Multinomial logit processes and preference discovery: inside and outside the black box

S. Cerreia-Vioglio, F. Maccheroni, M. Marinacci
Università Bocconi

A. Rustichini
University of Minnesota

September 14, 2018

Abstract

We provide both an axiomatic and a neuropsychological characterization of the dependence of choice probabilities on time in the *softmax* (or *Multinomial Logit Process*) form

$$p_t(a, A) = \frac{e^{\lambda(t)u(a)}}{\sum_{b \in A} e^{\lambda(t)u(b)}} \tag{MLP}$$

where: $p_t(a, A)$ is the probability that alternative a is selected from the set A of feasible alternatives if t is the time available to decide, u is a utility function on the set of all alternatives, and λ is an accuracy parameter on a set of time points.

MLP is a very popular model of random choice in many fields of decision making, from Quantal Response Equilibrium theory to Discrete Choice Analysis, Marketing, Psychophysics, and Neuroscience. Our axiomatic characterization of softmax permits to empirically test its descriptive validity and to better understand its conceptual underpinnings as a theory of agents' rationality. Our neuropsychological foundation provides a computational model that may explain softmax emergence in human behavior and that naturally extends to multialternative choice the classical Diffusion Model paradigm of binary choice. As we discuss in the paper, these complementary approaches provide a complete perspective on softmaximization as a model of preference discovery, both in terms of internal (neuropsychological) causes and external (behavioral) effects.

Keywords: Discrete Choice Analysis, Drift Diffusion Model, Heteroscedastic Extreme Value Models, Luce Model, Metropolis Algorithm, Multinomial Logit Model, Quantal Response Equilibrium

Acknowledgements: A first draft of this paper was circulated under the title Law of Demand and Forced Choice, IGIER Working Paper 593, 2016. We thank Jerome Adda, Carlo Baldassi, Pierpaolo Battigalli, Patrick Beissner, Andrei Borodin, Roberto Corrao, Federico Echenique, Agata Farina, Marcelo Fernandez, Loic Grenie, Philip Holmes, Ryota Iijima, Michael Konig, Giacomo Lanzani, Laura Maspero, Thomas Palfrey, Charles Plott, Antonio Rangel, Roger Ratcliff, Giorgia Romagnoli, Kota Saito, Larry Samuelson, Vaibhav Srivastava, Jakub Steiner, Tomasz Strzalecki, the Editor (Joel Sobel) and four Anonymous Referees, as well as the audiences at Caltech, Princeton, Queen Mary University, University of Southampton and at the workshops "Time Uncertainties and Strategies" (December 2016, PSE Paris), "Decision Making and Imprecise Beliefs" (March 2017, HEC Paris), "Decision: Theory, Experiments, and Applications" (May 2017, IHP Paris), "European Working Group on Multicriteria Decision Aiding" (September 2017, Dauphine Paris), "INFORMS Annual Meeting" (October 2017, Houston TX), "VT Decision Workshop" (April 2018, Virginia Tech), "Mathematics of Behavioral Economics and Knightian Uncertainty in Financial Markets" (May 2018, Bielefeld University), "Risk, Uncertainty and Decision" (June 2018, Heidelberg University), "LOFT" (July 2018, Bocconi University) for helpful suggestions. We especially thank Jay Lu with whom we proved early versions of some of the results of Section 3.2. Simone Cerreia-Vioglio and Massimo Marinacci gratefully acknowledge the financial support of ERC (grants SDDM-TEA and INDIMACRO, respectively).

1 Introduction

Both human and machine decisions must be often made under external time constraints. Introspection and empirical evidence agree that time scarcity leads to mistakes, even when the agent's objectives are clearly defined.

In this paper, we provide both an axiomatic and a neuropsychological characterization of the dependence of choice probabilities on time in the *softmax* (or *Multinomial Logit Process*) form

$$p_t(a, A) = \frac{e^{\lambda(t)u(a)}}{\sum_{b \in A} e^{\lambda(t)u(b)}} \tag{MLP}$$

where:

- $p_t(a, A)$ is the probability that alternative a is selected from the set A of feasible alternatives if t is the time available to decide,[1]
- $u : X \to \mathbb{R}$ is a utility function on the set X of all alternatives,
- $\lambda : T \to (0, \infty)$ is an accuracy parameter on a set T of time points.

For each fixed time t, our theory coincides with the one of Luce (1959), that is,

$$p_t(a, A) = \frac{e^{v_t(a)}}{\sum_{b \in A} e^{v_t(b)}}$$

where v_t is a function from X to $\mathbb{R}$.

Our **axiomatic contribution** consists in characterizing, in terms of observables, the time dynamics of the functions v_t that corresponds to MLP. Specifically, by observing choice frequencies, we are able to establish whether there exist a **time-independent** utility u and a **time-dependent** accuracy parameter λ such that

$$v_t(x) = \lambda(t)\, u(x) \qquad \forall (x, t) \in X \times T$$

and, if so, to identify them. Moreover, for $T = (0, \infty)$ we show under which conditions λ is increasing and bijective. In this case, when the agent has no time to decide, all alternatives become equiprobable

$$\lim_{t \to 0} p_t(a, A) = 1/|A| \qquad \forall a \in A$$

and choice is completely random. In contrast, longer times allow for greater differences in the selection probability for alternatives that differ in their utilities until, without time pressure, only the utility maximizing ones survive

$$\lim_{t \to \infty} p_t(a, A) = \begin{cases} 1/|\arg\max_A u| & \text{if } a \in \arg\max_A u \\ 0 & \text{else} \end{cases}$$

and so choice is perfectly rational.

[1] Or the experience level of the agent, see below.

In terms of random utility (see McFadden, 1973), expression (MLP) amounts to

$$p_t(a, A) = \Pr\left\{u(a) + \frac{\epsilon_a}{\lambda(t)} > u(b) + \frac{\epsilon_b}{\lambda(t)} \quad \text{for all } b \in A \setminus \{a\}\right\}$$

where $\{\epsilon_x\}_{x \in X}$ is a collection of independent errors with type I extreme value distribution, mean 0 and variance $\pi^2/6$. Therefore, our axioms characterize the behavior of an agent who is trying to maximize u over A, but – because of time pressure – makes mistakes in evaluating the utility of the various alternatives. These mistakes are i.i.d. and their standard deviation is inversely proportional to $\lambda(t)$.

Our **neuropsychological contribution** consists in presenting a computational model that extends the classical *Drift Diffusion Model* (*DDM*) of Ratcliff (1978) to multialternative decision tasks and leads to choice probabilities that are again described by (MLP). Specifically, we start form the observation of Russo and Rosen (1975) – confirmed by more recent research (see below) – that multialternative choice procedures are composed primarily of sequential **pairwise comparisons**, in which actual evaluative processing takes place, and that agents' **exploration strategies** are based on the similarity and proximity of available alternatives. Formally, we assume that:

- pairwise comparisons can be described by the DDM with parameters u and $\lambda(t)$; so that in confronting a and b, one of the two alternatives is selected when the net evidence in its favor reaches the decision threshold $\lambda(t)$;
- exploration strategies follow a Markovian symmetric protocol *à la* Metropolis et al. (1953); which describes the probability of considering a new proposal c after having compared a and b and selected one of them;
- search terminates when t units of time have elapsed, at which point the incumbent solution is chosen.

We show that the decision process that we just described leads, approximately, to (MLP).

Therefore, softmaximization can be seen as the stylized output of a search procedure that looks for a maximum under time pressure, when pairwise comparisons are based on the accumulation of noisy evidence (and hence they are time consuming and subject to mistakes).

As we further discuss below, softmax is a very popular model of random choice in many fields of decision making, from Quantal Response Equilibrium theory to Discrete Choice Analysis, Marketing, Psychophysics, and Neuroscience. Our axiomatic characterization of Multinomial Logit Processes permits to empirically test their descriptive validity and to better understand their conceptual underpinnings as a theory of agents' (possibly bounded) rationality. Our neuropsychological foundation provides a computational model that may explain softmax emergence in human behavior and that naturally extends to multialternative choice the classical Diffusion Model paradigm of binary choice.

1.1 Different fields of application and our contribution

1.1.1 Discovered Preference Hypothesis and Quantal Response Equilibrium

On the theoretical side, softmax can be regarded as a formalization of the Discovered Preference Hypothesis (DPH) outlined in Plott (1996). According to the DPH, agents acquire an understanding of how their basic needs are satisfied by the different alternatives in the choice environment through a process of reflection and practice; a process which, in the long run, leads to optimizing behavior.

In our model, both reflection and practice are synthesized by the passage of time. This is best seen by considering the two aspects separately: If t represents the time limit within which an alternative a must be chosen from a feasible set A, then preference discovery occurs by reflection only. This is our main interpretation of t throughout the paper. Else, if t represents the number of times the agent has been facing choice problem A and a is an action to be taken, then preference discovery mainly occurs by practice. In this case, the typical finding that λ is increasing captures the dynamics of exploration in earlier decision stages and exploitation in later ones.

By *preference discovery* we mean both the **(partially observable) process** that leads the agent to consider one alternative superior to another and its **(fully observable) outcome**, that is, the resulting choice probabilities. For example, in the DDM case, the discovery process is given by Brownian evidence accumulation (see, e.g., Rangel and Clithero, 2014, Shadlen and Shohamy, 2016, and Fudenberg, Strack, and Strzalecki, 2017) and the discovery outcome is a binomial logit probability (see, e.g., Webb, 2018).

Softmaximization is the form that preference discovery takes in the *Quantal Response Equilibrium (QRE)* theory of McKelvey and Palfrey (1995). In this case, t is the experience level of the player, that is, the number of times he played the game,[2] $u(a)$ is the expected payoff of action a, and $\lambda(t)$ captures the player's degree of rationality. From the original data analysis of McKelvey and Palfrey (1995) to the recent Agranov, Caplin, and Tergiman (2015), Goeree, Holt, and Palfrey (2016), and Ortega and Stocker (2016) evidence seems to suggest that, for sophisticated players, the function λ increases as time passes and the decision making environment is better understood.[3]

Our axiomatic and neuropsychological characterizations of softmax can thus be seen as two alternative foundations of QRE. The first identifies the discovery outcome, the second explains the discovery process. QRE is thus the equilibrium concept that corresponds to the decision theory we develop in this paper. Goeree, Holt, and Palfrey (2016) give a broad perspective of its different applications.

---

[2] Our use of the pronoun "he" is completely nature-neutral, our agents might be algorithms.

[3] Interestingly, in Agranov, Caplin, and Tergiman (2015) and Ortega and Stocker (2016), t is not the experience level, but the time the player had to contemplate the altenatives in A before choosing.

1.1.2 Discrete Choice Analysis and Heteroscedastic Multinomial Logit Models

Ideas that are similar to the DPH emerged, in the same years, in the field of *Discrete Choice Analysis (DCA)*.[4] Relative to this literature, the model we propose falls in the category of Heteroscedastic Multinomial Logit Models. According to these models

$$p_s(a, A) = \frac{e^{\lambda(s)u(a)}}{\sum_{b \in A} e^{\lambda(s)u(b)}} = \Pr\left\{u(a) + \frac{\epsilon_a}{\lambda(s)} > u(b) + \frac{\epsilon_b}{\lambda(s)} \quad \forall b \in A \setminus \{a\}\right\} \qquad \text{(HMNL)}$$

where:

- $p_s(a, A)$ is the probability that alternative a is selected from the set A of feasible alternatives in **experimental condition** s, or, more in general, according to **data source** s,
- for each $x \in X$, $u(x)$ is the systematic component of the random utility $u(x) + \epsilon_x/\lambda(s)$,
- for each source $s \in S$, $\pi^2/6\lambda^2(s)$ is the variance of the random utility $u(x) + \epsilon_x/\lambda(s)$.

The main experimental conditions s that we consider in this paper are externally imposed deliberation times. Chen, Chorus, Molin, and van Wee (2015) present an application of the HMNL to travel behavior under these conditions. Instead, in Bradley and Daly (1994), again on transportation preferences, s is the number of pairwise choices completed in a survey, so it is similar in concept to the experience level of QRE; but they find a decreasing λ that they attribute to responders' fatigue. Under the same interpretation of s, Savage and Waldman (2008) and Campbell, Boeri, Doherty, and Hutchinson (2015) estimate the values of λ in early, middle, and late phases of different experiments with the aim of capturing preference learning (decreasing variance – increasing λ) and fatigue (increasing variance – decreasing λ).

From a DCA viewpoint, our axiomatic characterization of the HMNL allows to test for misspecification of the model itself. Moreover, we also provide simple techniques to directly identify parameters from data (see Proposition 4). In return, the DCA literature started by Ben-Akiva and Morikawa (1991), Swait and Louviere (1993), Hensher and Bradley (1993), and Bhat (1995) provides a number of methods to estimate the parameters of our model.

Given the broad use of HMNL, also in cases in which the set S of sources is not a subset of $(0, \infty)$ and is not endowed with a natural order,[5] in Appendix A.2 we complete our analysis with an axiomatic characterization of its most general version. As Theorem 3 of Luce (1959) provides a theoretical foundation for the applications of the (homoscedastic) MNL, our Theorem 16 provides one for the applications of the (heteroscedastic) HMNL.

[4]For example, Hensher, Louviere, and Swait (1999) write: "Consumers first become aware of needs and/or problems, and then search for information about products/services that could satisfy their need(s)/solve their problems. As consumers learn about products, product attributes and attribute levels, and associated uncertainties, they form beliefs about products. Eventually, they have sufficient information to form a decision strategy (or utility function) for valuing and trading off various product attributes."

[5]For example, S is a set of locations in Train (2009, pp. 24-25), it is a doubleton distinguishing between stated intentions and market choices in Ben-Akiva and Morikawa (1990), and it can also be a set of levels of (one or more) personality traits like in Proto, Rustichini, and Sofianos (2017). In healthcare experiments, the elements of S have been used to represent different levels of medical training (see, e.g., Vass, Wright, Burton, and Payne, 2018).

1.1.3 Psychophysics and Neuroscience

In discrimination tasks where pairs $A = \{a, b\}$ of alternatives are compared, (MLP) describes the "logistic psychometric function family" and is the most used parameterization of response probabilities.[6] In this case, λ is the slope of the psychometric function and it measures accuracy of discrimination or – in the neuroscientific study of choice behavior – sensitivity to utility differences.[7]

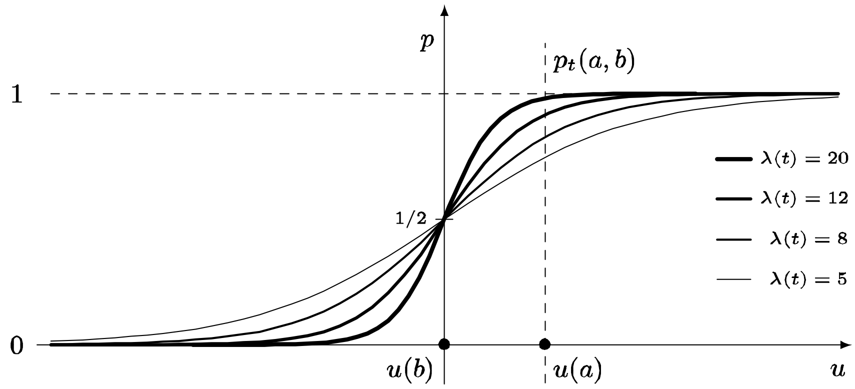


Figure 1: $p_t(a, b) = p_t(a, \{a, b\})$ plotted as a function of $u = u(a)$, for fixed $u(b)$.[8]

When deliberation time t is externally controlled (and scarce relative to the difficulty of the decision task), it is a well established fact of classical psychophysics that the slope of the psychometric function increases in t. In other words, externally decreasing deliberation time has the effect of diminishing the probability of a correct response. More in general, and irrespective of the number of alternatives, from the pioneering experiments of Cattel (1902) on perceived intensity of light, to the classical Link (1978), to the recent study of ALQahtani et al. (2016) on the effect of time pressure on diagnostic decisions, evidence has been systematically confirming that accuracy is an increasing function of the time available to decide. In particular, Ortega and Stocker (2016) calibrate softmax and find that λ is approximately linear in log-time, a special case that we are able to axiomatically characterize as a byproduct of our analysis (see Appendix A.2).

Based on a sequence of DDM pairwise comparisons, our neuropsychological foundation of softmax provides a new model of choice among multiple alternatives that captures the evidence above. Moreover, this model seems to be consistent with both old and new eye-tracking findings and with the known capacity constraints of working memory (see Section 4). Our axiomatic characterization, instead, allows to externally validate the model itself.

1.2 Related literature

An important distinction in the modelling of stochastic choice processes is whether the final decision time t is exogenous (the time limit set by the experimenter or by the environment)

[6] See Falmagne (1985) and Iverson and Luce (1998).

[7] Dovencioglu, Ban, Schofield, and Welchman (2013) and Tsunada, Liu, Gold, and Cohen (2016) are recent neuroscience works that use this specification of the psychometric function and adopt the corresponding interpretation of the parameter λ.

[8] Geometrically, $\lambda(t)/4$ is the slope of the tangent to the graph at $(u(b), 1/2)$.

or endogenous – the actual response time of a subject who chooses when to make his decision without external time pressure. This paper considers exogenous deliberation times and the present discussion of the literature focuses on models with this feature. In Economics, models where decision time is endogenously – say, optimally – chosen are the subject of an active literature and we refer the reader to Woodford (2014), Steiner, Stewart, and Matejka (2017), and Fudenberg, Strack, and Strzalecki (2017) for updated perspectives. In Psychology, these models have a long history, arguably beginning with Laming (1968). We refer to Bogacz et al. (2006) for more recent developments.

