa (1957), Games and Decisions. Wiley, New York.

[25] Manzini, Paola and Marco Mariotti (2014) "Stochastic Choice and Consideration Sets", *Econometrica*, 82(3):1153–1176.

[26] McClelland, Gary H., William D. Schulze, and Don L. Coursey (1993) "Insurance for low-probability hazards: A bimodal response to unlikely events", *Journal of risk and uncertainty*, 7(1): 95-116.

[27] McCausland, William J. and A.A.J. Marley (2014) "Bayesian inference and model comparison for random choice structures", *Journal of Mathematical Psychology* (62–63): 33-46.

[28] McCausland, William J., Clintin Davis-Stober, A. A. J. Marley, Sanghyuk Park and Nicholas Brown (2015) "Testing Random Utility Using Falmagne's Conditions Directly", http://apps.olin.wustl.edu/conf/SBIES/Files/pdf/2016/194.pdf

[29] Mochon, Daniel and Shane Frederick (2012) "Interpreting Compromise Effects", mimeo, Tulane University and Yale University School of Management.

[30] Plerou, Vasiliki, Parameswaran Gopikrishnan and H. Eugene Stanley (2003) "Two-phase behaviour of financial markets", *Nature*, 421: 130.

[31] Proto, Eugenio, Aldo Rustichini and Andis Sofianos (2014) "Higher Intelligence Groups Have Higher Cooperation Rates in the Repeated Prisoner's Dilemma", IZA DP No. 8499.

[32] Reutskaya, Elena, Rosemarie Nagel, Colin F. Camerer, and Antonio Rangel (2011) "Search Dynamics in Consumer Choice under Time Pressure: An Eye-Tracking Study", *American Economic Review*, 101:900-926.

[33] Rubinstein, Ariel (2007) "Instinctive and Cognitive Reasoning: A Study of Response Times", *Economic Journal*,117: 1243-1259.

[34] Rubinstein, Ariel (2016) "Typology of Players: Between Instinctive and Contemplative", *Quarterly Journal of Economics*, 131 (2): 859-890.

[35] Sur, Pragya , Galit Shmuel, Smarajit Bose and Paromita Dubey (2015) "Modeling Bimodal Discrete Data Using Conway-Maxwell-Poisson Mixture Models", *Journal of Business & Economic Statistics*, 33: 352-365.

# Appendices

## A Independence of the axioms - Theorem 1

The stochastic choice rules in cases A.1 and A.3 are not dRUMs since they take on more than two interior values. The dRUM in A.2 is not asymmetric.

**A.1: Regularity and Constant Expansion, but not Negative Expansion**

| | $\{a,b,c\}$ | $\{a,b\}$ | $\{a,c\}$ | $\{b,c\}$ |
|---|---|---|---|---|
| $a$ | $\frac{1}{3}$ | $\frac{1}{2}$ | $\frac{1}{3}$ | – |
| $b$ | 0 | $\frac{1}{2}$ | – | $\frac{1}{3}$ |
| $c$ | $\frac{2}{3}$ | – | $\frac{2}{3}$ | $\frac{2}{3}$ |

Table 2

**A.2: Regularity and Negative Expansion, but not Constant Expansion**

| | $\{a,b,c\}$ | $\{a,b\}$ | $\{a,c\}$ | $\{b,c\}$ |
|---|---|---|---|---|
| $a$ | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{1}{2}$ | – |
| $b$ | $\frac{1}{2}$ | $\frac{1}{2}$ | – | $\frac{1}{2}$ |
| $c$ | 0 | – | $\frac{1}{2}$ | $\frac{1}{2}$ |

Table 3

**A.3: Constant Expansion and Negative Expansion, but not Regularity**

| | $\{a,b,c\}$ | $\{a,b\}$ | $\{a,c\}$ | $\{b,c\}$ |
|---|---|---|---|---|
| $a$ | 0 | 0 | $\frac{1}{2}$ | – |
| $b$ | 1 | 1 | – | $\frac{2}{3}$ |
| $c$ | 0 | – | $\frac{1}{2}$ | $\frac{1}{3}$ |

Table 4

## B Independence of the axioms - Theorem 2

The stochastic choice rules in cases B.2 and B.3 are not dRUMs since they take on more than two interior values.