Natenzon (2017) proposes a Multinomial Bayesian Probit model with the aim of jointly accommodating similarity effects, attraction effects, and compromise effects in a preference learning perspective. According to Natenzon's model when facing a menu of alternatives the decision maker (who has a priori i.i.d. standard normally distributed beliefs the on the possible utilities of alternatives) receives a random vector of – jointly normally distributed – signals that represents how much he is able to learn about the ranking of alternatives before making a choice (say within time t). The decision maker updates the prior according to Bayes' rule and chooses the option with the highest posterior mean utility. Conceptually, both Natenzon's model and ours can be seen as formal embodiments of the Discovered Preference Hypothesis. Natenzon's Multinomial Bayesian Probit predates ours and gains descriptive power by abandoning the random utility paradigm and augmenting the number of parameters. At the same time, our model is easier to falsify thanks to its axiomatic nature and because of the abundance of routine techniques to estimate it. From a neuropsychological viewpoint, both approaches belong to the diffusion models' tradition that started with the seminal works of Ratcliff, Busemeyer, and coauthors (see the reviews of Rieskamp, Busemeyer, and Mellers, 2006, Fehr and Rangel, 2011, Ratcliff, Smith, Brown, and McKoon, 2016).

In a general Random Expected Utility perspective, Lu (2016) captures preference learning through increasingly informative priors on the set of probabilistic beliefs of the decision maker, while in Caplin and Dean (2011) preference learning occurs through sequential search.

Fudenberg and Strzalecki (2015) axiomatize a discounted adjusted logit model. Differently from the present work and the ones discussed above, their paper studies stochastic choice in a dynamic setting where choices made today can influence the possible choices available tomorrow, and consumption may occur in multiple periods. Frick, Iijima, and Strzalecki (2017) characterize its general random utility counterpart. Finally, Caplin and Dean (2013) and Matejka and McKay (2015) identify the Multinomial Logit Model in terms of optimal information acquisition, while Saito (2017) obtains several characterizations of the Mixed Logit Model.

As to the neuropsychological modelling of choice tasks with $N > 2$ alternatives, the vast majority of extensions of the DDM considers **simultaneous evidence accumulation** for all the N alternatives in the menu. In these models, the choice task is assumed to **simultaneously** activate N accumulators, each of which is primarily sensitive to one of the alternatives and integrates evidence relative to that alternative; choices are then made based on absolute or relative evidence levels. See, e.g., Roe, Busemeyer, and Townsend (2001), Anderson, Goeree, and Holt (2004), McMillen and Holmes (2006), Bogacz, Usher, Zhang, and McClelland (2007), Ditterich (2010), and Krajbich and Rangel (2011).

Alternatively, Reutskaja, Nagel, Camerer, and Rangel (2011) propose three two-stage models in which subjects randomly search through the feasible set during an initial search phase, and when this phase is concluded they select the best item that was encountered during the search, up to some noise. This approach can be called **quasi-exhaustive search** in that the presence of a deadline may terminate the search phase before all alternatives have been evaluated and introduces an error probability.

Here, instead, we focus on **sequential pairwise comparison** as advocated by Russo and Rosen (1975) in a seminal eye fixation study (see especially their concluding section). Although different from the models considered by Krajbich and Rangel (2011) and Reutskaja, Nagel, Camerer, and Rangel (2011), our model is consistent with some of their experimental findings about the menu-exploration process and shares the classical choice theory approach according to which multialternative choice relies on binary comparison and elimination. The computational model we adopt is extended and numerically studied in Baldassi, Cerreia-Vioglio, Maccheroni, and Marinacci (2017), who also axiomatize the binary DDM.

1.3 Paper outline

The paper is organized as follows: Section 2 presents the mathematical setup and a short review of the classical Luce Model. In Section 3, we axiomatically characterize Multinomial Logit Processes as well as some special cases (for example, the classical Multinomial Logit Model with affine u). Section 4 shows how softmax distributions emerge from a neuropsychologically inspired decision procedure that combines Markovian search of the menu (*à la* Metropolis et al., 1953) and DDM binary comparisons of its alternatives. The final Section 5 concludes, and proofs are relegated to the appendix. The latter also contains a characterization of general Heteroscedastic Multinomial Logit Models and of the special case in which $T = (0, \infty)$ and λ is increasing and bijective.

2 Random choice rules

Let $\mathcal{A}$ be the collection of all non-empty finite subsets A of a universal set X of possible alternatives. The elements of $\mathcal{A}$ are called choice sets (or menus, or choice problems). We denote by $\Delta(X)$ the set of all finitely supported probability measures on X and, for each $A \subseteq X$, by $\Delta(A)$ the subset of $\Delta(X)$ consisting of the measures assigning mass 1 to A.

Definition 1 *A* random choice rule *is a function*

$$\begin{array}{rccc} p: & \mathcal{A} & \to & \Delta(X) \\ & A & \mapsto & p_A \end{array}$$

such that $p_A \in \Delta(A)$ *for all* $A \in \mathcal{A}$.

Given any alternative $a \in A$, we interpret $p_A(\{a\})$, also denoted by $p(a, A)$, as the probability that an agent chooses a when the set of available alternatives is A. More generally, if B is a subset of A, we denote by $p_A(B)$ or $p(B, A)$ the probability that the selected element lies in B.[9] This probability can be viewed as the frequency with which an element in B is chosen. As usual, given any a and b in X, we set

$$p(a,b) = p(a,\{a,b\}), \quad r(a,b) = \frac{p(a,b)}{p(b,a)}, \quad \ell(a,b) = \ln r(a,b) \tag{1}$$

So, $r(a,b)$ denotes the *odds* for a against b and $\ell(a,b)$ the *log-odds*. The advantage of using log-odds for a against b is that they are strictly positive if and only if odds are favorable to a. Indeed,

$$p(a,b) > p(b,a) \iff r(a,b) > 1 \iff \ell(a,b) > 0 \tag{2}$$

Starting with the seminal Davidson and Marschak (1959), any of the above three equivalent relations has been viewed as revealing the *stochastic preference* for a over b, denoted $a \succ_p b$.

Luce (1959) proposes the most classical random choice model. Its assumptions on p are:

Positivity $p(a,b) > 0$ *for all* $a, b \in X$.

Choice Axiom $p(a,A) = p(a,B)\,p(B,A)$ *for all* $B \subseteq A$ *in* $\mathcal{A}$ *and all* $a \in B$.

The latter axiom says that the probability of choosing an alternative a from menu A is the probability of first selecting B from A, then choosing a from B (provided a belongs to B). As observed by Luce, formally this assumption corresponds to the fact that $\{p_A : A \in \mathcal{A}\}$ is a conditional probability system in the sense of Renyi (1956).[10] Remarkably, Luce's Choice Axiom is also equivalent to:

Independence from Irrelevant Alternatives

$$\frac{p(a,b)}{p(b,a)} = \frac{p(a,A)}{p(b,A)} \tag{IIA}$$

for all $A \in \mathcal{A}$ *and all* $a, b \in A$ *such that* $p(a,A)/p(b,A)$ *is well defined.*[11]

This axiom says that the likelihood ratio for a against b is independent of the available alternatives that are different from a and b themselves.

Theorem 1 (Luce) *A random choice rule* $p : \mathcal{A} \to \Delta(X)$ *satisfies Positivity and the Choice Axiom if and only if there exists* $v : X \to \mathbb{R}$ *such that*

$$p(a,A) = \frac{e^{v(a)}}{\sum_{b \in A} e^{v(b)}} \tag{LM}$$

for all $A \in \mathcal{A}$ *and all* $a \in A$.

In this case, v *is unique up to an additive constant.*

[9] Formally, $x \mapsto p(x,A)$ for all $x \in X$ is the discrete density of p_A, but notation will be abused and $p_A(\cdot)$ identified with $p(\cdot,A)$.

[10] See Lemma 2 of Luce (1959) and Lemma 13 in the appendix.

[11] That is, different from $0/0$. See Lemma 3 of Luce (1959) for the case in which Positivity holds and our Lemma 13 in Appendix A.1 for the general case.

The function v is called *(Luce-McFadden) value.* Since

$$\ell(a,b) = v(a) - v(b) \tag{3}$$

the value v represents stochastic preferences, that is,

$$a \succ_p b \iff v(a) > v(b) \tag{4}$$

This simple observation is at the core of the analysis of discrete choice experiments, as far as, preference estimation is concerned.[12]

Theorem 1 also shows that, under the Choice Axiom, Positivity is equivalent to the stronger assumption that p_A has full support for all $A \in \mathcal{A}$.[13]

Full Support $p(a,A) > 0$ *for all* $A \in \mathcal{A}$ *and all* $a \in A$.

3 Random choice processes

Let $\varnothing \neq T \subseteq (0,\infty)$ be a – discrete or continuous – set of points of time.

Definition 2 *A* random choice process *is a collection* $\{p_t\}_{t\in T}$ *of random choice rules.*

For each t, we interpret $p_t(a,A)$ as the probability that an agent chooses alternative a from menu A if t is the maximum amount of time he is given to decide.[14] For this reason, we call t *deliberation time.*

N.B. *Throughout the paper, deliberation time is exogenously given and fixed.*

An important alternative interpretation of t, especially when T is discrete and panel data are considered, is the number of times that the agent has been facing choice problem A, called *experience level* by McKelvey and Palfrey (1995).

We maintain the following assumptions:

Positivity p_t *satisfies Positivity for all* $t \in T$.

Choice Axiom p_t *satisfies the Choice Axiom for all* $t \in T$.

By the Luce Theorem, for each $t \in T$ there exists a value $v_t : X \to \mathbb{R}$ such that

$$p_t(a,A) = \frac{e^{v_t(a)}}{\sum_{b\in A} e^{v_t(b)}} \qquad \forall a \in A \in \mathcal{A}$$

Each of the v_t's is unique up to an additive constant, so that time-dependent value differences are meaningful and captured by log-odds.

[12] See, e.g., Louviere, Hensher, and Swait (2000, p. 113).

[13] The *support* of p_A is $\operatorname{supp} p_A = \{a \in A : p(a,A) > 0\}$. In Cerreia-Vioglio, Maccheroni, Marinacci, and Rustichini (2018), we relax the Full Support assumption and characterize general random choice rules in terms of "optimality "of their support (see Theorem 14 in Appendix A.1).

[14] E.g., by an experimenter, by a script, or by a spouse. See Agranov, Caplin, and Tergiman (2015) for a simple protocol that allows to observe these probabilities for human agents.

3.1 Multinomial Logit Processes

The problem now is to characterize the separation of the value $v_t(x)$ in a **time-dependent** accuracy component $\lambda(t)$ and a **time-independent** utility component $u(x)$.

Our first novel assumption requires that favorable odds for a against b remain favorable. Formally:

Weak Consistency *Given any $s > t$ in T,*

$$p_t(a,b) > p_t(b,a) \implies p_s(a,b) > p_s(b,a) \tag{5}$$

for all $a, b \in X$.

In terms of stochastic preferences, this is equivalent to

$$a \succ_t b \implies a \succ_s b \tag{6}$$

for all $a, b \in X$ and all $s > t$ in T. Thus, Weak Consistency means that stochastic preferences are weakly stable: as t increases, stochastic preference is not reverted. This feature is coherent with the idea that correct-on-average, yet noisy, evidence is accumulating to inform the decision maker's choice between the two alternatives (see also Section 4). Moreover, by (4), Weak Consistency of stochastic preference yields weak consistency of values:

$$v_t(a) > v_t(b) \implies v_s(a) > v_s(b)$$

The next and final assumption, instead, requires – in view of 3 – that relative values be time invariant.

Log-odds Ratio Invariance *Given any $s > t$ in T,*

$$\frac{\ell_t(a,c)}{\ell_t(b,c)} = \frac{\ell_s(a,c)}{\ell_s(b,c)}$$

for all $a, b, c \in X$ such that either ratio is well defined.

We are ready to state our first representation result.

Theorem 2 *A random choice process $\{p_t\}_{t\in T}$ satisfies Positivity, the Choice Axiom, Weak Consistency, and Log-odds Ratio Invariance if and only if there exist $u : X \to \mathbb{R}$ and $\lambda : T \to (0,\infty)$ such that*

$$p_t(a,A) = \frac{e^{\lambda(t)u(a)}}{\sum_{b\in A} e^{\lambda(t)u(b)}} \tag{MLP}$$

for all $A \in \mathcal{A}$, all $a \in A$, and all $t \in T$.

In this case, u is cardinally unique, and λ is unique given u unless the latter is constant.

In what follows, we call a random choice process that satisfies (MLP) a *Multinomial Logit Process with utility u and accuracy λ*.

As anticipated in the introduction, this theorem provides a preferential foundation for the multinomial logit representation of choice behavior under time pressure (when t represents deliberation time) and for Quantal Response Equilibrium theory (when t represents experience level). Such a foundation clarifies the conceptual underpinnings of these theories and makes them falsifiable. Moreover, together with the forthcoming Proposition 4, Theorem 2 presents a theoretical framework for the multinomial logit analysis of discrete choice experiments in which multiple decisions are taken. In these experiments, both u and λ are estimated and the monotonicity behavior of the latter is interpreted in terms of preference learning (increasing λ – decreasing error variance) and fatigue (decreasing λ – increasing error variance).[15] Moreover, the independence of u from t is interpreted as stability of the "true preferences".[16]

Our axioms can thus be viewed as testable implications of this popular (heteroscedastic) logit model, while our identification result below (Proposition 4) says that, if the model is not misspecified (axioms are not violated), then its parameters are easily retrievable from choice frequencies.

The last part of the statement of Theorem 2 guarantees that accuracy is determined only if utility is not constant. The next simple proposition shows that utility is constant if and only if choice is completely random (irrespective of deliberation time) if and only if stochastic preference is trivial (again, irrespective of deliberation time).

Proposition 3 *Let $\{p_t\}_{t\in T}$ be a Multinomial Logit Process with utility u and accuracy λ. The following conditions are equivalent:*

1. *u is constant;*
2. *$p_t(a, A) = 1/|A|$ for all $A \in \mathcal{A}$, all $a \in A$, and all $t \in T$;*
3. *there are no $a, b \in X$ and $t \in T$ such that $\ell_t(a, b) > 0$;*
4. *$\ell_t(a, b) = 0$ for all $a, b \in X$ and all $t \in T$.*

The importance of point 3 is that, contrapositively, if there are $\hat{a}, \hat{b} \in X$ and $\hat{t} \in T$ such that $\ell_{\hat{t}}(\hat{a}, \hat{b}) > 0$, then u is non-constant and λ unique (up to a strictly positive multiplicative constant). The next proposition shows that identification is straightforward.[17]

Proposition 4 *Let $\{p_t\}_{t\in T}$ be a Multinomial Logit Process for which there exist $\hat{a}, \hat{b} \in X$ and $\hat{t} \in T$ such that $\ell_{\hat{t}}(\hat{a}, \hat{b}) > 0$. Then setting*

$$\begin{cases} \hat{\lambda}(t) = \ell_t(\hat{a}, \hat{b}) & \forall t \in T \\ \hat{u}(x) = \dfrac{\ell_{\hat{t}}(x, \hat{b})}{\ell_{\hat{t}}(\hat{a}, \hat{b})} & \forall x \in X \end{cases} \tag{7}$$

[15] See Savage and Waldman (2008) and Campbell, Boeri, Doherty, and Hutchinson (2015).

[16] See, e.g., Hensher, Louviere, and Swait (1999).

[17] Estimation of u and λ is standard, typically carried out by maximum likelihood. See, e.g., Ben-Akiva and Lerman (1985) on the econometric side and McKelvey and Palfrey (1995) on the game theoretic one.

the functions $\hat{\lambda} : T \to (0, \infty)$ and $\hat{u} : X \to \mathbb{R}$ are well defined and

$$p_t(a, A) = \frac{e^{\hat{\lambda}(t)\hat{u}(a)}}{\sum_{b \in A} e^{\hat{\lambda}(t)\hat{u}(b)}} \tag{$\widehat{\text{MLP}}$}$$

for all $A \in \mathcal{A}$, all $a \in A$, and all $t \in T$.

This result has two very important empirical implications. First, it allows a novel econometric test of our theory: Expression (7) says that for any two distinct pairs $(\hat{a}, \hat{b})$ and $(\tilde{a}, \tilde{b})$ such that for some $\hat{t}$ and $\tilde{t}$ we have both $p_{\hat{t}}(\hat{a}, \hat{b}) > p_{\hat{t}}(\hat{b}, \hat{a})$ and $p_{\tilde{t}}(\tilde{a}, \tilde{b}) > p_{\tilde{t}}(\tilde{b}, \tilde{a})$, it must be the case that the log-likelihood ratios $\hat{\lambda}(t)$ and $\tilde{\lambda}(t)$ are collinear (testing collinearity is one of the most classical exercises in econometrics).[18] Second, it allows to test the properties that the economic literature makes on λ and u. The obvious example is the monotonicity of λ, which we indeed characterize with the next axiom. In the next section we will also characterize the linearity of both λ and u.

Weak Discovered Preference Axiom *Given any $s > t$ in T,*

$$p_t(a, b) > p_t(b, a) \implies p_s(a, b) \geq p_t(a, b) \tag{8}$$

for all $a, b \in X$.