**B.1: Regularity and Negative Expansion, but not Contraction Consistency**

See example in Table **??**.

**B.2: Regularity and Contraction Consistency, but not Negative Expansion**

| | $\{a,b,c\}$ | $\{a,b\}$ | $\{a,c\}$ | $\{b,c\}$ |
|---|---|---|---|---|
| $a$ | $\frac{1}{3}$ | 0 | $\frac{1}{2}$ | – |
| $b$ | $\frac{1}{3}$ | $\frac{1}{3}$ | – | $\frac{2}{3}$ |
| $c$ | $\frac{1}{3}$ | – | $\frac{1}{2}$ | $\frac{1}{3}$ |

Table 5

**B.3: Negative Expansion and Contraction Consistency, but not Regularity**

| | $\{a,b,c\}$ | $\{a,b\}$ | $\{a,c\}$ | $\{b,c\}$ |
|---|---|---|---|---|
| $a$ | 1 | 0 | $\frac{1}{2}$ | – |
| $b$ | 0 | $\frac{1}{3}$ | – | $\frac{2}{3}$ |
| $c$ | 0 | – | $\frac{1}{2}$ | $\frac{1}{3}$ |

Table 6

## C Independence of the axioms - Theorem 3

**C.1: Modal Regularity (i) and (ii), Modal Contraction Consistency but not Modal Impact Consistency**

The stochastic choice rule in Table 7 cannot be a mdRUM since in $\{a,b,c\}$ three alternatives are chosen with positive probability, which cannot be all at the top of two rankings.

| | $\{a,b,c\}$ | $\{a,b\}$ | $\{a,c\}$ | $\{b,c\}$ |
|---|---|---|---|---|
| $a$ | $\frac{1}{3}$ | $\frac{1}{2}$ | $\frac{1}{2}$ | – |
| $b$ | $\frac{1}{3}$ | $\frac{1}{2}$ | – | $\frac{1}{2}$ |
| $c$ | $\frac{1}{3}$ | – | $\frac{1}{2}$ | $\frac{1}{2}$ |

Table 7

**C.2: Modal Regularity (i), Modal Impact Consistency, Modal Contraction Consistency but not Modal Regularity (ii)**

The stochastic choice rule in Table 8 cannot be a mdRUM since on the one hand $p(a, \{a,b,c\}) = 1$ implies that $a$ is the highest alternative in both rankings, while $p(b, \{a,b\}) > 0$ would require $b$ to be above $a$ in at least one ranking.

| | $\{a,b,c\}$ | $\{a,b\}$ | $\{a,c\}$ | $\{b,c\}$ |
|---|---|---|---|---|
| $a$ | 1 | $\frac{1}{2}$ | 1 | – |
| $b$ | 0 | $\frac{1}{2}$ | – | $\frac{1}{2}$ |
| $c$ | 0 | – | 0 | $\frac{1}{2}$ |

Table 8

**C.3: Modal Regularity (ii), Modal Impact Consistency, Modal Contraction Consistency but not Modal Regularity (i)**

| | $\{a,b,c\}$ | $\{a,b\}$ | $\{a,c\}$ | $\{b,c\}$ |
|---|---|---|---|---|
| $a$ | $\frac{1}{2}$ | 0 | 1 | – |
| $b$ | $\frac{1}{2}$ | 1 | – | $\frac{1}{2}$ |
| $c$ | 0 | – | 0 | $\frac{1}{2}$ |

Table 9

The stochastic choice rule in Table 9 cannot be a mdRUM since $p(b, \{a,b\}) = 1$ requires $b$ to be above $a$ in both rankings, while $p(a, \{a,b,c\}) > 0$ would require $a$ to be above $b$ in at least one ranking.