The right hand side is where (8) strengthens (5) by requiring that favorable odds for a against b become **more** favorable, rather than just remaining favorable, as deliberation time increases.[19] This provides a possible formalization of the DPH and captures the intuition that evidence accumulation reduces the possibility of mistakes. Specifically, if a is stochastically preferred to b given time t to decide, the probability of a mistake is the probability of choosing b from $\{a, b\}$, that is, $p_t(b, a)$. By Weak Consistency, a is stochastically preferred to b also when the time s to decide is greater than t; in this case, the probability of a mistake is $p_s(b, a)$. The Weak Discovered Preference Axiom requires that $p_s(b, a) \leq p_t(b, a)$, that is, error rates decrease over time.

Another compelling formalization of the DPH is as stochastic dominance of payoffs as time increases. Consider a Multinomial Logit Process with utility u and an arbitrarily fixed $A \in \mathcal{A}$. For each $t \in T$, by selecting alternatives according to $p_t(\cdot, A)$, the agent's payoffs (in utils) are described by the random variable

$$\begin{array}{rccc} U_A(t): & A & \to & \mathbb{R} \\ & a & \mapsto & u(a) \end{array}$$

Given any $s > t$ in T, the random variable $U_A(s)$ *stochastically dominates* the random variable $U_A(t)$ if and only if

$$p_s(\{a \in A : u(a) \geq h\}, A) \geq p_t(\{a \in A : u(a) \geq h\}, A) \qquad \forall h \in \mathbb{R} \tag{9}$$

That is, for any given h, the probability of obtaining a payoff greater than h after deliberating for s units of time is greater than after deliberating for t units of time (when choice from A is considered).

[18] For more on the testing of MLP, see Hensher, Louviere, and Swait (1999).

[19] Formally, $p_t(a, b) > p_t(b, a) \iff r_t(a, b) > 1$ and $p_s(a, b) \geq p_t(a, b) \iff r_s(a, b) \geq r_t(a, b)$.

Proposition 5 *Let $\{p_t\}_{t\in T}$ be a Multinomial Logit Process with utility u and accuracy λ. The following conditions are equivalent:*

1. *$\{p_t\}_{t\in T}$ satisfies the Weak Discovered Preference Axiom;*

2. *$U_A(s)$ stochastically dominates $U_A(t)$ for all $A \in \mathcal{A}$ and all $s > t$ in T;*

3. *λ is (weakly) increasing on T.*[20]

In a dual version of this result, increasing error rates are easily seen to correspond stochastically dominated payoffs (as time increases) and to decreasing λ's.

3.2 The Multinomial Logit Model

Let $T = (0, \infty)$. The characterization of Multinomial Logit Processes with linear accuracy, that is,

$$p_t(a, A) = \frac{e^{tu(a)}}{\sum_{b\in A} e^{tu(b)}} \qquad \forall a \in A \in \mathcal{A} \quad \forall t \in (0, \infty) \tag{MNL}$$

provides an axiomatic foundation to the classic Multinomial Logit Model, the interest of which goes well beyond the preference discovery point of view of this paper.

This can be achieved in two ways. First, by adding to the axioms of Luce the following linearity condition:

Log-odds Linearity *Given any $s > t$ in $(0, \infty)$,*

$$\frac{\ell_t(a, b)}{\ell_s(a, b)} = \frac{t}{s}$$

for all $a, b \in X$ such that $\ell_t(a, b) / \ell_s(a, b)$ is well defined.

This assumption has clear empirical content; yet, its conceptual significance is somewhat limited. For this reason, we investigate a second way of characterizing the MNL that bears on its interpretation as a theory of optimization subject to disturbances the variance of which vanishes as scale diverges (see Ben-Akiva and Lerman, 1985, Chapter 5).

Continuity $\lim_{t\to s} p_t(a, A)$ *exists for all $A \in \mathcal{A}$, all $a \in A$, all $s \in [0, \infty]$, and it coincides with $p_s(a, A)$ if $s \in (0, \infty)$.*

Continuity guarantees that, as t tends to either 0 or ∞, two limit random choice rules p_0 and p_∞ are defined that extend the domain of the random choice process $\{p_t\}$ to $[0, \infty]$. These limit rules satisfy the Choice Axiom if $\{p_t\}$ does, but they do not necessarily inherit the Positivity property. They have a simple interpretation: p_0 describes choice when no time is available to deliberate (or the impulsive irrational agent of Becker, 1962), while p_∞ describes time unconstrained choice (or the neoclassical rational agent).

[20] If u is constant, then λ is undetermined and this condition must be considered vacuous.

Consistency *Given any* t *in* $(0,\infty)$,

$$p_t(a,b) > p_t(b,a) \implies p_\infty(a,b) > p_\infty(b,a)$$

for all $a, b \in X$.

The interpretation of Consistency is analogous to that of Weak Consistency.

Asymptotic Uniformity

$$p_\infty(a,b) \neq 0,1 \implies p_\infty(a,b) = 1/2$$

for all $a, b \in X$.

Asymptotic Uniformity postulates that, if the decision maker is unable to make up his mind between alternatives a and b irrespectively of the time available to do so, then he will choose by flipping a fair coin. This is a classical notion of indifference that can be traced back to the early days of experimental economics (see again Davidson and Marschak, 1959, who in turn attribute it to Mosteller and Nogee, 1951).

Boundedness

$$\sup_{t,s\in(0,\infty)} |r_{t+s}(a,b) - r_t(a,b)\, r_s(a,b)| < \infty$$

for all $a, b \in X$.

Boundedness is a "grain of exponentiality" that requires the time combination of odds not to be "infinitely far" from additive. This interpretation of Boundedness bears on the neural analogy of Gold and Shadlen (2002, 2007) between the comparison of a and b and the testing of the "hypothesis" that a be preferable to b, on the one hand, and on the analogy between deliberation and the gathering of evidence about these hypotheses (see also Shadlen and Shohamy, 2016). Under this interpretation, $r_{t+s}(a,b)$ are the odds for a against b resulting from $t+s$ time units of continuous deliberation, while $r_t(a,b)\, r_s(a,b)$ are those that would result by combining t and s time units of deliberation. Boundedness requires that these two ways of combining evidence do not diverge from one another.[21]

[21] See Frederic, Di Bacco, and Lad (2012) for a formal interpretation of the product of odds in terms of combination of evidence. In Appendix A.2, we show that replacing, in the Boundedness axiom, the sum $t+s$ with the superadditive concatenation $t \oplus s = t + s + ts$ corresponds to the logarithmic case $\lambda(t) = \ln(1+t)$. We also show that a generic time concatenation $t \oplus s$ (in the sense of measurement theory; see, Luce and Suppes, 2002) corresponds to the general case of an increasing and bijective λ.

Theorem 6 *Given a random choice process $\{p_t\}_{t\in(0,\infty)}$, the following conditions are equivalent:*

1. *$\{p_t\}_{t\in(0,\infty)}$ satisfies Positivity, the Choice Axiom, and Log-odds Linearity;*
2. *$\{p_t\}_{t\in(0,\infty)}$ satisfies Positivity, the Choice Axiom, Continuity, Consistency, Asymptotic Uniformity, and Boundedness;*
3. *there exists $u : X \to \mathbb{R}$ such that*

$$p_t(a, A) = \frac{e^{tu(a)}}{\sum_{b\in A} e^{tu(b)}} \tag{MNL}$$

for all $A \in \mathcal{A}$, all $a \in A$, and all $t \in (0,\infty)$.

In this case, u is unique up to an additive constant.

For obvious reasons, we call a process $\{p_t\}$ that satisfies (MNL) a *Multinomial Logit Model with utility u*. This model is the workhorse of Discrete Choice Analysis, where it is typically also assumed that $X \subseteq \mathbb{R}^n$ and $u(x) = \beta \cdot x$ for some $\beta \in \mathbb{R}^n$. The following proposition characterizes this fundamental special case.[22]

Proposition 7 *Let X be a convex set and $\{p_t\}_{t\in(0,\infty)}$ be a Multinomial Logit Model with utility u. The following conditions are equivalent:*

1. *u is affine;*
2. *there exists $t \in (0,\infty)$ such that*

$$p_t(a, b) = p_{\frac{t}{\alpha}}(\alpha a + (1-\alpha) b, b)$$

for all $a, b \in X$ and all $\alpha \in (0,1)$;

3. *given any $t \in (0,\infty)$ and any $c \in X$,*

$$p_t(a, A) = p_{\frac{t}{\alpha}}(\alpha a + (1-\alpha) c, \alpha A + (1-\alpha) c)$$

for all $A \in \mathcal{A}$, all $a \in A$, and all $\alpha \in (0,1)$.

In Discrete Choice Analysis, the elements of X are often viewed as vectors of attributes.[23] The equivalence between points 1 and 3 above thus shows that the affinity of u corresponds to an inverse relation between the proximity of attribute levels and the degree of choice accuracy. To fix ideas, assume $A = \{a, b\}$ with $u(a) > u(b)$. Shrinkage by a factor $\alpha = 1/2$ of attribute levels in the direction of c then doubles the time to achieve the same probability of choosing the optimal alternative.

[22] It also characterizes the case in which X is a set of mixed actions and $u(x)$ is the expected payoff of x.

[23] For example, in a typical travel demand application the components x_0 and x_1 of x represent travel time and travel cost of alternative x, respectively (Ben-Akiva and Lerman, 1985, and Train, 2009).

4 The Metropolis-DDM algorithm

So far, we regarded the components p_t of a random choice process $\{p_t\}$ as the output of a black box: our axioms characterize Multinomial Logit Processes, but remain silent about what decision procedure may generate the corresponding choice probabilities. In this section, we address this issue by combining Markovian search of the menu and DDM pairwise comparisons of its alternatives. Both assumptions, inspired by the seminal eye-tracking study of Russo and Rosen (1975), find support in recent theories about memory (see, e.g., Luck and Vogel, 1997, Vogel and Machizawa, 2004, and Shadlen and Shohamy, 2016) as well as in recent eye-tracking experiments on multialternative choice (see, e.g., Krajbich and Rangel, 2011, and Reutskaja, Nagel, Camerer, and Rangel, 2011).

Specifically, we consider a decision maker who has to choose an optimal alternative from a finite menu A within an **exogenously controlled deliberation time** t. For extra clarity, t is the time given to the agent to think through a single decision episode involving the choice problem, so it is, the moment in which the deliberation process is externally terminated and a choice must be made.[24]

In what follows, we first introduce the different parts of the decision procedure and we then assemble them. Notation is eased by assuming, without loss of generality, that $A = \{1, 2, ..., |A|\}$, with $|A| \geq 2$, and by identifying elements μ of $\Delta(A)$ with vectors in $\mathbb{R}^{|A|}$.

4.1 Exploration

Exploration of menu A has a classic Metropolis et al. (1953) format. The agent starts with a first (automatically accepted) candidate solution b drawn from an initial distribution $\mu \in \Delta(A)$. Then, given an incumbent solution b, the agent considers an alternative candidate solution $a \neq b$ with probability $Q(a \mid b)$. The only requirements we make on the probability transition matrix Q are irreducibility and symmetry (see, e.g., Madras, 2002). The independence of μ and Q from the utility function u are the formal counterparts of the aforementioned eye-tracking evidence (see also Subsection 4.4.3 below).

A natural assumption that simultaneously guarantees both irreducibility and symmetry of Q is that the subjective distance which the decision maker perceives between alternatives be described by a *semimetric* d on A, and that $Q(a \mid b)$ be a strictly positive function of $d(a, b)$.[25] For example, d can be the discrete metric if A is a set of abstract alternatives, the Euclidean metric if A is a set of multiattribute alternatives (like transportation modes), the shortest-path distance if A is a connected graph (like a wine rack or a vending machine). In particular, one may consider a perceptual[26] semimetric and the following parametric form

$$Q(a \mid b) = \frac{1}{|A| - 1} \frac{1}{d(a, b)^{\rho}} \qquad \forall a \neq b \text{ in } A$$

[24] The interpretation of t as experience level does not apply to this section.

[25] A *semimetric* has all the properties of a metric except the triangle inequality. It can represent either a physical distance affecting attention or a psychological distance describing the mental landscape in which the decision maker organizes the elements of A (possibly incorporating similarity considerations), or a mix of the two. See Russo and Rosen (1975) and Roe, Busemeyer, and Townsend (2001).

[26] A semimetric is *perceptual* if the minimum distance between two distinct alternatives is 1.

where $\rho \in (0,\infty)$ is an exploration aversion parameter: for large ρ only the nearest neighbours of the incumbent solution are considered, while for small ρ exploration is essentially uniform across alternatives.

4.2 Binary comparison

Once proposed, alternative a is compared with incumbent b via the Drift Diffusion Model of Ratcliff (1978). According to this model, an alternative is selected as soon as the net evidence in its favor reaches a posited decision threshold $z \in (0,\infty)$, which in our case depends on the time constraint t via $z = \lambda(t)$.[27] Specifically, the comparison of a and b is believed to activate two neuronal populations whose activities (firing rates) provide evidence for the two alternatives.[28] If the mean activities $u(a)$ and $u(b)$ of these populations experience instantaneous and independent white noise fluctuations, then evidence accumulation in favor of a and b is represented by two uncorrelated Brownian motions with drift

$$V_a(\tau) = u(a)\tau + W_a(\tau) \qquad \text{and} \qquad V_b(\tau) = u(b)\tau + W_b(\tau)$$

defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. With this,

- the *net evidence* in favor of a against b is given by the difference

$$Z_{a,b}(\tau,\omega) = [u(a) - u(b)]\tau + \sqrt{2}W(\tau,\omega) \qquad \forall (\tau,\omega) \in (0,\infty)\times\Omega \tag{DDM}$$

  where W is the Wiener process $(W_a - W_b)/\sqrt{2}$;

- the comparison ends when $Z_{a,b}(\tau)$ reaches either the threshold z or $-z$; so the *response time* is the random variable

$$\mathrm{RT}^z_{a,b}(\omega) = \min\{\tau \in (0,\infty) : |Z_{a,b}(\tau,\omega)| = z\} \qquad \forall \omega \in \Omega$$

- at which time, if the upper bound z has been reached, the agent accepts proposal a; otherwise, the lower bound $-z$ has been reached, proposal a is rejected and the agent maintains incumbent b; so, the *comparison outcome* is the random variable

$$\mathrm{CO}^z_{a,b}(\omega) = \begin{cases} a & \text{if } Z_{a,b}\left(\mathrm{RT}^z_{a,b}(\omega),\omega\right) = z \\ b & \text{if } Z_{a,b}\left(\mathrm{RT}^z_{a,b}(\omega),\omega\right) = -z \end{cases}$$

In a Discovered Preference Hypothesis perspective, it is important to observe that the DDM does not assume $u(a)$ and $u(b)$ to be known to the agent. Indeed, the sign of the utility difference $u(a) - u(b)$ is discovered by accumulating (noisy) evidence from either the external environment or from memory until the threshold z is reached — see Shadlen and Shohamy (2016) and Fudenberg, Strack, and Strzalecki (2017).

[27] See Milosavljevic et al. (2010) and Karsilar, Simen, Papadakis, and Balci (2014, especially p. 14) for a discussion of the threshold-reducing effects of time pressure.

[28] See Bogacz et al. (2006) and Shadlen and Shohamy (2016) for the neurophysiological and neuropsychological analyses of this mechanism.

4.3 Decision

We now combine Metropolis exploration and DDM pairwise comparisons. The resulting procedure describes an agent who, given time t to decide, automatically adjusts $z = \lambda(t)$, explores the menu according to Q, and compares alternatives according to CO^z. His search continues until time t is reached, at which point he chooses the incumbent solution.

Metropolis-DDM Algorithm

Input: *Given* $t \in (0, \infty)$ *and* $z = \lambda(t) \in (0, \infty)$ *.*

Start: *Draw* a_0 *from* A *according to* μ *and*

- *set* $\tau_0 = 0$*,*
- *set* $b_0 = a_0$*;*

Repeat: *Draw* a_{n+1} *from* A *according to* $Q(\cdot \mid b_n)$ *and compare it to* b_n *via the DDM, so:*

- *set* $\tau_{n+1} = \tau_n + \mathrm{RT}^z_{a_{n+1},b_n}$*,*
- *set* $b_{n+1} = \mathrm{CO}^z_{a_{n+1},b_n}$*;*

until $\tau_{n+1} > t$*.*

Stop: *Set* $b^* = b_n$*.*

Output: *Choose* b^* *from* A*.*

Since, at each iteration of the "repeat-until" loop, the evaluation of the sign of the utility difference $u(a) - u(b)$ is performed according to the DDM, the algorithm we propose preserves the interpretation in terms of preference discovery that we discussed at the end of last section. In particular, after comparing incumbent b with proposal a and selecting $\mathrm{CO}^z_{a,b}$ as the new incumbent, the agent has not learned $u(a)$ and $u(b)$, but rather has performed a "test" of the "hypothesis" that a be preferable to b (see Gold and Shadlen, 2002, 2007).

The fact that this test is time-consuming and subject to error represents the main difference between the Metropolis-DDM algorithm and the standard brute force comparison-and-elimination algorithm of classical optimization (termed *standard revision* in marketing). According to standard revision, multiple alternatives are compared in a pairwise fashion and one alternative is permanently eliminated after each binary comparison; thus, after $|A| - 1$ of these comparisons the incumbent solution is an optimal choice. The implicit assumption which this brute-force procedure rests upon is that the pairwise comparisons are instantaneous and exact. In the time-constrained Metropolis-DDM algorithm, instead, the fact that comparisons are time consuming may lead to incomplete exploration of the menu, while the fact that comparisons may be erroneous makes it inadvisable to eliminate permanently an alternative that was judged inferior at a previous stage.

In particular, the adoption of a Drift Diffusion Model with threshold $z = \lambda(t)$ entails a mean pairwise comparison time of

$$\frac{\lambda(t)}{|u(a) - u(b)|} \tanh\left(\frac{|u(a) - u(b)|\,\lambda(t)}{2}\right) \tag{10}$$

and an error probability of

$$\frac{1}{1 + e^{\lambda(t)|u(a)-u(b)|}} \tag{11}$$

So, large utility differences imply fast and accurate comparisons, while increasing the accuracy threshold slows down comparisons, thus diminishing the probability of exploring the entire menu.

4.4 Bringing theory to data

In this section we first show that the limit distribution of choice generated by the Metropolis-DDM algorithm is softmax. This fact has both theoretical and empirical implications. On the one hand, it puts forth the algorithm as a plausible internal cause of the softmaximizing behavior that we characterized in the first part of the paper. On the other, it presents a first test for the algorithm itself. We then strengthen this test by an estimate of the number of iterations performed by the algorithm within time t (the larger this number the closer the final distribution is to softmax) and by numerical simulations of the algorithm with biological parameters.

Subsequently, we compare the predictions of the Metropolis-DDM and of the standard Multi-Alternative-DDM algorithm and we find that, while the former predicts Multinomial Logit choice probabilities, the latter predicts Multinomial Probit ones. This is another test for the Metropolis-DDM algorithm.

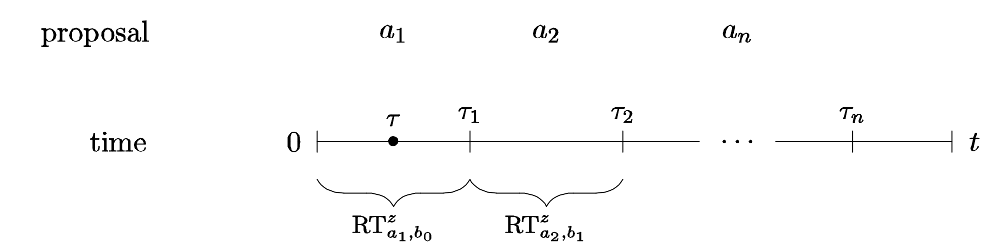
A third and last test is presented by approximating the eye-fixation frequencies implied by the Metropolis-DDM algorithm and by matching them with some recent empirical data of Krajbich and Rangel (2011).

4.4.1 Stationarity and simulations

The Metropolis-DDM algorithm randomly produces a sequence

$$(b_0, a_1, \tau_1, b_1, ..., b_n, a_{n+1}, \tau_{n+1}, b_{n+1}, ...)$$

of incumbents b_n, proposals a_{n+1}, and elapsed times τ_{n+1},



which is truncated at the chosen alternative $b^* = b_n$ by the stopping rule $\tau_{n+1} > t$.

At each iteration of the "repeat-until" loop, the proposal a is accepted as the new incumbent with probability

$$\alpha^{\lambda(t)}(a \mid b) = \mathbb{P}\left(\omega \in \Omega : \mathrm{CO}_{a,b}^{\lambda(t)}(\omega) = a\right)$$

while a is rejected and the old incumbent b is maintained with the complementary probability $1 - \alpha^{\lambda(t)}(a \mid b)$. Therefore, the resulting probability of selecting a as a new incumbent given old incumbent b is

$$P_t(a \mid b) = Q(a \mid b)\,\alpha^{\lambda(t)}(a \mid b) \qquad \forall a \neq b \text{ in } A$$

This transition probability combines the stochasticity of the (Metropolis) proposal mechanism and that of the (DDM) acceptance/rejection rule. Specifically, after n iterations of the "repeat-until" loop, the probability that $b^* = a$ is the a-th component of the vector $P_t^n \mu$ in $\Delta(A)$.[29]

Proposition 8 *Let $u \in \mathbb{R}^A$, $t \in (0, \infty)$ and $\lambda(t) \in (0, \infty)$. If Q is irreducible and symmetric, then P_t is aperiodic, irreducible, and its stationary distribution is the softmax*

$$p_t(a, A) = \frac{e^{\lambda(t)u(a)}}{\sum_{b \in A} e^{\lambda(t)u(b)}} \qquad \forall a \in A \tag{softmax}$$

In particular, $\lim_{n\to\infty} P_t^n \mu = p_t(\cdot, A)$ for all $\mu \in \Delta(A)$.

This result completes the interpretation of Multinomial Logit Processes as a theory of the Discovered Preference Hypothesis by presenting softmax as the ideal behavior of a decision maker who follows the Metropolis-DDM algorithm.

On the other hand, the behavioral axioms of the first part of the paper present an "asymptotic" test for this algorithm, in that softmax is the output of the algorithm provided it "converges" within deliberation time t. The degree of convergence, in turn, depends on the stochastic number N_t of iterations performed by the algorithm within time t.

The next result delivers an expected lower bound for this number and paves the way for our convergence result (Proposition 10 below).

Proposition 9 *Let $u \in \mathbb{R}^A$, $t \in (0, \infty)$ and $\lambda(t) \in (0, \infty)$. Then*

$$\mathbb{E}(N_t) \geq \frac{2t}{\lambda^2(t)} - 1 \tag{12}$$

In particular, $\lambda(t) = o\left(\sqrt{t}\right)$ implies that $\mathbb{E}(N_t) \to \infty$ as $t \to \infty$.

Qualitatively, this result says that, for the algorithm to converge, accuracy $\lambda(t)$ must be an increasing and concave function of deliberation time t. Quantitatively, it permits to prove convergence for logarithmic λ's – like the ones found by Ortega and Stocker (2016). This is proved in the next proposition.[30]

[29] Here $P_t = [P_t(a \mid b)]_{a,b\in A}$ is regarded as the probability transition matrix of a Markov chain with initial distribution μ, and P_t^n as its n-th power.

[30] Recall that a matrix is strictly quasipositive if and only if all of its off-diagonal terms are strictly positive (clearly such a matrix is irreducible).

Proposition 10 *Let $u \in \mathbb{R}^A$ and Q be symmetric and strictly quasipositive. If $\lambda : (0, \infty) \to (0, \infty)$ diverges at ∞ and is such that*

$$\limsup_{t \to \infty} \frac{\lambda(t)}{\ln(t)} < \frac{1}{\max_{a,b \in A} |u(a) - u(b)|}$$

then, as $t \to \infty$,

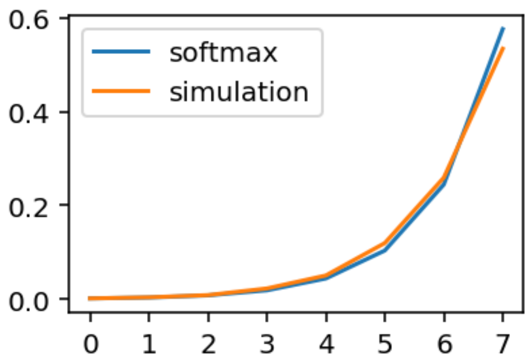
$$\exp \mathbb{E} \left(\ln \frac{1}{2} \sum_{a \in A} \left| \left[P_t^{N_t} \mu \right](a) - p_A^{(u,\lambda(t))}(a) \right| \right) \to 0$$

That is, the total variation distance between $P_t^{N_t}\mu$ and p_A vanishes in geometric mean.

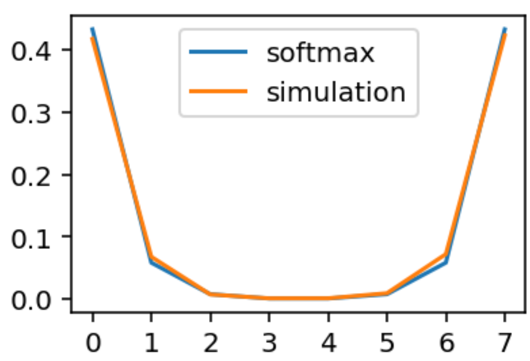
Finally, in terms of experimental testability, a **key observation** is that the parameters u and $\lambda(t)$ which describe the stationary distribution (softmax) and appear in expression (12) are the **same** ones that govern binary DDM comparisons in our algorithm. The latter parameters can be easily estimated by eye-tracking techniques (see, e.g., Milosavljevic et al., 2010). So, Propositions 8 and 9 become powerful tests for the Metropolis-DDM algorithm and its convergence.

For example, the estimates of Milosavljevic et al. (2010) on binary DDM comparisons, under both high and low time pressure, correspond to $\max_{a,b \in A} |u(a) - u(b)| \approx 7.071$, with $\lambda(t) \approx 0.849$ for high time pressure and $\lambda(t) \approx 1.442$ for low time pressure. In the plots below we report the output of some simulations with $A = \{0, 1, ..., 7\}$.

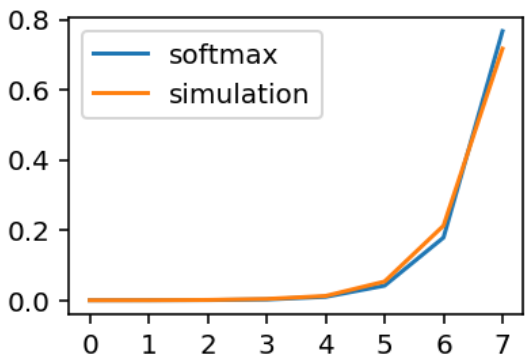
Choose in 4 seconds with $u(x) \propto x - 3.5$



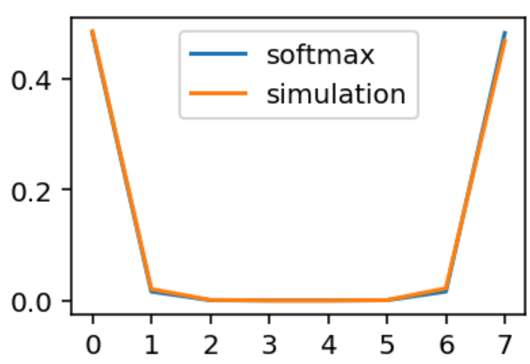
Choose in 4 seconds with $u(x) \propto |x - 3.5|$



Choose in 12 seconds with $u(x) \propto x - 3.5$



Choose in 12 seconds with $u(x) \propto |x - 3.5|$



These simulations show that, when calibrated with the experimental data of Milosavljevic et al. (2010), our Metropolis-DDM algorithm converges numerically to its limit distribution.

4.4.2 Metropolis-DDM vs. MA-DDM

Before studying other testable implications of the Metropolis-DMM algorithm, it is important to compare its prediction about the asymptotic choice distribution with the prediction of the Multi-Alternative-DDM (MA-DDM), which is the most popular competing model. According to the MA-DDM,[31] $|A|$ neuronal populations are **simultaneously activated** when an agent faces a menu A of more than two alternatives and a time constraint t. Each population is primarily sensitive to an alternative $a \in A$, and evidence accumulation in favor of each a is represented by a Brownian motion with drift

$$V_a^{\sigma}(\tau) = u(a)\,\tau + \sigma W_a(\tau)$$

The additional diffusion parameter σ accounts for the fact that the multiplicity of alternatives involved might affect the noise of evidence accumulation. In this model, the strategy is simply to choose the alternative supported by the largest amount of (gross) evidence at time t: no thresholds are involved. That is, choice probabilities are given, for all alternatives $a \in A$, by

$$q_t(a, A) = \mathbb{P}\left(V_a^{\sigma}(t) > V_b^{\sigma}(t) \quad \text{for all } b \in A \setminus \{a\}\right) \tag{MA-DDM}$$

The next result shows that, while the choice probability $p_t(a, A)$ is a **Multinomial Logit** with parameters u and $\lambda(t)$, the choice probability $q_t(a, A)$ is a **Multinomial Probit** with parameters u and $\sigma^{-1}\sqrt{t}$.

Proposition 11 *Let $u \in \mathbb{R}^A$, $t \in (0, \infty)$ and $\sigma \in (0, \infty)$. If $A = \{a_1, ..., a_k\}$, then*

$$q_t(a_i, A) = \Phi\left(\frac{\sqrt{t}}{\sigma}[u(a_i) - u(a_1)], ..., \frac{\sqrt{t}}{\sigma}[u(a_i) - u(a_{i-1})], \frac{\sqrt{t}}{\sigma}[u(a_i) - u(a_{i+1})], ...\right)$$

for all $i = 1, ..., k$, where Φ is the $k-1$ dimensional normal distribution with mean 0 and covariance matrix of all 1's off the diagonal and 2's on it.

Multinomial logit and probit are arguably the most popular discrete choice distributions and there are standard econometric techniques to distinguish between them. These techniques thus provide a way to compare the Metropolis-DDM and the MA-DDM.

Remark 1 *Notice that the Metropolis-DDM only requires the simultaneous representation of two objects at any given moment, while the MA-DDM requires the representation of $|A|$ objects. This representation has to be implemented in working memory; but – as well known – the number of items in active representation in working memory is bounded by a number around four. This is noted by both Russo and Rosen (1975, pp. 267 and 275) and Krajbich and Rangel (2011, p. 13856). See also Luck and Vogel (1997) and Vogel and Machizawa (2004).*

[31] See, e.g., Roe, Busemeyer, and Townsend (2001), McMillen and Holmes (2006), and Bogacz, Usher, Zhang, and McClelland (2007).

4.4.3 Matching eye-tracking evidence

Eye-tracking data come in the form of fixation sequences, say

$$x_0 - x_1 - x_2 - x_3 - \cdots \tag{13}$$

Subsequences $x - y$, or more importantly refixation subsequences $x - y - x$ (or $x - y - x - y$, ...), are hypothesized to correspond to pairwise comparisons, while the appearance of a new alternative, say $x - y - x - z$, relates to exploration. The sequence of comparisons generated by the Metropolis-DDM is

$$b_0 \leftrightsquigarrow a_1 \to b_1 \leftrightsquigarrow a_2 \to b_2 \leftrightsquigarrow \cdots \tag{14}$$

where the b_n are incumbents (temporary solutions to the decision problem) and the a_{n+1} are proposals (outputs of the exploration process, generated according to $Q(\cdot \mid b_n)$).

These two facts imply that, if the Metropolis-DDM is the actual decision process, then (14) can be obtained from the observed sequence (13) of eye fixations by replacing subsequences:

- $x - y - z$ with $x \leftrightsquigarrow y \to y \leftrightsquigarrow z$,
- $x - y - x - z$ with $x \leftrightsquigarrow y \to x \leftrightsquigarrow z$,
- $x - y - x - y - z$ with $x \leftrightsquigarrow y \to y \leftrightsquigarrow z$,

and so on.

In any case, (14) is observable by the analyst and its long-run frequency is qualitatively similar to the long-run fraction of times in which the subject's gaze fixates the various alternatives. The next proposition computes this long-run frequency with the objective to compare it with the eye-tracking findings of Krajbich and Rangel (2011, henceforth KR). Since KR have three equally spaced alternatives, we consider a uniform exploration matrix Q.[32] In reading it, remember that only half of the fixations (those corresponding to the proposals a) are related to exploration, the other half (those corresponding to the incumbents b) are the results of binary comparisons.

Proposition 12 *Let $u \in \mathbb{R}^A$, $t \in (0, \infty)$ and $\lambda(t) \in (0, \infty)$. If Q is uniform, then the long-run frequency of the sequence*

$$(b_0, a_1, b_1, ..., b_n, a_{n+1}, b_{n+1}, ...) \tag{15}$$

generated by the Metropolis-DDM algorithm is

$$\frac{1}{2} p_t(\cdot, A) + \frac{1}{2} \frac{1}{|A|} \tag{16}$$

almost surely.

[32] On page 13855 they remark that "The influence of value on the fixation process was taken into account in our numerical simulations, although the results do not change if we instead assume that fixations are random and independent of value."

The first implication of this result is that, though subjects are more likely to fixate on the best item as the trial progresses, they still continue to look at the worst item x_*. KR (pp. 13854-13855) find a rate of no less than 20% and a final probability of choosing x_* of 6%. Recalling that our Proposition 8 predicts a final probability $p_t(x_*, A)$, and thus assuming $p_t(x_*, A) = 6/100$, our theoretical estimate is

$$\frac{1}{2} \cdot \frac{6}{100} + \frac{1}{2} \cdot \frac{1}{3} = 19.\bar{6}\%$$

Remarkably, it almost perfectly matches their empirical 20%.[33] They also find that items that are more fixated on are more likely to be chosen, and that the last fixation is usually to the best item. Propositions 8 and 12 predict precisely this fixation pattern, and the Metropolis-DDM algorithm chooses either the last or the second last fixated item. Our predictions are also consistent with their observations that subjects most often choose the best item, and that the likelihood of this decreases with set size and increases conditional on it being either the last seen or the one seen before the last.[34]

Summing up, the Metropolis-DDM algorithm is consistent with the qualitative findings of Krajbich and Rangel (2011), who adopt a variation of the MA-DDM to explain them.[35]

5 Concluding remarks

In this paper, we axiomatically characterized Multinomial Logit Processes

$$p_t(a, A) = \frac{e^{\lambda(t)u(a)}}{\sum_{b \in A} e^{\lambda(t)u(b)}} \qquad \forall a \in A \in \mathcal{A} \quad \forall t \in T \tag{MLP}$$

that is, processes $\{p_t\}_{t \in T}$ of random choice rules that have Multinomial Logit distributions with t-dependent scale parameter $\lambda(t)$.[36] We also showed how, when t is interpreted as an exogenously given and fixed deliberation time, softmax distributions emerge as stationary selection probabilities of a search algorithm *à la* Metropolis in which the acceptance/rejection rule is dictated by the Drift Diffusion Model. We thus provided a neuropshychological foundation for (MLP) as a theory of preference discovery.