**C.4: Modal Regularity (i) and (ii), Modal Impact Consistency, but not Modal Contraction Consistency**

| | $\{a,b,c,d\}$ | $\{a,b,c\}$ | $\{a,b,d\}$ | $\{a,c,d\}$ | $\{b,c,d\}$ | $\{a,b\}$ | $\{a,c\}$ | $\{a,d\}$ | $\{b,c\}$ | $\{b,d\}$ | $\{c,d\}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $a$ | 0 | $\frac{1}{2}$ | $\frac{1}{2}$ | 0 | – | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{1}{2}$ | – | – | – |
| $b$ | 0 | 0 | 0 | – | 0 | $\frac{1}{2}$ | – | – | $\frac{1}{2}$ | $\frac{1}{2}$ | – |
| $c$ | $\frac{1}{2}$ | $\frac{1}{2}$ | – | $\frac{1}{2}$ | $\frac{1}{2}$ | – | $\frac{1}{2}$ | – | $\frac{1}{2}$ | – | $\frac{1}{2}$ |
| $d$ | $\frac{1}{2}$ | – | $\frac{1}{2}$ | $\frac{1}{2}$ | $\frac{1}{2}$ | – | – | $\frac{1}{2}$ | – | $\frac{1}{2}$ | $\frac{1}{2}$ |

Suppose that the example in Table **??** is an mdRUM. Then $p(a, \{a,b,c,d\}) = 0$ and $p(a, \{a,b,c\}) = \frac{1}{2}$ imply that $d$ is an immediate predecessor of $a$ in one of the two rankings on $\{a,b,c,d\}$. This must remain the case when the rankings are restricted to $\{a,b,d\}$, yet $p(a, \{a,b,d\}) = \frac{1}{2} = p(a, \{a,b\})$ imply that $d$ is not an immediate predecessor of $a$ in $\{a,b,d\}$ in either ranking.

# D An Implication

**Proposition 1** *Let $p$ be a stochastic choice rule that satisfies Constant Expansion, Regularity and Negative Expansion. Then it satisfies Contraction Consistency.*

**Proof**. Let $p$ satisfy the axioms (so that in view of Theorem 1 $p$ is an asymmetric dRUM), let $a \in B \subseteq A$ and $b \in X$, and suppose by contradiction that $p(a, A) > p(a, A \cup \{b\})$ but $p(a, B) \leq p(a, B \cup \{b\})$. Since by Regularity $p(a, B) \geq p(a, B \cup \{b\})$, it can only be $p(a, B) = p(a, B \cup \{b\})$. Also, by Regularity it must be $p(a, B) \geq p(a, A)$. However it cannot be $p(a, B) = p(a, A)$, for in that case $p(a, B \cup \{b\}) = p(a, A)$ and Constant Expansion would imply the contradiction

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

Then it must be $p(a, B) > p(a, A)$, so that the only possible configuration of choice probabilities is

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

Let $\alpha > 1 - \alpha$ with $\alpha \in (0, 1)$. In the above expression, it cannot be that $p(a, B \cup \{b\}) = \alpha$ and $p(a, A) = 1 - \alpha$, since then Negative Expansion and $p(a, B) \neq p(a, A)$ would imply the contradiction $p(a, A \cup B) = p(a, A) = 0 > p(a, A \cup \{b\})$. Similarly, it cannot be $p(a, A) = \alpha > p(a, A \cup \{b\}) = 1 - \alpha$, for again Negative Expansion would imply the contradiction

$$p(a, A \cup \{b\}) = p(a, A \cup (A \cup \{b\})) = 0 \neq 1 - \alpha = p(a, A \cup \{b\})$$

The only remaining possibility is $p(a, B \cup \{b\}) = p(a, B) = 1$, $p(a, A) \in \{\alpha, 1 - \alpha\}$ and $p(a, A \cup \{b\}) = 0$. By Regularity, $p(a, B \cup \{b\}) = 1$ implies $p(a, \{a, b\}) = 1$. Moreover if $p(a, A) > p(a, A \cup \{b\})$ it must be $p(b, A \cup \{b\}) > 0$ (if not, then there exist $x, y \in A$ such that $p(x, A \cup \{b\}), p(y, A \cup \{b\}) > 0$ and $p(x, A \cup \{b\}) + p(y, A \cup \{b\}) = 1$, where possibly $x = y$ while $a \neq x, y$ since $p(a, A \cup \{b\}) = 0$. But then by Regularity $p(x, A) \geq p(x, A \cup \{b\})$ and $p(y, A) \geq p(y, A \cup \{b\})$, implying the contradiction $p(a, A) = 0$). But if $p(b, A \cup \{b\}) > 0$ the contradiction $p(b, \{a, b\}) > 0$ follows. We conclude that Contraction Consistency holds. $\square$

# E Proof of Theorem 2

Necessity is straightforward, so we only show sufficiency.