The Metropolis-DDM algorithm that we presented here relies on the simplest versions of Markovian exploration and of the Drift Diffusion Model. A natural extension of our neuropsychological analysis consists in considering more general two-alternative forced-choice models, such as the ones discussed in Diederich and Busemeyer (2003), Ratcliff and Smith (2004), Bogacz et al. (2006), Krajbich, Armel, and Rangel (2010), and Rustichini and Padoa-Schioppa

[33] Reutskaja, Nagel, Camerer, and Rangel (2011, p. 910) also find "many" refixations on already scrutinized inferior alternatives.

[34] See Figure 4 on page 13856 of KR, but also page 908 and Figure 3 on page 909 of Reutskaja, Nagel, Camerer, and Rangel (2011). We thank a referee for pointing this out.

[35] Similar considerations apply to Reutskaja, Nagel, Camerer, and Rangel (2011), where another family of (non-DDM based) models is examined.

[36] Specifically, see Theorems 2 and 6 in the main text and their more general versions Theorem 16 and Lemma 20 in the appendix.

(2015).[37] This goes beyond the scope of the present paper, whose objective is to build a first bridge between the axiomatic and the neuro-computational approaches to random choice. But it is the object of current research (see Baldassi, Cerreia-Vioglio, Maccheroni, and Marinacci, 2017).

Another natural extension – that we are currently investigating – is the possibility of softening the hard time constraint t of the Metropolis-DDM algorithm and making it endogenous. A natural direction of investigation is suggested by the Simulated Annealing heuristics of the seminal Kirkpatrick, Gelatt, and Vecchi (1983). *Simulated Annealing* consists in performing a Metropolis algorithm while "slowly" increasing the threshold $z = \lambda(t)$ instead of keeping it fixed.[38] So that, as $\lambda(t) \to \infty$, the limit distribution of incumbents concentrates on the optimal solution:[39]

$$\frac{e^{\lambda(t)u(a)}}{\sum_{b \in A} e^{\lambda(t)u(b)}} \to \delta_{\arg\max_A u}(a) \qquad \forall a \in A$$

Analogously, in the *Simulated DDM-Annealing* the initial threshold $\lambda(s_0)$ has to be maintained from $\tau = 0$ to $\tau = s_0^-$, with s_0 sufficiently large to obtain a frequency

$$\hat{p}_{s_0}(a, A) \approx \frac{e^{\lambda(s_0)u(a)}}{\sum_{b \in A} e^{\lambda(s_0)u(b)}} \qquad \forall a \in A$$

Subsequently, $\lambda(s_1) > \lambda(s_0)$ has to be maintained from $\tau = s_0$ to $\tau = s_1^-$, with s_1 sufficiently large to obtain a frequency

$$\hat{p}_{s_1}(a, A) \approx \frac{e^{\lambda(s_1)u(a)}}{\sum_{b \in A} e^{\lambda(s_1)u(b)}} \qquad \forall a \in A$$

and so on, with the objective of generating a long-run frequency

$$\lim_{i \to \infty} \hat{p}_{s_i}(a, A) \approx \lim_{i \to \infty} \frac{e^{\lambda(s_i)u(a)}}{\sum_{b \in A} e^{\lambda(s_i)u(b)}} = \delta_{\arg\max_A u}(a) \qquad \forall a \in A$$

In particular, the strictly increasing sequence $\{s_i, \lambda(s_i)\}_{i \in \mathbb{N}}$ corresponds to the annealing schedule of Kirkpatrick, Gelatt, and Vecchi (1983).[40] Because of the properties of the DDM that we discussed above, as time passes and λ increases, pairwise comparisons will take longer and they will become more precise, with the probability of accepting new proposals decreasing to 0. Like the original version, this variation of the Simulated Annealing heuristics is assumed to run until the system "freezes", that is, the incumbent stops changing.[41] The selected alternative is the agent's final choice. Interestingly, and differently from the Metropolis-DDM algorithm, response time in the Simulated DDM-Annealing procedure becomes **endogenous**.

[37] Also the modelling of menu exploration can be made more realistic, for example by taking into account the visual saliency of alternatives (see Milosavljevic, Navalpakkam, Koch, and Rangel, 2012).

[38] "At each temperature [here $1/z$], the simulation must proceed long enough for the system to reach a steady state...", as Kirkpatrick, Gelatt, and Vecchi (1983) put it.

[39] For simplicity of exposition, here $\arg\max_A u = \{a^*\}$ is a singleton.

[40] Here $\{s_i\}$ is a sequence of times, whereas in the original it is a sequence of numbers of iterations.

[41] Here "stops changing" means it does not change for a "large" amount of time, whereas in the original "stops changing" means that it does not change for a "large" number of iterations.

A Proofs

A.1 Preliminaries

Lemma 13 *Let $p : \mathcal{A} \to \Delta(X)$ be a random choice rule. The following conditions are equivalent:*

1. *p is such that, $p_A(C) = p_B(C)\, p_A(B)$ for all $C \subseteq B \subseteq A$ in $\mathcal{A}$;*
2. *p satisfies the Choice Axiom;*
3. *p is such that $p(b, B)\, p(a, A) = p(a, B)\, p(b, A)$ for all $B \subseteq A$ in $\mathcal{A}$ and all $a, b \in B$;*
4. *p satisfies Independence from Irrelevant Alternatives;*
5. *p is such that $p(Y \cap B, A) = p(Y, B)\, p(B, A)$ for all $B \subseteq A$ in $\mathcal{A}$ and all $Y \subseteq X$.*

Moreover, in this case, p satisfies Positivity if and only if it satisfies Full Support.

Proof *1 implies 2.* Choose as C the singleton a appearing in the statement of the axiom.

2 implies 3. Given any $B \subseteq A$ in $\mathcal{A}$ and any $a, b \in B$, by the Choice Axiom, $p(a, A) = p(a, B)\, p(B, A)$, but then $p(b, B)\, p(a, A) = p(a, B)\, p(b, B)\, p(B, A) = p(a, B)\, p(b, A)$ where the second equality follows from another application of the Choice Axiom.

3 implies 4. Let $A \in \mathcal{A}$ and arbitrarily choose $a, b \in A$ such that $p(a, A) / p(b, A) \neq 0/0$. By 3,

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

three cases have to be considered:

- $p(b, a) \neq 0$ and $p(b, A) \neq 0$, then $p(a, A) / p(b, A) = p(a, b) / p(b, a)$;
- $p(b, a) = 0$, then $p(a, b)\, p(b, A) = 0$, but $p(a, b) \neq 0$ (because $p(a, b) / p(b, a) \neq 0/0$), thus $p(b, A) = 0$ and $p(a, A) \neq 0$ (because $p(a, A) / p(b, A) \neq 0/0$); therefore
$$\frac{p(a, b)}{p(b, a)} = \infty = \frac{p(a, A)}{p(b, A)}$$
- $p(b, A) = 0$, then $p(b, a)\, p(a, A) = 0$, but $p(a, A) \neq 0$ (because $p(a, A) / p(b, A) \neq 0/0$), thus $p(b, a) = 0$ and $p(a, b) \neq 0$ (because $p(a, b) / p(b, a) \neq 0/0$); therefore
$$\frac{p(a, A)}{p(b, A)} = \infty = \frac{p(a, b)}{p(b, a)}$$

4 implies 3. Given any $B \subseteq A$ in $\mathcal{A}$ and any $a, b \in B$:

- If $p(a, A) / p(b, A) \neq 0/0$ and $p(a, B) / p(b, B) \neq 0/0$, then by IIA
$$\frac{p(a, A)}{p(b, A)} = \frac{p(a, b)}{p(b, a)} = \frac{p(a, B)}{p(b, B)}$$

- $\circ$ If $p(b,A) \neq 0$, then $p(b,B) \neq 0$ and $p(b,B)\,p(a,A) = p(a,B)\,p(b,A)$.
- $\circ$ Else $p(b,A) = 0$, then $p(b,B) = 0$ and again $p(b,B)\,p(a,A) = p(a,B)\,p(b,A)$.

- Else, either $p(a,A)/p(b,A) = 0/0$ or $p(a,B)/p(b,B) = 0/0$ and in both cases

$$p(b,B)\,p(a,A) = p(a,B)\,p(b,A)$$

3 implies 5. Given any $B \subseteq A$ in $\mathcal{A}$ and any $Y \subseteq X$, since $p(B,B) = 1$, it follows $p(Y,B) = p(Y \cap B, B)$. Therefore

$$\begin{aligned} p(Y \cap B, A) &= \sum_{y \in Y \cap B} p(y,A) = \sum_{y \in Y \cap B} \left(\sum_{x \in B} p(x,B) \right) p(y,A) = \sum_{y \in Y \cap B} \left(\sum_{x \in B} p(x,B)\,p(y,A) \right) \\ [\text{by } 3] &= \sum_{y \in Y \cap B} \left(\sum_{x \in B} p(y,B)\,p(x,A) \right) = \sum_{y \in Y \cap B} p(y,B) \left(\sum_{x \in B} p(x,A) \right) \\ &= \sum_{y \in Y \cap B} p(y,B)\,p(B,A) = p(Y \cap B, B)\,p(B,A) = p(Y,B)\,p(B,A) \end{aligned}$$

5 implies 1. Take $Y = C$.

Finally, let p satisfy the Choice Axiom. Assume – *per contra* – Positivity holds and $p(a,A) = 0$ for some $A \in \mathcal{A}$ and some $a \in A$. Then $A \neq \{a\}$ and, for all $b \in A \setminus \{a\}$, the Choice Axiom implies $0 = p(a,A) = p(a,\{a,b\})\,p(\{a,b\},A) = p(a,b)\,(p(a,A) + p(b,A)) = p(a,b)\,p(b,A)$ whence $p(b,A) = 0$ (because $p(a,b) \neq 0$), contradicting $p(A,A) = 1$. Therefore Positivity implies Full Support. The converse is trivial. ■

The next result is a special case of the general Luce's model of Echenique and Saito (2018). Specifically, Theorem 14 extends Theorem 1 by maintaining the assumption of Independence from Irrelevant Alternatives while removing that of Full Support. In the subsequent analysis this theorem will allow us to distill the utility function u starting from choice frequencies. In reading it, recall that a *choice correspondence* is a map $\Gamma : \mathcal{A} \to \mathcal{A}$ such that $\Gamma(A) \subseteq A$ for all $A \in \mathcal{A}$. A choice correspondence is called *rational* if and only if it satisfies the *Weak Axiom of Revealed Preference* of Arrow (1959), that is,

$$B \subseteq A \in \mathcal{A} \text{ and } \Gamma(A) \cap B \neq \varnothing \text{ imply } \Gamma(B) = \Gamma(A) \cap B \tag{WARP}$$

Theorem 14 *A random choice rule $p : \mathcal{A} \to \Delta(X)$ satisfies the Choice Axiom if and only if there exist a function $v : X \to (0,\infty)$ and a rational choice correspondence $\Gamma : \mathcal{A} \to \mathcal{A}$ such that*

$$p(a,A) = \begin{cases} \dfrac{v(a)}{\sum_{b \in \Gamma(A)} v(b)} & \text{if } a \in \Gamma(A) \\ 0 & \text{else} \end{cases} \tag{GLM}$$

for all $A \in \mathcal{A}$ and all $a \in A$.

In this case, Γ is unique and $\Gamma(A) = \operatorname{supp} p_A$ for all $A \in \mathcal{A}$.

Proof See Cerreia-Vioglio, Maccheroni, Marinacci, and Rustichini (2018). ■

The final result of this section characterizes the possibility of considering a constant v in (GLM). In reading it, recall that a random choice rule is *uniform* if and only if

$$p(a, A) = \begin{cases} \dfrac{1}{|\operatorname{supp} p_A|} & \text{if } a \in \operatorname{supp} p_A \\ 0 & \text{else} \end{cases} \qquad \forall a \in A \in \mathcal{A}$$

and that the (binary) relation $\succsim$ generated by a choice correspondence $\Gamma : \mathcal{A} \to \mathcal{A}$ is defined by

$$a \succsim b \iff a \in \Gamma(\{a, b\})$$

Corollary 15 *A uniform random choice rule* $p : \mathcal{A} \to \Delta(X)$ *satisfies the Choice Axiom if and only if* $\operatorname{supp} p : \mathcal{A} \to \mathcal{A}$ *is a rational choice correspondence.*

In this case, the relation $\succsim$ *generated by* $\operatorname{supp} p$ *is characterized by*

$$a \succsim b \iff p(a, b) > 0 \iff p(a, b) \geq p(b, a)$$

and it is a weak order.

A.2 Proofs of the results of Section 3

Proof of Theorem 2 *Only if.* Since $\{p_t\}_{t \in T}$ satisfies Positivity and the Choice Axiom, by Theorem 1, for each $t \in T$, there exists $u_t : X \to \mathbb{R}$ such that

$$p_t(a, A) = \frac{e^{u_t(a)}}{\sum_{b \in A} e^{u_t(b)}} \qquad \forall a \in A \in \mathcal{A} \tag{17}$$

Arbitrarily choose $\bar{c} \in X$ and replace each u_t with $u_t - u_t(\bar{c})$. With this, $u_t(\bar{c}) = 0$ for all $t \in T$ and (17) still holds.

If, for all $t \in T$, u_t is constant, then MLP holds (e.g., with $u(x) = 0$ for all $x \in X$). Otherwise, there exists $\bar{t} \in T$ such that $u_{\bar{t}}$ is not constant, so that $u_{\bar{t}}(\bar{b}) \neq 0 = u_{\bar{t}}(\bar{c})$ for some $\bar{b} \in X$. This implies that

$$\frac{\ell_{\bar{t}}(a, \bar{c})}{\ell_{\bar{t}}(\bar{b}, \bar{c})} = \frac{u_{\bar{t}}(a) - u_{\bar{t}}(\bar{c})}{u_{\bar{t}}(\bar{b}) - u_{\bar{t}}(\bar{c})} = \frac{u_{\bar{t}}(a)}{u_{\bar{t}}(\bar{b})}$$

is a well defined real number for all $a \in X$. By Log-odds Ratio Invariance, $\ell_t(a, \bar{c}) / \ell_t(\bar{b}, \bar{c})$ is well defined too,[42] and

$$\frac{u_t(a)}{u_t(\bar{b})} = \frac{\ell_t(a, \bar{c})}{\ell_t(\bar{b}, \bar{c})} = \frac{\ell_{\bar{t}}(a, \bar{c})}{\ell_{\bar{t}}(\bar{b}, \bar{c})} = \frac{u_{\bar{t}}(a)}{u_{\bar{t}}(\bar{b})} \in \mathbb{R} \qquad \forall (a, t) \in X \times T$$

[42] Notice that Log-odds Ratio Invariance is equivalent to the stronger:

Strong Log-odds Ratio Invariance *Given any s and t in T,*

$$\frac{\ell_t(a, c)}{\ell_t(b, c)} = \frac{\ell_s(a, c)}{\ell_s(b, c)}$$

for all $a, b, c \in X$ such that either ratio is well defined.

Therefore, $u_t(\bar{b}) \neq 0 = u_t(\bar{c})$ for all $t \in T$, and

$$u_t(a) = \frac{u_t(\bar{b})}{u_{\bar{t}}(\bar{b})} u_{\bar{t}}(a) \qquad \forall (a,t) \in X \times T. \tag{18}$$

Consider the case in which $u_{\bar{t}}(\bar{b}) > 0 = u_{\bar{t}}(\bar{c})$. If $t > \bar{t}$, then (17) and Weak Consistency imply

$$p_{\bar{t}}(\bar{b}, \bar{c}) > p_{\bar{t}}(\bar{c}, \bar{b}) \implies p_t(\bar{b}, \bar{c}) > p_t(\bar{c}, \bar{b}) \implies u_t(\bar{b}) > u_t(\bar{c}) = 0$$

thus $u_t(\bar{b}) / u_{\bar{t}}(\bar{b}) > 0$. This is clearly true also if $t = \bar{t}$. Else $t < \bar{t}$, assume *per contra* $u_t(\bar{b}) < 0 = u_t(\bar{c})$, then (17) and Weak Consistency imply

$$p_t(\bar{c}, \bar{b}) > p_t(\bar{b}, \bar{c}) \implies p_{\bar{t}}(\bar{c}, \bar{b}) > p_{\bar{t}}(\bar{b}, \bar{c}) \implies u_{\bar{t}}(\bar{b}) < u_{\bar{t}}(\bar{c}) = 0$$

a contradiction. Thus $u_t(\bar{b}) / u_{\bar{t}}(\bar{b}) > 0$ holds for all $t \in T$ provided $u_{\bar{t}}(\bar{b}) > 0$. It is easy to show that the same is true if $u_{\bar{t}}(\bar{b}) < 0$.[43] This shows that

$$\begin{array}{rccc} \lambda : & T & \to & (0, \infty) \\ & t & \mapsto & \frac{u_t(\bar{b})}{u_{\bar{t}}(\bar{b})} \end{array}$$

is well defined. Moreover, the function $u = u_{\bar{t}} : X \to \mathbb{R}$ is non-constant and relation (18) implies

$$u_t(a) = \lambda(t) u(a) \qquad \forall (a,t) \in X \times T$$

which together with (17) shows that the axioms imply representation MLP.

If. It is easy to verify that the converse implication holds too. For the sake of completeness, we check that representation MLP implies Strong Log-odds Ratio Invariance. Let $t, s \in T$ and $a, b, c, x, y \in X$. Notice that

$$\ell_t(x,y) = \lambda(t)[u(x) - u(y)]$$

so that $\ell_t(x,y) = 0$ if and only if $u(x) = u(y)$, and the same considerations hold with s in place of t. Assume $\ell_s(a,c) / \ell_s(b,c)$ is well defined:

- If $\ell_s(b,c) = 0$, then $u(b) = u(c)$ and $\ell_s(a,c) \neq 0$, so $u(a) \neq u(c)$, then
  - $\ell_s(a,c) = \lambda(s)[u(a) - u(c)] \neq 0$ and since $\lambda(s) > 0$, then

$$\frac{\ell_s(a,c)}{\ell_s(b,c)} = \frac{\lambda(s)[u(a) - u(c)]}{0} = \frac{u(a) - u(c)}{0}$$

[43] If $t > \bar{t}$, then, by (17) and Weak Consistency, we have

$$u_{\bar{t}}(\bar{b}) < 0 = u_{\bar{t}}(\bar{c}) \implies p_{\bar{t}}(\bar{c}, \bar{b}) > p_{\bar{t}}(\bar{b}, \bar{c}) \implies p_t(\bar{c}, \bar{b}) > p_t(\bar{b}, \bar{c}) \implies u_t(\bar{b}) < u_t(\bar{c}) = 0$$

thus $u_t(\bar{b}) / u_{\bar{t}}(\bar{b}) > 0$. This is clearly true also if $t = \bar{t}$. Else $t < \bar{t}$, assume *per contra* $u_t(\bar{b}) > 0 = u_t(\bar{c})$, then (17) and Weak Consistency imply

$$p_t(\bar{b}, \bar{c}) > p_t(\bar{c}, \bar{b}) \implies p_{\bar{t}}(\bar{b}, \bar{c}) > p_{\bar{t}}(\bar{c}, \bar{b}) \implies u_{\bar{t}}(\bar{b}) > u_{\bar{t}}(\bar{c}) = 0$$

a contradiction. Thus $u_t(\bar{b}) / u_{\bar{t}}(\bar{b}) > 0$ holds for all $t \in T$.

- $\ell_t(b,c) = \lambda(t)[u(b) - u(c)] = 0$, because $u(b) = u(c)$,
- $\ell_t(a,c) = \lambda(t)[u(a) - u(c)] \neq 0$, because $u(a) \neq u(c)$, and since $\lambda(t) > 0$, then

$$\frac{\ell_t(a,c)}{\ell_t(b,c)} = \frac{\lambda(t)[u(a) - u(c)]}{0} = \frac{u(a) - u(c)}{0} = \frac{\ell_s(a,c)}{\ell_s(b,c)}$$

- Else $\ell_s(b,c) \neq 0$, then $u(b) \neq u(c)$ and $\ell_t(b,c) = \lambda(t)[u(b) - u(c)] \neq 0$, so that

$$\frac{\ell_s(a,c)}{\ell_s(b,c)} = \frac{\lambda(s)[u(a) - u(c)]}{\lambda(s)[u(b) - u(c)]} = \frac{u(a) - u(c)}{u(b) - u(c)} = \frac{\lambda(t)[u(a) - u(c)]}{\lambda(t)[u(b) - u(c)]} = \frac{\ell_t(a,c)}{\ell_t(b,c)}$$

The case in which $\ell_t(a,c)/\ell_t(b,c)$ is well defined is analogous.

As to uniqueness of u and λ, notice that, if also $\bar{u}$ and $\bar{\lambda}$ represent $\{p_t\}_{t \in T}$ in the sense of MLP, then

$$e^{\lambda(t)(u(a)-u(b))} = r_t(a,b) = e^{\bar{\lambda}(t)(\bar{u}(a)-\bar{u}(b))}$$

for all $t \in T$ and all $a, b \in X$. Therefore $\lambda(t)(u(a) - u(b)) = \bar{\lambda}(t)(\bar{u}(a) - \bar{u}(b))$ for all $t \in T$ and all $a, b \in X$. Arbitrarily choose $t^* \in T$ and $b^* \in X$ to conclude that

$$\bar{u}(a) = \frac{\lambda(t^*)}{\bar{\lambda}(t^*)}(u(a) - u(b^*)) + \bar{u}(b^*) = ku(a) + h \qquad \forall a \in X$$

with $k > 0$ and $h \in \mathbb{R}$. Since the converse is also true,[44] cardinal uniqueness of u follows. Moreover, if u is not constant, choosing $a, b \in X$ with $u(a) \neq u(b)$, the previous argument yields $\lambda(t)(u(a) - u(b)) = \bar{\lambda}(t)(ku(a) - ku(b))$ for all $t \in T$, so that $\bar{\lambda} = k^{-1}\lambda$ if $\bar{u} = ku + h$. Finally, if $\bar{u} = u$ is given, so that $k = 1$, it follows $\bar{\lambda} = \lambda$. ■

Inspection of the proof shows that it is possible to obtain the softmax representation for any non-singleton index set T by replacing Weak Consistency with:

Strong Consistency *Given any s and t in T,*

$$p_t(a,b) > p_t(b,a) \implies p_s(a,b) > p_s(b,a)$$

for all $a, b \in X$.

As we discussed in the main text, the theorem below characterizes the most general version of the Heteroscedastic Multinomial Logit Model.

Theorem 16 *A collection $\{p_t\}_{t \in T}$ of random choice rules satisfies Positivity, the Choice Axiom, Strong Consistency, and Strong Log-odds Ratio Invariance if and only if there exist $u : X \to \mathbb{R}$ and $\lambda : T \to (0, \infty)$ such that*

$$p_t(a,A) = \frac{e^{\lambda(t)u(a)}}{\sum_{b \in A} e^{\lambda(t)u(b)}} \tag{HMNL}$$

for all $A \in \mathcal{A}$, all $a \in A$, and all $t \in T$.

In this case, u is cardinally unique, and λ is unique given u unless the latter is constant.

[44] That is, if $\bar{u} = ku + h$ with $k > 0$ and $h \in \mathbb{R}$, then $\bar{u}$ and $\bar{\lambda} = k^{-1}\lambda$ represent $\{p_t\}_{t \in T}$ in the sense of MLP.

Propositions 3 and **4** are simple corollaries of Theorem 2 (we leave the routine proofs to the reader). The proof of Proposition 5 is less trivial and builds on the following:

Proposition 17 *Let* $\{p_t\}_{t\in(0,\infty)}$ *be a Multinomial Logit Model with utility* u. *Then*

$$p_s(\{a\in A : u(a)\geq h\}, A) \geq p_t(\{a\in A : u(a)\geq h\}, A) \qquad \forall h\in\mathbb{R}$$

for all $s>t$ *in* $(0,\infty)$ *and all* $A\in\mathcal{A}$.

Proof Arbitrarily choose $A\in\mathcal{A}$. If we prove that, given any $h\in\mathbb{R}$ such that $\varnothing\subsetneq\{c\in A : u(c)\geq h\}\subsetneq A$, it holds

$$\frac{\mathrm{d}}{\mathrm{d}t}\, p_t(\{c\in A : u(c)\geq h\}, A) > 0 \qquad \forall t\in(0,\infty) \tag{19}$$

then the statement follows. Indeed, in this case, the function $t\mapsto p_t(\{c\in A : u(c)\geq h\}, A)$ is strictly increasing on $(0,\infty)$ for all $h\in\mathbb{R}$ such that $\varnothing\subsetneq\{c\in A : u(c)\geq h\}\subsetneq A$, while it is constantly equal to 0 or 1 if h is such that $\{c\in A : u(c)\geq h\}=\varnothing$ or $\{c\in A : u(c)\geq h\}=A$.

Next we show that (19) holds. Set $[u\geq h]=\{c\in A : u(c)\geq h\}$, assume $h\in\mathbb{R}$ is such that $\varnothing\subsetneq[u\geq h]\subsetneq A$, and notice that in this case $[u<h]$ must be non-empty. Given any $t\in(0,\infty)$, with the abbreviation $\sum\limits_{u(c)\geq h}=\sum\limits_{c\in A:u(c)\geq h}$, we have

$$\begin{aligned}
0<\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\sum\limits_{u(c)\geq h} e^{tu(c)}}{\sum\limits_{b\in A} e^{tu(b)}}\right) &= \frac{\left(\sum\limits_{b\in A} e^{tu(b)}\right)\sum\limits_{u(c)\geq h} u(c)\,e^{tu(c)} - \sum\limits_{u(c)\geq h} e^{tu(c)}\left(\sum\limits_{b\in A} u(b)\,e^{tu(b)}\right)}{\left(\sum\limits_{b\in A} e^{tu(b)}\right)^2}\\
\iff 0 &< \sum_{u(c)\geq h} u(c)\,e^{tu(c)}\sum_{u(b)<h} e^{tu(b)} - \sum_{u(c)<h} u(c)\,e^{tu(c)}\sum_{u(b)\geq h} e^{tu(b)}\\
\iff 0 &< \frac{\sum\limits_{u(c)\geq h} u(c)\,e^{tu(c)}}{\sum\limits_{u(b)\geq h} e^{tu(b)}} - \frac{\sum\limits_{u(c)<h} u(c)\,e^{tu(c)}}{\sum\limits_{u(b)<h} e^{tu(b)}}\\
\iff 0 &< \sum_{u(c)\geq h} u(c)\left(\frac{e^{tu(c)}}{\sum\limits_{u(b)\geq h} e^{tu(b)}}\right) - \sum_{u(c)<h} u(c)\left(\frac{e^{tu(c)}}{\sum\limits_{u(b)<h} e^{tu(b)}}\right)\\
\iff 0 &< \sum_{c\in[u\geq h]} u(c)\,p_t(c,[u\geq h]) - \sum_{c\in[u<h]} u(c)\,p_t(c,[u<h])
\end{aligned}$$

and this concludes the proof, because – in the last step above – the minuend is an average (i.e., a convex combination) of values $u(c)\geq h$, so it cannot be smaller than h itself, the subtrahend is an average of values $u(c)<h$, so it is strictly smaller than h itself, hence the difference on the right hand side is strictly positive, irrespectively of the value of t. ■

Proof of Proposition 5 If u is constant, the requirements of points 1 and 2 are automatically satisfied, while that of point 3 is vacuous. Assume then that u is non-constant.

1 implies 3. Let $s > t$ in T. Arbitrarily choose $a, b \in X$ such that $u(a) > u(b)$. Since $p_t(a,b) > p_t(b,a)$, the Weak Discovered Preference Axiom implies that $p_s(a,b) \geq p_t(a,b)$, but then

$$\frac{p_s(a,b)}{1 - p_s(a,b)} \geq \frac{p_t(a,b)}{1 - p_t(a,b)}$$

that is, $r_s(a,b) \geq r_t(a,b)$. Passing to the logarithms, it follows that $\lambda(t)(u(a) - u(b)) \geq \lambda(s)(u(a) - u(b))$ and $\lambda(t) \geq \lambda(s)$.

3 implies 2. Let $s > t$ in T and observe that $\lambda(s) \geq \lambda(t)$. Denote by $\{q_l\}_{l \in (0,\infty)}$ the Multinomial Logit Model with utility u. By Proposition 17, it follows that, for every $A \in \mathcal{A}$, if $l \geq l'$, then

$$q_l(\{a \in A : u(a) \geq h\}, A) \geq q_{l'}(\{a \in A : u(a) \geq h\}, A) \qquad \forall h \in \mathbb{R}$$

Now, taking $l = \lambda(s)$ and $l' = \lambda(t)$, we have

$$q_{\lambda(s)}(\{a \in A : u(a) \geq h\}, A) \geq q_{\lambda(t)}(\{a \in A : u(a) \geq h\}, A) \qquad \forall h \in \mathbb{R}$$

and

$$p_s(\{a \in A : u(a) \geq h\}, A) \geq p_t(\{a \in A : u(a) \geq h\}, A) \qquad \forall h \in \mathbb{R}$$

that is, $U_A(s)$ stochastically dominates $U_A(t)$.

2 implies 1. Let $s > t$ in T. If $p_t(a,b) > p_t(b,a)$, then $u(a) > u(b)$ and there exists $h \in \mathbb{R}$ such that $u(a) > h > u(b)$. Since $U_{\{a,b\}}(s)$ stochastically dominates $U_{\{a,b\}}(t)$, then

$$p_s(\{x \in \{a,b\} : u(x) \geq h\}, \{a,b\}) \geq p_t(\{x \in \{a,b\} : u(x) \geq h\}, \{a,b\})$$

but $\{x \in \{a,b\} : u(x) \geq h\} = \{a\}$, therefore $p_s(a,b) \geq p_t(a,b)$. ■

Lemma 18 *Let $\{p_t\}_{t \in (0,\infty)}$ be a random choice process that satisfies Positivity, the Choice Axiom, Continuity, Consistency, and Asymptotic Uniformity. Then:*

(i) the relation defined on X by $a \succsim b$ if and only if $p_\infty(a,b) > 0$ is a weak order, and

$$a \succsim b \iff p_\infty(a,b) \geq p_\infty(b,a)$$

(ii) given any $a, b \in X$, the function $\varphi_{a,b} : (0,\infty) \to (0,\infty)$ defined by

$$\varphi_{a,b}(t) = r_t(a,b) \qquad \forall t \in (0,\infty)$$

is continuous and either constantly equal to 1 (if $a \sim b$), or divergent at ∞ as $t \to \infty$ (if $a \succ b$), or vanishing as $t \to \infty$ (if $b \succ a$).

Proof (i) By Theorem 1, for each $t \in (0,\infty)$, there exists $v_t : X \to (0,\infty)$ such that

$$p_t(a,A) = \frac{v_t(a)}{\sum_{b \in A} v_t(b)} \tag{20}$$

for all $A \in \mathcal{A}$ and all $a \in A$.[45] While p_∞ satisfies the Choice Axiom because it is defined by Continuity and p_t satisfies the Choice Axiom for all $t \in (0, \infty)$. By Theorem 14, there exist $v : X \to (0, \infty)$ and a rational choice correspondence $\Gamma : \mathcal{A} \to \mathcal{A}$ such that

$$p_\infty (a, A) = \begin{cases} \dfrac{v (a)}{\sum_{b \in \Gamma(A)} v (b)} & \text{if } a \in \Gamma (A) \\ 0 & \text{else} \end{cases}$$

for all $A \in \mathcal{A}$ and all $a \in A$; moreover, $\Gamma (A) = \operatorname{supp} (p_\infty)_A$ for all $A \in \mathcal{A}$. Given any $A \in \mathcal{A}$ and any $a, b \in \operatorname{supp} (p_\infty)_A = \Gamma (A)$, we have that $p_\infty (a, A), p_\infty (b, A) \neq 0$ and so $p_\infty (a, A) / p_\infty (b, A)$ is well defined. By Lemma 13 (the equivalence of the Choice Axiom and Independence from Irrelevant Alternatives), we have that

$$\frac{v (a)}{v (b)} = \frac{p_\infty (a, A)}{p_\infty (b, A)} = \frac{p_\infty (a, b)}{p_\infty (b, a)}$$

If $a \neq b$, since $v (a), v (b) \in (0, \infty)$, it must be the case that $p_\infty (a, b) \neq 0, 1$, then, by Asymptotic Uniformity, $p_\infty (a, b) = 1/2 = p_\infty (b, a)$ and so $v (a) = v (b)$. Since the choice of $A \in \mathcal{A}$ and $a, b \in \operatorname{supp} (p_\infty)_A$ was arbitrary, it follows that v is constant on $\Gamma (A)$ for all $A \in \mathcal{A}$, and so p_∞ is a uniform random choice rule. But then Corollary 15 guarantees that the relation defined on X by

$$a \succsim b \iff p_\infty (a, b) > 0 \iff p_\infty (a, b) \geq p_\infty (b, a) \tag{21}$$

is a weak order.[46]

(ii) Given any $t \in (0, \infty)$,

$$\varphi_{a,b} (t) = r_t (a, b) = \frac{p_t (a, b)}{p_t (b, a)} = \frac{v_t (a)}{v_t (b)} \in (0, \infty)$$

for all $a, b \in X$, thus $\varphi_{a,b} : (0, \infty) \to (0, \infty)$ is well defined. Moreover, by Continuity, $\varphi_{a,b}$ is continuous on $(0, \infty)$ too.

- If $a \sim b$, and *per contra* $\varphi_{a,b} (t) \neq 1$ for some $t \in (0, \infty)$, then
  - either $\varphi_{a,b} (t) > 1$, thus $p_t (a, b) > p_t (b, a)$ and, by Consistency, $p_\infty (a, b) > p_\infty (b, a)$, contradicting $a \sim b$,
  - or $\varphi_{a,b} (t) < 1$, thus $p_t (a, b) < p_t (b, a)$ and, by Consistency, $p_\infty (a, b) < p_\infty (b, a)$, contradicting $a \sim b$,

  so we can conclude $\varphi_{a,b} (t) = 1$ for all $t \in (0, \infty)$.
- If $a \succ b$, by (21) it follows $p_\infty (b, a) = 0$ and $p_\infty (a, b) = 1$, then

  $$\lim_{t \to \infty} \varphi_{a,b} (t) = \lim_{t \to \infty} \frac{p_t (a, b)}{p_t (b, a)} = \frac{p_\infty (a, b)}{p_\infty (b, a)} = \infty$$

  thus $\varphi_{a,b}$ diverges at ∞ as $t \to \infty$.

[45] The v_t's appearing in (20) are the exponentials of the v's appearing in (LM).