***Preliminaries and outline*** Let $p$ satisfy Regularity, Negative Expansion and Contraction Consistency. The proof consists of several blocks. First (block 0), we show that the axioms imply the following weakening of Constant Expansion:

**Weak Constant Expansion**: If $p(a, A) = p(a, B) = \alpha$ and $p(a, A \cup B) > 0$ then $p(a, A \cup B) = \alpha$.

Unlike Constant Expansion, the weaker version is necessary for a general dRUM. Observe that Lemma 1 and Lemma 2 continue to hold when Constant Expansion is replaced by Weak Constant Expansion.[14] In the proof of Theorem 1 we did not use Constant Expansion explicitly (only through the lemmas). So in view of Theorem 1 we shall only need to consider the case $p(a, A) \in \left\{0, \frac{1}{2}, 1\right\}$ for all $A$ and $a \in A$, with $p(a, A) \in (0, 1)$

[14] To check, note that Constant Expansion is invoked only in the proof of Lemma 1, and only once, in a step where the additional clause of Weak Constant Expansion is satisfied by hypothesis.

for some $A$ and $a \in A$, and at most two alternatives chosen with positive probability in any menu $A$. We define an algorithm to construct two rankings $r_1$ and $r_2$ explicitly for this case (block 1). Next, we show that the algorithm is well defined (block 2), and finally we show that $r_1$ and $r_2$ so constructed retrieve $p$ (block 3).

*0.* ***If** $p$ **satisfies Regularity and Negative Expansion then it satisfies Weak Constant Expansion.***

*Proof.* By contradiction, suppose that $p(a, A \cup B) = \beta > 0$ and $p(a, A) = p(a, B) = \alpha \neq \beta$. Negative Expansion (applied to $A$ and $A \cup B$) is contradicted unless $\beta = 1$. Assuming this is the case, by Regularity $p(a, A) \geq p(a, A \cup B) = 1$ and $p(a, B) \geq p(a, A \cup B) = 1$, so that $\alpha = \beta$, a contradiction.

*1.* ***The algorithm to construct the two rankings when** $p(a, A) \in \left\{0, \frac{1}{2}, 1\right\}$.*

Let $a, b \in X$ be such that $p(a, X), p(b, X) > 0$. Enumerate the elements of the rankings $r_1$ and $r_2$ to be constructed by $r_1(i) = x_i$ and $r_2(i) = y_i$ for $i = 1, \ldots n$. Set $x_1 = a$ and $y_1 = b$ (were possibly $x_1 = a = b = y_1$).

The rest of the rankings $r_1$ and $r_2$ for $i \geq 2$ are defined recursively. Let

$$L_i^1 = X \setminus \bigcup_{j=1}^{i-1} x_j$$

Next, define the set $S_i$ of alternatives that are impacted by $x_{i-1}$ in $L_i^1$:

$$S_i = \left\{s \in L_i^1 : p\left(s, L_i^1\right) > p\left(s, L_i^1 \cup \{x_{i-1}\}\right)\right\}$$

where observe that by lemmas 1 and 2 we have $0 < |S_i| \leq 2$.

If $S_i = \{c\}$ for some $c \in X$, then let $x_i = c$.