[46] And $a \sim b$ if and only if $p_\infty (a, b) = p_\infty (b, a)$, while $a \succ b$ if and only if $p_\infty (a, b) > p_\infty (b, a)$.

- If $b \succ a$, then (since $\varphi_{a,b} = 1/\varphi_{b,a}$ for all $a, b \in X$)

$$\lim_{t\to\infty} \varphi_{b,a}(t) = \infty \implies \lim_{t\to\infty} \varphi_{a,b}(t) = \lim_{t\to\infty} \frac{1}{\varphi_{b,a}(t)} = 0$$

thus $\varphi_{a,b}$ is vanishing as $t \to \infty$. ■

As we discussed in the main text, the Boundedness axiom imposes some discipline on the time combination of odds. To better put this idea into perspective we need the concept of concatenation, which is one of the classical ingredients of the psychophysical theory of measurement.[47]

Definition 3 *A binary operation $\oplus$ on $\mathbb{R}_+$ which is associative, commutative, with identity element* 0*, and such that*

$$t > t' \implies t \oplus s > t' \oplus s \qquad \forall s > 0$$

is called concatenation*.*

Simple examples of concatenation are $t \oplus s = t + s + \kappa ts$ and $t \oplus s = \sqrt[\kappa]{t^\kappa + s^\kappa}$ for some $\kappa > 0$. The concept of concatenation allows us to state a general version of Boundedness.

Weak Boundedness *There exists a continuous concatenation $\oplus$ such that*

$$\sup_{t,s\in(0,\infty)} \left| r_{t\oplus s}(a,b) - r_t(a,b)\, r_s(a,b) \right| < \infty \tag{22}$$

for all $a, b \in X$.

This version of Boundedness retrospectively clarifies the meaning of the original axiom: the time combination of odds should not to be "infinitely far" from being a concatenation. Moreover, it permits to go beyond the case of linear λ's and characterize, for example, logarithmic and power ones. Before entering the details of our most delicate proof, we report a beautiful theorem of Aczel (1948) that characterizes continuous concatenations.

Theorem 19 (Aczel) *A binary operation $\oplus$ on $\mathbb{R}_+$ is a continuous concatenation if and only if there exists and increasing bijection $w : (0,\infty) \to (0,\infty)$ such that*

$$t \oplus s = w\left(w^{-1}(t) + w^{-1}(s)\right) \qquad \forall t, s > 0$$

This allows to rephrase Weak Boundedness as:

Ordinal Boundedness *There exists an increasing bijection $w : (0,\infty) \to (0,\infty)$ such that*

$$\sup_{t,s\in(0,\infty)} \left| r_{w(t+s)}(a,b) - r_{w(t)}(a,b)\, r_{w(s)}(a,b) \right| < \infty \tag{OB}$$

for all $a, b \in X$.

Ordinal Boundedness requires that time can be rescaled so that the corresponding variation of odds is not exponentially unbounded. It reduces to Boundedness when $w(t) = t$ for all t.[48]

[47] See, e.g., Falmagne (1985, Chapter 2).

[48] Ordinal Boundedness stands to Boundedness as the Ordinal IIA assumption of Fudenberg, Iijima, and Strzalecki (2015) stands to IIA.

Lemma 20 *A random choice process $\{p_t\}_{t\in(0,\infty)}$ satisfies Positivity, the Choice Axiom, Continuity, Consistency, Asymptotic Uniformity, and Ordinal Boundedness if and only if there exist $u : X \to \mathbb{R}$ and an increasing bijective $\lambda : (0,\infty) \to (0,\infty)$ such that*

$$p_t(a, A) = \frac{e^{\lambda(t)u(a)}}{\sum_{b\in A} e^{\lambda(t)u(b)}} \tag{23}$$

for all $A \in \mathcal{A}$, all $a \in A$, and all $t \in (0,\infty)$.

In this case,

(i) $p_0(a, A) = \dfrac{1}{|A|}$ *for all $A \in \mathcal{A}$ and all $a \in A$;*

(ii) $p_\infty(a, A) = \dfrac{1}{|\arg\max_A u|}\delta_a(\arg\max_A u)$ *for all $A \in \mathcal{A}$ and all $a \in A$;*

(iii) u is cardinally unique;

(iv) if u is non-constant, then λ is unique given u;

(v) $w = \lambda^{-1}$ is increasing, bijective, and $r_{\lambda^{-1}(t+s)}(a, b) = r_{\lambda^{-1}(t)}(a, b)\, r_{\lambda^{-1}(s)}(a, b)$ for all $a, b \in X$ and all $s, t \in (0,\infty)$, thus w realizes OB.

Proof Let $\{p_t\}$ be a random choice process that satisfies Positivity, the Choice Axiom, Continuity, Consistency, Asymptotic Uniformity, and Ordinal Boundedness. As in Lemma 18, we define, for all $a, b \in X$,

$$\varphi_{a,b}(t) = r_t(a, b) \qquad \forall t \in (0,\infty)$$

and we show that, thanks to the additional assumption of Ordinal Boundedness,

$$\varphi_{a,b}(w(t+s)) = \varphi_{a,b}(w(t))\, \varphi_{a,b}(w(s)) \qquad \forall t, s \in (0,\infty) \tag{24}$$

Three cases have to be considered, depending on whether $a \sim b$, $a \succ b$, or $b \succ a$ according to the weak order $\succsim$ defined in Lemma 18.

- If $a \sim b$, then $\varphi_{a,b}(t) = 1$ for all $t \in (0,\infty)$ and (24) holds.
- If $a \succ b$, then $\varphi_{a,b}$ is unbounded above and so is $\varphi_{a,b} \circ w : (0,\infty) \to (0,\infty)$. Moreover, by Ordinal Boundedness, there exists $M > 0$ such that

  $$|\varphi_{a,b}(w(t+s)) - \varphi_{a,b}(w(t))\, \varphi_{a,b}(w(s))| < M \qquad \forall t, s \in (0,\infty) \tag{25}$$

  But $(0,\infty)$ is a semigroup with respect to addition and $\varphi_{a,b} \circ w$ is unbounded above. Therefore, Baker (1980, Theorem 1) implies that (24) holds.
- Else, $b \succ a$, then the previous point shows

  $$\varphi_{b,a}(w(t+s)) = \varphi_{b,a}(w(t))\, \varphi_{b,a}(w(s))$$

  for all $t, s \in (0,\infty)$, but then

  $$\varphi_{a,b}(w(t+s)) = \frac{1}{\varphi_{b,a}(w(t+s))} = \frac{1}{\varphi_{b,a}(w(t))\, \varphi_{b,a}(w(s))} = \varphi_{a,b}(w(t))\, \varphi_{a,b}(w(s))$$

  for all $t, s \in (0,\infty)$, and (24) holds also in this case.

We conclude that the functional equation (24) holds for all $a, b \in X$. Continuity of $\varphi_{a,b} \circ w$, its strict positivity, and (24), imply that

$$\varphi_{a,b}(w(t)) = e^{h(a,b)t} \qquad \forall t \in (0, \infty)$$

for a unique $h(a, b) \in \mathbb{R}$ (see, e.g., Aczel, 1966, Theorem 2.1.2.1, p. 38). Setting $\lambda = w^{-1}$ it follows that $\varphi_{a,b}(s) = e^{h(a,b)\lambda(s)}$ for all $s \in (0, \infty)$, and λ is an increasing bijection like w.

Now fix some $a^* \in X$ and define $u : X \to \mathbb{R}$ by $u(x) = h(x, a^*)$ for all $x \in X$. Given any $t \in (0, \infty)$ and any $x, y \in X$, by (20),

$$\varphi_{x,y}(t) = \frac{v_t(x)}{v_t(y)} = \frac{v_t(x)}{v_t(a^*)} \frac{v_t(a^*)}{v_t(y)} = \frac{\varphi_{x,a^*}(t)}{\varphi_{y,a^*}(t)} = \frac{e^{h(x,a^*)\lambda(t)}}{e^{h(y,a^*)\lambda(t)}} = \frac{e^{u(x)\lambda(t)}}{e^{u(y)\lambda(t)}}$$

Therefore, for every $t \in (0, \infty)$, $A \in \mathcal{A}$, and $a \in A$, arbitrarily choosing $y \in A$,

$$p_t(a, A) = \frac{v_t(a)}{\sum_{b \in A} v_t(b)} = \frac{\frac{v_t(a)}{v_t(y)}}{\sum_{b \in A} \frac{v_t(b)}{v_t(y)}} = \frac{\frac{e^{u(a)\lambda(t)}}{e^{u(y)\lambda(t)}}}{\sum_{b \in A} \frac{e^{u(b)\lambda(t)}}{e^{u(y)\lambda(t)}}} = \frac{e^{\lambda(t)u(a)}}{\sum_{b \in A} e^{\lambda(t)u(b)}}$$

and (23) holds.

Annotation 1 *Notice that from the axioms we derived (23) with $\lambda = w^{-1}$.*

Since p_0 and p_∞ are defined by Continuity and λ vanishes as $t \to 0$ while it diverges at ∞ as $t \to \infty$, then (23) implies (i) and (ii).

As to uniqueness of u and λ, notice that, if also $\bar{u}$ and $\bar{\lambda}$ represent $\{p_t\}$ in the sense of (23), then

$$e^{\lambda(t)(u(a)-u(b))} = \frac{e^{\lambda(t)u(a)}}{e^{\lambda(t)u(b)}} = r_t(a, b) = \frac{p_t(a, b)}{p_t(b, a)} = e^{\bar{\lambda}(t)(\bar{u}(a)-\bar{u}(b))}$$

for all $t \in (0, \infty)$ and all $a, b \in X$. Therefore $\lambda(t)(u(a) - u(b)) = \bar{\lambda}(t)(\bar{u}(a) - \bar{u}(b))$ for all $t \in (0, \infty)$ and all $a, b \in X$. Choose $t = 1$ and arbitrarily fix $b^* \in X$ to conclude that

$$\bar{u}(a) = \frac{\lambda(1)}{\bar{\lambda}(1)}(u(a) - u(b^*)) + \bar{u}(b^*) = ku(a) + h \qquad \forall a \in A$$

with $k > 0$ and $h \in \mathbb{R}$. Point (iii) follows. If u is not constant, by choosing $a, b \in X$ with $u(a) \neq u(b)$, the previous argument, yields $\lambda(t)(u(a) - u(b)) = \bar{\lambda}(t)(ku(a) - ku(b))$ for all $t \in (0, \infty)$, so that $\bar{\lambda} = k^{-1}\lambda$. If $\bar{u} = u$ is given, then $k = 1$ and $\bar{\lambda} = \lambda$. This proves point (iv).

Clearly, λ^{-1} is increasing and bijective, and by (23)

$$r_{\lambda^{-1}(t+s)}(a, b) = r_{\lambda^{-1}(t)}(a, b)\, r_{\lambda^{-1}(s)}(a, b) \qquad \forall a, b \in X \quad \forall t, s \in (0, \infty)$$

which implies point (v).

The rest is trivial. ■

The equivalence of points 2 and 3 of **Theorem 6** is obtained by considering the special case in which $w(t) = t$ for all $t \in (0, \infty)$.[49] While that of points 1 and 3 is a routinary verification.

[49] See Annotation 1.

Logarithmic λ's, the role of which we discussed in the main text, correspond to Exponential Boundedness,[50] and give softmax the special form

$$p_t(a, A) = \frac{(1+\kappa t)^{u(a)}}{\sum_{b \in A} (1+\kappa t)^{u(b)}} \qquad \forall a \in A \in \mathcal{A} \quad \forall t \in (0, \infty) \tag{LSP}$$

where κ is a strictly positive constant. The analogous characterization of power λ's is left to the reader.

The next statement expresses the form **Proposition 7** takes for the (more general) Multinomial Logit Processes characterized in Lemma 20. In reading it, notice that points 2 and 3 can be equivalently stated in terms of λ or $w = \lambda^{-1}$.

Proposition 21 *Let X be a convex set and $\{p_t\}_{t\in(0,\infty)}$ be a Multinomial Logit Process with utility u and increasing bijective accuracy λ. The following conditions are equivalent:*

1. *u is affine;*
2. *there exists $t \in (0, \infty)$ such that*
$$p_t(a, b) = p_{\lambda^{-1}\left(\frac{\lambda(t)}{\alpha}\right)}(\alpha a + (1-\alpha) b, b) \tag{26}$$
for all $a, b \in X$ and all $\alpha \in (0, 1)$;
3. *given any $t \in (0, \infty)$ and any $c \in X$,*
$$p_t(a, A) = p_{\lambda^{-1}\left(\frac{\lambda(t)}{\alpha}\right)}(\alpha a + (1-\alpha) c, \alpha A + (1-\alpha) c)$$
for all $A \in \mathcal{A}$, all $a \in A$, and all $\alpha \in (0, 1)$.

Proof *2 implies 1.* Let $t \in (0, \infty)$ be such that (26) holds, then for all $a \neq b$ in X, and all $\alpha \in (0, 1)$,

$$\begin{aligned} p_{\lambda^{-1}\left(\frac{\lambda(t)}{\alpha}\right)}(b, \alpha a + (1-\alpha) b) &= 1 - p_{\lambda^{-1}\left(\frac{\lambda(t)}{\alpha}\right)}(\alpha a + (1-\alpha) b, b) \\ &= 1 - p_t(a, b) \\ &= p_t(b, a) \end{aligned}$$

whence

$$\begin{aligned} e^{\frac{\lambda(t)}{\alpha}[u(\alpha a + (1-\alpha)b) - u(b)]} &= r_{\lambda^{-1}\left(\frac{\lambda(t)}{\alpha}\right)}(\alpha a + (1-\alpha) b, b) \\ &= r_t(a, b) = e^{\lambda(t)[u(a)-u(b)]} \end{aligned}$$

[50] That is, Ordinal Boundedness with $w(t) = (e^t - 1)/\kappa$ for some strictly positive constant κ; which yields $\lambda(t) = \ln(1 + \kappa t)$ in (23), and corresponds to the superadditive concatenation $t \oplus s = t + s + \kappa ts$.

so

$$\frac{1}{\alpha}\left[u\left(\alpha a+(1-\alpha)b\right)-u\left(b\right)\right]=u\left(a\right)-u\left(b\right)$$
$$u\left(\alpha a+(1-\alpha)b\right)=\alpha u\left(a\right)+(1-\alpha)u\left(b\right)$$

which delivers affinity of u.

1 implies 3. Given any $t\in(0,\infty)$ and any $c\in X$, we have that, for all $A\in\mathcal{A}$, all $a\in A$, and all $\alpha\in(0,1)$,

$$\begin{aligned} p_{\lambda^{-1}\left(\frac{\lambda(t)}{\alpha}\right)}\left(\alpha a+(1-\alpha)c,\alpha A+(1-\alpha)c\right) &= \frac{e^{\frac{\lambda(t)}{\alpha}u(\alpha a+(1-\alpha)c)}}{\sum_{b\in A}e^{\frac{\lambda(t)}{\alpha}u(\alpha b+(1-\alpha)c)}}=\frac{e^{\frac{\lambda(t)}{\alpha}\alpha u(a)+\frac{\lambda(t)}{\alpha}(1-\alpha)u(c)}}{\sum_{b\in A}e^{\frac{\lambda(t)}{\alpha}\alpha u(b)+\frac{\lambda(t)}{\alpha}(1-\alpha)u(c)}} \\ &= \frac{e^{\lambda(t)u(a)}}{\sum_{b\in A}e^{\lambda(t)u(b)}}=p_t\left(a,A\right) \end{aligned}$$

3 implies 2 is trivial. ■

A.3 Proofs of the results of Section 4

We refer to Madras (2002) for notation and basic definitions regarding Markov Chain Monte Carlo methods.

Proof of Propositions 8 and 12 Once observed that

$$\alpha^{\lambda(t)}\left(a\mid b\right)=\frac{e^{\lambda(t)u(a)}}{e^{\lambda(t)u(a)}+e^{\lambda(t)u(b)}}\qquad\forall a\neq b\text{ in }A$$

(see Ratcliff, 1978), Proposition 8 is mathematically due to Barker (1965) who proves it in the context of plasma physics, here we report a simple proof for completeness. Given any $t,\lambda(t)\in(0,\infty)$, the explicit form of P_t is

$$P_t\left(a\mid b\right)=\begin{cases} Q\left(a\mid b\right)\dfrac{e^{\lambda(t)u(a)}}{e^{\lambda(t)u(a)}+e^{\lambda(t)u(b)}} & \text{if } a\neq b \\ 1-\sum_{c\in A\setminus\{b\}}Q\left(c\mid b\right)\dfrac{e^{\lambda(t)u(c)}}{e^{\lambda(t)u(c)}+e^{\lambda(t)u(b)}} & \text{if } a=b \end{cases}$$

and so P_t is irreducible because Q is. Moreover, again by irreducibility of Q,

$$\sum_{c\in A\setminus\{b\}}Q\left(c\mid b\right)>0\qquad\forall b\in A$$

(otherwise it would follow $Q\left(b\mid b\right)=1$ for some $b\in A$, violating irreducibility). But then $P_t\left(b\mid b\right)>0$ for all $b\in A$, which implies aperiodicity of P_t.