If $|S_i| = 2$, then let $S_i = \{c, d\}$ with $c \neq d$ and consider two cases:

*(1.i)* $p\left(c, L_i^1\right) > p\left(c, L_i^1 \cup \{x_j\}\right)$ (resp., $p\left(d, L_i^1\right) > p\left(d, L_i^1 \cup \{x_j\}\right)$) for all $j = 1, \ldots i-1$ and $p\left(d, L_i^1\right) = p\left(d, L_i^1 \cup \{x_j\}\right)$ (resp., $p\left(c, L_i^1\right) = p\left(c, L_i^1 \cup \{x_j\}\right)$) for some $j \in \{1, \ldots, i-2\}$ (that is, one alternative is impacted by *all* predecessors $x_j$ in $L_i^1$ while the other alternative is not).

In this case let $x_i = c$ (resp., $x_i = d$).

*(1.ii)* $p\left(c, L_i^1\right) > p\left(c, L_i^1 \cup \left\{x_j\right\}\right)$ and $p\left(d, L_i^1\right) > p\left(d, L_i^1 \cup \left\{x_j\right\}\right)$ for all $j = 1, ...i-1$ (both alternatives are impacted by all predecessors $x_j$ in $L_i^1$).

In this case, let $x_i = c$.

Proceed in an analogous way for the construction of $r_2$, starting from $y_2$, that is for all $i = 2, ...n$ define recursively

$$L_i^2 = X \setminus \bigcup_{j=1}^{i-1} y_j$$

$$T_i = \left\{t \in L_i^2 : p\left(t, L_i^2\right) > p\left(t, L_i^2 \cup \{y_{i-1}\}\right)\right\}$$

and as before $0 < |S_i| \leq 2$.

If $|T_i| = \{e\}$for some $e \in X$, then let $y_i = e$, while if $|T_i| = 2$ , then letting $T_i = \{e, f\}$:

*(2.i)* $p\left(e, L_i^2\right) > p\left(e, L_i^2 \cup \left\{y_j\right\}\right)$ (resp. $p\left(e, L_i^2\right) > p\left(e, L_i^2 \cup \left\{y_j\right\}\right)$) for all $j = 1, ...i-1$, and $p\left(f, L_i^2\right) = p\left(f, L_i^2 \cup \left\{y_j\right\}\right)$ (resp. $p\left(e, L_i^2\right) = p\left(e, L_i^2 \cup \left\{y_{i-g}\right\}\right)$) for some $j \in \{1, ..., i-2\}$.

In this case let $y_i = e$ (resp., $f$).

*(2.ii)* $p\left(e, L_i^2\right) > p\left(e, L_i^2 \cup \left\{y_j\right\}\right)$ and $p\left(f, L_i^2\right) > p\left(f, L_i^2 \cup \left\{y_j\right\}\right)$ for all $j = 1, ...i-1$.

In this case let $y_i = f$ (resp., $y_i = e$) whenever $e \succ_1 f$ (resp., $f \succ_1 e$) (i.e. we require consistency with the construction of the first order).

***2. Showing that the algorithm is well-defined***

We show that cases *(1.i)* and *(1.ii)* are exhaustive, which means showing that if $|S_i| = 2$ at least one alternative in $S_i$ is impacted by *all* its predecessors in $r_1$. We proceed by induction on the index $i$. If $i = 2$ there is nothing to prove. Now consider the step $i = k+1$. If $|S_{k+1}| = 1$ again there is nothing to prove, so let $|S_{k+1}| = 2$, with $S_{k+1} = \{c, d\}$. By construction $p\left(c, L_{k+1}^1\right) > p\left(c, L_{k+1}^1 \cup \{x_k\}\right)$ and $p\left(d, L_{k+1}^1\right) > p\left(d, L_{k+1}^1 \cup \{x_k\}\right)$, so that by Lemma 1 and Lemma 2 it must be

$$p\left(c, L_{k+1}^1\right) = \frac{1}{2} = p\left(d, L_{k+1}^1\right) \tag{1}$$

By contradiction, suppose that there exist $u, v \in \{x_1, ...x_{k-1}\}$ for which $u$ does not impact $c$ in $L_{k+1}^1$ and $v$ does not impact $d$ in $L_{k+1}^1$, i.e.

$$p\left(c, L_{k+1}^1 \cup \{u\}\right) = p\left(c, L_{k+1}^1\right) = \frac{1}{2} \tag{2}$$

and

$$p\left(d, L_{k+1}^{1} \cup \{v\}\right) = p\left(d, L_{k+1}^{1}\right) = \frac{1}{2} \tag{3}$$

(note that by construction we have $u, v \neq x_k$). We can rule out the case $u = v$. For suppose $u = v = x_j$ for some $j < k$. Then, recalling (1),

$$p\left(d, L_{k+1}^{1} \cup \{x_j\}\right) = p\left(d, L_{k+1}^{1}\right) = \frac{1}{2} = p\left(c, L_{k+1}^{1}\right) = p\left(c, L_{k+1}^{1} \cup \{x_j\}\right)$$

and thus $p\left(x_j, L_{k+1}^{1} \cup \{x_j\}\right) = 0$, which contradicts Regularity and $p\left(x_j, L_j^{1}\right) > 0$ with $L_{k+1}^{1} \cup \{x_j\} \subset L_j^{1}$.

Since by construction and Regularity $p\left(v, L_{k+1}^{1} \cup \{v\}\right) > 0$, Lemma 1, Lemma 2 and (3) imply

$$p\left(c, L_{k+1}^{1} \cup \{v\}\right) = 0 \tag{4}$$

Similarly, $p\left(v, L_{k+1}^{1} \cup \{v\}\right) > 0$ (by construction and Regularity), Lemma 1, Lemma 2 and (2) imply

$$p\left(d, L_{k+1}^{1} \cup \{u\}\right) = 0 \tag{5}$$

(i.e. $u$ impacts $d$ in $L_{k+1}^{1} \cup \{u\}$ and $v$ impacts $c$ in $L_{k+1}^{1}$). Now consider the menu $L_{k+1}^{1} \cup \{u, v\}$. By (4), (5) and Regularity it must be:

$$p\left(c, L_{k+1}^{1} \cup \{u, v\}\right) = 0 = p\left(d, L_{k+1}^{1} \cup \{u, v\}\right) \tag{6}$$

It cannot be that $p\left(u, L_{k+1}^{1} \cup \{u, v\}\right) = 1$, for otherwise Regularity would imply $p\left(u, L_{k+1}^{1} \cup \{u\}\right) = 1$, contradicting $p\left(c, L_{k+1}^{1} \cup \{u\}\right) = p\left(c, L_{k+1}^{1}\right) = \frac{1}{2}$. Similarly, it cannot be that $p\left(v, \{u, v\} \cup L_{k+1}^{1}\right) = 1$. Finally, if either $p\left(v, \cup L_{k+1}^{1} \cup \{u, v\}\right) = 0$ or $p\left(u, L_{k+1}^{1} \cup \{u, v\}\right) = 0$, then in view of (6) it would have to be $p\left(w, L_{k+1}^{1} \cup \{u, v\}\right) > 0$ for some $w \in L_{k+1}^{1}$. But this is impossible since $c \neq w \neq d$ by (6), and then by Regularity the contradiction $p\left(w, L_{k+1}^{1}\right) > 0$ would follow. Therefore it must be

$$p\left(u, L_{k+1}^{1} \cup \{u, v\}\right) = \frac{1}{2} = p\left(v, L_{k+1}^{1} \cup \{u, v\}\right)$$

It follows that both

$$p\left(u, L_{k+1}^{1} \cup \{u\}\right) = p\left(u, L_{k+1}^{1} \cup \{u, v\}\right) \tag{7}$$

and

$$p\left(v, L_{k+1}^{1} \cup \{v\}\right) = p\left(v, L_{k+1}^{1} \cup \{u, v\}\right) \tag{8}$$

i.e. neither does $u$ impact $v$ in $L^1_{k+1} \cup \{v\}$, nor does $v$ impact $u$ in $L^1_{k+1} \cup \{u\}$. Suppose w.l.o.g. that $v$ is a predecessor of $u$, and let $u = x_j$. By the inductive hypothesis $v$ impacts $u$ in $L^1_j$, so that by Contraction Consistency it also impacts $u$ in $L^1_{k+1} \cup \{u\} \subset L^1_j$, i.e.

$$p\left(u, L^1_{k+1} \cup \{u\}\right) > p\left(u, L^1_{k+1} \cup \{u, v\}\right)$$

a contradiction with (7). A symmetric argument applies if $u$ is a predecessor of $v$ using (8).

A straightforward adaptation of the argument above shows that cases *(2.i)* and *(2.ii)* are exhaustive.

***3. Showing that the algorithm retrieves the observed choice***.

Let $p_{\frac{1}{2}}$ be the dRUM generated by $\left(r_1, r_2, \frac{1}{2}\right)$. For any alternative $x$ denote $L^i_x$ its (weak) lower contour set in ranking $r_i$, that is

$$L^i_x = \{x\} \cup \{s \in X : x \succ_i s\}$$

and note that by construction $p\left(x, L^i_x\right) > 0$. We examine the possible cases of failures of the algorithm in succession.

<u>3.1</u>: $p(a, A) = 0$ and $p_{\frac{1}{2}}(a, A) > 0$.

Then $a \succ_i a'$ for some $i$, for all $a' \in A \setminus \{a\}$, hence $A \subseteq L^i_a$, and thus by Regularity and $p\left(a, L^i_a\right) > 0$ we have $p(a, A) > 0$, a contradiction.

<u>3.2</u>: $p(a, A) = 1$ and $p_{\frac{1}{2}}(a, A) < 1$.

Then there exists $b \in A$ such that $b \succ_i a'$ for some $i$, for all $a' \in A \setminus \{b\}$, hence $A \subseteq L^i_b$, and thus by Regularity and $p\left(b, L^i_a\right) > 0$ we have $p(b, A) > 0$, a contradiction.

<u>3.3:</u> $p(a, A) = \frac{1}{2}$ and $p_{\frac{1}{2}}(a, A) = 1$.

Let $\mathcal{I}\left(p_{\frac{1}{2}}\right)$ be the set of all pairs $(a, A)$ satisfying the conditions of this case. Fix an $(a, A) \in \mathcal{I}\left(p_{\frac{1}{2}}\right)$ that is maximal in $\mathcal{I}\left(p_{\frac{1}{2}}\right)$ in the sense that $(b, B) \in \mathcal{I}\left(p_{\frac{1}{2}}\right) \Rightarrow a \succ_i b$ for some $i$.

Because $p_{\frac{1}{2}}(a, A) = 1$ we have $A \subseteq L^1_a \cap L^2_a$. If $p\left(a, L^1_a \cap L^2_a\right) = 1$ we have an immediate contradiction, since by Regularity $p(a, A) = 1$. So it must be $p\left(a, L^1_a \cap L^2_a\right) = \frac{1}{2}$. We show that this also leads to a contradiction.

Suppose first that $L_a^1 \setminus L_a^2 \neq \varnothing$, and take $z \in L_a^1 \setminus L_a^2$ such that $a \succ_1 z$ and $z \succ_2 a$. By construction this implies $p\left(a, L_a^2\right) > p\left(a, L_a^2 \cup \{z\}\right)$, so that by Contraction Consistency $p\left(a, L_a^1 \cap L_a^2\right) > p\left(a, \left(L_a^1 \cap L_a^2\right) \cup \{z\}\right)$. Since by Regularity $p\left(a, \left(L_a^1 \cap L_a^2\right) \cup \{z\}\right) \geq p\left(a, L_a^2\right) = \frac{1}{2}$ (where recall $z \in L_a^1$, so that $L_a^1 \cup \{z\} = L_a^1$), we have $p\left(a, L_a^1 \cap L_a^2\right) = 1$, a contradiction. A symmetric argument applies if $L_a^2 \setminus L_a^1 \neq \varnothing$.

Finally, consider the case $L_a^1 = L_a^2 = L_a$ for some $L_a \subset X$. Then $X \setminus L_a^1 = X \setminus L_a^2$. Define $U_a = X \setminus L_a$ (where observe that $u \in U_a \Rightarrow u \succ_i a$, $i = 1, 2$). Since we assumed that $p\left(a, L_a^1 \cap L_a^2\right) = \frac{1}{2}$, there exists a $w \in \left(L_a^1 \cap L_a^2\right)$ for which $p\left(w, L_a^1 \cap L_a^2\right) = \frac{1}{2}$.

Let $x_1, ..., x_n$ be the predecessors of $a$ in the ranking $r_1$ and let $y_1, ..., y_n$ be the predecessors of $a$ in the ranking $r_2$ (where obviously $\{x_1, ..., x_n\} = U_a = \{y_1, ..., y_n\}$). So $a = x_{n+1} = y_{n+1}$ (note that since $L_a^1 = L_a^2$, $a$ must have the same position in both rankings).

By the construction of the algorithm $x_n$ and $y_n$ impact $a$ in $L_a$, that is $p\left(a, L_a\right) > p\left(a, L_a \cup \{x_n\}\right)$ and $p\left(a, L_a\right) > p\left(a, L_a \cup \{y_n\}\right)$.

We claim that there exists a $z \in U_a$ that does not impact $w$ in $L_a$, that is $p\left(w, L_a\right) = p\left(w, L_a \cup \{z\}\right)$. To see this, if $p\left(w, L_a\right) = p\left(w, L_a \cup \{x_n\}\right)$ (resp. $p\left(w, L_a\right) = p\left(w, L_a \cup \{y_n\}\right)$) simply set $z = x_n$ (resp. $z = y_n$). If instead $p\left(w, L_a\right) > p\left(w, L_a \cup \{x_n\}\right)$ and $p\left(w, L_a\right) > p\left(w, L_a \cup \{y_n\}\right)$ then, since $a = x_{n+1} = y_{n+1}$, $p\left(w, L_a\right) = p\left(w, L_a \cup \{z\}\right)$ for some $z \in U_a$ by the construction of the algorithm (otherwise, given that $x_n$ and $y_n$ impact $w$ in $L_a$, it should be $w = x_{n+1}$ or $w = y_{n+1}$). Fix such a $z$.

Regularity and the construction imply that $p\left(z, L_a \cup \{z\}\right) = \frac{1}{2}$. But since $z \in U_a$ we have $z \succ_i a'$, $i = 1, 2$, for all $a' \in L_a$, so that $p_{\frac{1}{2}}\left(z, L_a \cup \{z\}\right) = 1$ and therefore $z \in \mathcal{I}\left(p_{\frac{1}{2}}\right)$. And since in particular $z \succ_i a$, $i = 1, 2$, the initial hypothesis that $(a, A)$ is maximal in $\mathcal{I}\left(p_{\frac{1}{2}}\right)$ is contradicted.

<u>3.4</u>: $p\left(a, A\right) = \frac{1}{2}$ and $p_{\frac{1}{2}}\left(a, A\right) = 0$.

By construction there exist alternatives $b$ and $c$ (where possibly $b = c$) such that $b \succ_1 a'$ for all $a' \in A \setminus \{b\}$ and $c \succ_2 a'$ for all $a' \in A \setminus \{c\}$, so that $A \subseteq L_b^1$ and $A \subseteq L_c^2$. If $b \neq c$, then by Regularity the contradiction $p\left(b, A\right) = \frac{1}{2} = p\left(c, A\right)$ follows. Thus let $b = c$, so that $b \succ_i a'$ for all $a' \in A \setminus \{b\}$ for $i = 1, 2$, and by construction $p\left(b, L_b^1 \cap L_b^2\right) > 0$. If $p\left(b, L_b^1 \cap L_b^2\right) = 1$ then by Regularity we have a contradiction, since $A \subseteq L_b^1 \cap L_b^2$, so that

$p\left(b, L_b^1 \cap L_b^2\right) = p\left(b, A\right) = \frac{1}{2}$, and otherwise we are back in case 3.3. ■

Note in passing that it would not be possible to simply weaken Constant Expansion to Weak Constant Expansion (defined in block 0 in the proof above) in Theorem 1 to obtain a characterisation of general dRUMs. The example in Table **??** shows that these axioms are not tight enough for the purpose.