Next we show stationarity of $p_t\left(\cdot,A\right)$. Notice that, for all $a\neq b$ in A,

$$\begin{aligned} P_t\left(a\mid b\right)p_t\left(b,A\right) &= Q\left(a\mid b\right)\frac{e^{\lambda(t)u(a)}}{e^{\lambda(t)u(a)}+e^{\lambda(t)u(b)}}\frac{e^{\lambda(t)u(b)}}{\sum_{x\in A}e^{\lambda(t)u(x)}} \\ &= \frac{Q\left(a\mid b\right)}{\sum_{x\in A}e^{\lambda(t)u(x)}}\frac{e^{\lambda(t)(u(a)+u(b))}}{e^{\lambda(t)u(a)}+e^{\lambda(t)u(b)}}=P_t\left(b\mid a\right)p_t\left(a,A\right) \end{aligned}$$

while if $a = b$, then $P_t(a \mid b) p_t(b, A) = P_t(b \mid a) p_t(a, A)$ is obvious, thus

$$P_t(a \mid b) p_t(b, A) = P_t(b \mid a) p_t(a, A) \qquad \forall a, b \in A$$

Therefore, P_t is reversible with respect to $p_t(\cdot, A)$ and *a fortiori* $P_t p_t(\cdot, A) = p_t(\cdot, A)$ (see, e.g., Madras, 2002, Proposition 4.4).

The proof of Proposition 8 is concluded by observing that, since P_t is aperiodic and irreducible, then stationarity of $p_t(\cdot, A)$ implies $P_t^n \mu \to p_t(\cdot, A)$ as $n \to \infty$ for all $\mu \in \Delta(A)$ (see, e.g., Madras, 2002, Theorem 4.2).

Again by stationarity (*ibidem*), the long-run frequency of the sequence

$$(b_0, b_1, ..., b_n, b_{n+1}, ...) \tag{27}$$

of incumbents generated by the Metropolis-DDM algorithm is almost surely $p_t(\cdot, A)$. Moreover, if Q is uniform, by the Glivenko-Cantelli Theorem, the long-run frequency of the sequence

$$(a_1, ..., a_{n+1}, ...)$$

of proposals generated by the Metropolis-DDM algorithm is almost surely uniform. Simple algebra concludes the proof of Proposition 12. ■

In order to prove Proposition 9, we need a general result on renewal processes in the spirit of Wald's equality (see, e.g., Ross, 1970). Let $\{X_i\}_{i\in\mathbb{N}^*}$ be a sequence of nonnegative independent random variables with $\mathbb{E}(X_i) \leq \ell$ for all $i \in \mathbb{N}^* = \{1, 2, ...\}$, and set $X_0 \equiv 0$. Given any $n \in \mathbb{N} = \{0, 1, ...\}$, let

$$S_n = \sum_{i=0}^{n} X_i$$

and notice that $S_0 = 0$ and $S_n = \sum_{i=1}^{n} X_i$ if $n \in \mathbb{N}^*$. For each $t \in (0, \infty)$, define

$$N_t = \sup\{n \in \mathbb{N} : S_n \leq t\} \in \{0, 1, ..., \infty\}$$

The process $\{N_t\}_{t\in(0,\infty)}$ is the *renewal process* of $\{X_i\}_{i\in\mathbb{N}^*}$.

Annotation 2 *In our case, $\{N_t\}_{t\in(0,\infty)}$ represents the number of iterations performed by an algorithm whose i-th cycle has duration X_i, when the algorithm is stopped at time t.*

Given any $t \in (0, \infty)$, for each $n \in \mathbb{N}$,

$$N_t = n \iff X_0 + X_1 + ... + X_n \leq t < X_0 + X_1 + ... + X_{n+1} \iff S_n \leq t < S_{n+1} \tag{28}$$

and $N_t = \infty$ if and only if $\sum_{i=0}^{\infty} X_i \leq t$. This proves that N_t is a random variable, and that $N_t + 1$ is a stopping time with respect to $\{X_i\}_{i\in\mathbb{N}^*}$.[51]

[51] In fact, a random variable η with values in $\{1, ..., \infty\}$ is a *stopping time* with respect to $\{X_i\}_{i\in\mathbb{N}^*}$ if and only if $[\eta = n] \in \sigma(X_1, ..., X_n)$ for all $n \in \mathbb{N}^*$. Now

$$N_t + 1 = n \iff N_t = n - 1 \iff S_{n-1} \leq t < S_n$$

and both S_{n-1} and S_n are $\sigma(X_1, ..., X_n)$-measurable for all $n \in \mathbb{N}^*$.

A Wald-type inequality *Let* $\{X_i\}_{i\in\mathbb{N}^*}$ *be a sequence of nonnegative independent random variables with* $\mathbb{E}(X_i) \leq \ell$ *for all* $i \in \mathbb{N}^*$*. If* η *is a stopping time with respect to it, then*

$$\mathbb{E}\left(\sum_{i=1}^{\eta} X_i\right) \leq \ell\mathbb{E}(\eta)$$

and equality holds if $\mathbb{E}(X_i) = \ell$ *for all* $i \in \mathbb{N}^*$*.*

Proof Notice that, for each $i \in \{1, ..., \infty\}$,

$$1_{[\eta\geq i]} = \begin{cases} 1 & \text{if } i \leq \eta \\ 0 & \text{if } i > \eta \end{cases}$$

and moreover, for $i = 1$, we have $1_{[\eta\geq 1]} \equiv 1$. Then

$$\sum_{i=1}^{\eta} X_i = \sum_{i=1}^{\infty} X_i 1_{[\eta\geq i]}$$

and so $\sum_{i=1}^{\eta} X_i$ is a *bona fide* random variable; moreover, by nonnegativity and the Monotone Convergence Theorem,

$$\mathbb{E}\left(\sum_{i=1}^{\eta} X_i\right) = \mathbb{E}\left(\sum_{i=1}^{\infty} X_i 1_{[\eta\geq i]}\right) = \sum_{i=1}^{\infty} \mathbb{E}\left(X_i 1_{[\eta\geq i]}\right) = \mathbb{E}(X_1) + \sum_{i=2}^{\infty} \mathbb{E}\left(X_i 1_{[\eta\geq i]}\right)$$

But, for $i \geq 2$, $[\eta \geq i] = [\eta > i-1] = [\eta \leq i-1]^c$ and, since η is a stopping time,

$$[\eta \leq i-1] = [\eta = 1] \cup ... \cup [\eta = i-1] \in \sigma(X_1, ..., X_{i-1})$$

hence X_i and $1_{[\eta\geq i]}$ are independent; it follows that

$$\begin{aligned}\mathbb{E}(X_1) + \sum_{i=2}^{\infty} \mathbb{E}\left(X_i 1_{[\eta\geq i]}\right) &= \mathbb{E}(X_1) + \sum_{i=2}^{\infty} \mathbb{E}(X_i)\,\mathbb{E}\left(1_{[\eta\geq i]}\right) = \mathbb{E}(X_1) + \sum_{i=2}^{\infty} \mathbb{E}(X_i)\,\mathbb{P}([\eta \geq i]) \\ &= \sum_{i=1}^{\infty} \mathbb{E}(X_i)\,\mathbb{P}([\eta \geq i]) \leq \sum_{i=1}^{\infty} \ell\mathbb{P}([\eta \geq i]) = \ell\sum_{i=1}^{\infty} \mathbb{P}([\eta > i-1]) \\ &= \ell\sum_{j=0}^{\infty} \mathbb{P}([\eta > j]) = \ell\mathbb{E}(\eta)\end{aligned}$$

where the last equality follows from Lemma 1.1.1 of Resnick (2002).[52] ■

Theorem 22 *Let* $\{X_i\}_{i\in\mathbb{N}^*}$ *be a sequence of nonnegative independent random variables with* $\mathbb{E}(X_i) \leq \ell$ *for all* $i \in \mathbb{N}^*$*. If* $\ell > 0$*, then*

$$\mathbb{E}(N_t) \geq \frac{t}{\ell} - 1 \tag{29}$$

for all $t \in (0, \infty)$*.*

[52] And the only inequality is an equality if $\mathbb{E}(X_i) = \ell$ for all $i \in \mathbb{N}^*$.

Proof Arbitrarily choose $t \in (0, \infty)$ and recall that $N_t + 1$ is a stopping time with respect to $\{X_i\}_{i \in \mathbb{N}^*}$. If $\mathbb{E}(N_t) = \infty$, (29) is trivially true. Else N_t takes values in $\mathbb{N}$ with probability 1, then (28) guarantees that $\sum_{i=1}^{N_t+1} X_i > t$ almost surely. Together with the Waldean inequality, this implies

$$\ell \mathbb{E}(N_t + 1) \geq \mathbb{E}\left(\sum_{i=1}^{N_t+1} X_i\right) \geq t$$

which is equivalent to (29). ∎

Annotation 3 *This means that: if the i-th cycle of an algorithm has duration X_i with $\mathbb{E}(X_i) \leq \ell$ (for all $i \in \mathbb{N}^*$), and the algorithm is stopped at time t, then the expected number of iterations performed is at least $t/\ell - 1$.*

Proof of Proposition 9 The state space of the Metropolis-DDM algorithm consists of sequences

$$(\boldsymbol{s}, \boldsymbol{r}) = (b_0, a_1, r_1, b_1, ..., b_n, a_{n+1}, r_{n+1}, b_{n+1}, ...)$$

where

$$\boldsymbol{s} = (b_0, a_1, b_1, ..., b_n, a_{n+1}, b_{n+1}, ...) \in A^\infty$$

is the sequence of generated incumbents and proposals in A and

$$\boldsymbol{r} = (r_1, ..., r_{n+1}, ...) \in \mathbb{R}_+^\infty$$

is the sequence of realized response times. Endowed with its product Borel σ-algebra, $A^\infty \times \mathbb{R}_+^\infty$ is a probability space with respect to the the probability measure $\mathbb{P}$ whose marginal ν on A^∞ is the Markovian probability

$$\begin{aligned}
\nu(b_0, a_1) &= \mu(b_0) Q(a_1 \mid b_0) \\
\nu(b_n, a_{n+1} \mid b_{n-1}, a_n) &= \left(\delta_{a_n}(b_n) \frac{e^{\lambda(t)u(a_n)}}{e^{\lambda(t)u(a_n)} + e^{\lambda(t)u(b_{n-1})}} + \delta_{b_{n-1}}(b_n) \frac{e^{\lambda(t)u(b_{n-1})}}{e^{\lambda(t)u(a_n)} + e^{\lambda(t)u(b_{n-1})}}\right) Q(a_{n+1} \mid b_n)
\end{aligned}$$

and whose conditionals are

$$\mathbb{P}_{\boldsymbol{s}} = \delta_{\boldsymbol{s}} \times \left(\rho_{a_1,b_0} \times \rho_{a_2,b_1} \times \cdots \times \rho_{a_{n+1},b_n} \times \cdots\right) \qquad \forall \boldsymbol{s} \in A^\infty$$

where the second factor is the independent product of the response times' distributions $\rho_{a,b}$ of $\mathrm{RT}_{a,b}^z$ (see, e.g., Chang and Pollard, 1997, for the technical details of conditioning and disintegration).

Denoting

$$\begin{array}{rccc}
X_i: & A^\infty \times \mathbb{R}_+^\infty & \to & \mathbb{R}_+ \\
 & (\boldsymbol{s}, \boldsymbol{r}) & \mapsto & r_i
\end{array}$$

for all $i \in \mathbb{N}^*$, the number of iterations of the "repeat-until" loop of the Metropolis-DDM algorithm is, for each $(\boldsymbol{s}, \boldsymbol{r}) \in A^\infty \times \mathbb{R}_+^\infty$,

$$\begin{aligned}
N_t(\boldsymbol{s}, \boldsymbol{r}) &= \sup\{n \in \mathbb{N} : r_0 + r_1 + \cdots + r_n \leq t\} \\
&= \sup\{n \in \mathbb{N} : X_0(\boldsymbol{s}, \boldsymbol{r}) + X_1(\boldsymbol{s}, \boldsymbol{r}) + \cdots + X_n(\boldsymbol{s}, \boldsymbol{r}) \leq t\}
\end{aligned}$$

where by convention $X_0(\boldsymbol{s},\boldsymbol{r}) \equiv r_0 \equiv 0$. By expression (10), for any given $\boldsymbol{s} \in A^\infty$,

$$\mathbb{E}_{\boldsymbol{s}}(X_i) = \mathbb{E}\left(\mathrm{RT}^z_{a_i,b_{i-1}}\right) \leq \frac{\lambda^2(t)}{2} \qquad \forall i \in \mathbb{N}^*$$

and the X_i's are independent with respect to $\mathbb{P}_{\boldsymbol{s}}$, then, by Theorem 22,

$$\mathbb{E}_{\boldsymbol{s}}(N_t) \geq \frac{2t}{\lambda^2(t)} - 1$$

Expression (12) follows by the law of total expectation. ■

Proof of Proposition 10 Arbitrarily choose $t \in (0,\infty)$. Let q be the minimum off-diagonal entry of Q and $\xi = \max_{a,b\in A}|u(a) - u(b)|$ be the oscillation of u on A. If $b = a$, then

$$P_t(a \mid b) = 1 - \sum_{c\in A\setminus\{b\}} Q(c \mid b)\frac{1}{1+e^{\lambda(t)[u(b)-u(c)]}}$$

but, for all $c \in A$, $u(c) - u(b) \leq \xi$ and so $u(b) - u(c) \geq -\xi$ and $e^{\lambda(t)[u(b)-u(c)]} \geq e^{-\lambda(t)\xi}$, thus

$$\frac{1}{1+e^{\lambda(t)[u(b)-u(c)]}} \leq \frac{1}{1+e^{-\lambda(t)\xi}}$$

hence $\sum_{c\in A\setminus\{b\}} Q(c \mid b)\left(1+e^{\lambda(t)[u(b)-u(c)]}\right)^{-1} \leq \sum_{c\in A\setminus\{b\}} Q(c \mid b)\left(1+e^{-\lambda(t)\xi}\right)^{-1} \leq \left(1+e^{-\lambda(t)\xi}\right)^{-1}$, thus

$$P_t(a \mid b) \geq 1 - \frac{1}{1+e^{-\lambda(t)\xi}} = \frac{e^{-\lambda(t)\xi}}{1+e^{-\lambda(t)\xi}} = \frac{1}{1+e^{\lambda(t)\xi}} \geq \frac{q}{1+e^{\lambda(t)\xi}}$$

Since the same inequality clearly holds if $b \neq a$, we can apply Doeblin's Theorem to obtain

$$\sum_{c\in A}\left|[P_t^n\mu](c) - p_A^{(u,\lambda(t))}(c)\right| \leq 2\left(1 - \frac{q}{1+e^{\lambda(t)\xi}}\right)^n \qquad \forall n \in \mathbb{N}$$

In particular,

$$\begin{aligned}
\frac{1}{2}\sum_{c\in A}\left|\left[P_t^{N_t}\mu\right](c) - p_A^{(u,\lambda(t))}(c)\right| &\leq \left(1 - \frac{q}{1+e^{\lambda(t)\xi}}\right)^{N_t} \\
\ln\frac{1}{2}\sum_{c\in A}\left|\left[P_t^{N_t}\mu\right](c) - p_A^{(u,\lambda(t))}(c)\right| &\leq N_t \ln\left(1 - \frac{q}{1+e^{\lambda(t)\xi}}\right) \\
\frac{\ln\frac{1}{2}\sum_{c\in A}\left|\left[P_t^{N_t}\mu\right](c) - p_A^{(u,\lambda(t))}(c)\right|}{\ln\left(1 - \frac{q}{1+e^{\lambda(t)\xi}}\right)} &\geq N_t
\end{aligned}$$

since $1 - q\left(1+e^{\lambda(t)\xi}\right)^{-1} \in (0,1)$. By Proposition 9,

$$\mathbb{E}(N_t) \geq \frac{t}{\lambda(t)^2} - 1$$

but then

$$\frac{\mathbb{E}\left(\ln\frac{1}{2}\sum_{c\in A}\left|\left[P_t^{N_t}\mu\right](c)-p_A^{(u,\lambda(t))}(c)\right|\right)}{\ln\left(1-\frac{q}{1+e^{\lambda(t)\xi}}\right)}\geq\frac{t}{\lambda(t)^2}-1$$

$$\mathbb{E}\left(\ln\frac{1}{2}\sum_{c\in A}\left|\left[P_t^{N_t}\mu\right](c)-p_A^{(u,\lambda(t))}(c)\right|\right)\leq\left(\frac{t}{\lambda(t)^2}-1\right)\ln\left(1-\frac{q}{1+e^{\lambda(t)\xi}}\right)$$

$$\exp\mathbb{E}\left(\ln\frac{1}{2}\sum_{c\in A}\left|\left[P_t^{N_t}\mu\right](c)-p_A^{(u,\lambda(t))}(c)\right|\right)\leq\left(1-\frac{q}{1+e^{\lambda(t)\xi}}\right)^{\frac{t}{\lambda(t)^2}-1}$$

and standard calculus techniques show that the latter quantity vanishes as $t\to\infty$ because of the assumptions we made on λ and the way in which we defined q and ξ. ■

Proof of Proposition 11 Without loss of generality assume $i=1$. Then

$$\begin{aligned}q_t(a_1,A)&=\mathbb{P}\left(u(a_1)t+\sigma W_{a_1}(t)>u(a_j)t+\sigma W_{a_j}(t)\quad\text{for all } j=2,...,k\right)\\&=\mathbb{P}\left(\frac{W_{a_j}(t)}{\sqrt{t}}-\frac{W_{a_1}(t)}{\sqrt{t}}<\frac{\sqrt{t}}{\sigma}\left[u(a_1)-u(a_j)\right]\quad\text{for all } j=2,...,k\right)\\&=\mathbb{P}\left(X_j-X_1<\frac{\sqrt{t}}{\sigma}\left[u(a_1)-u(a_j)\right]\quad\text{for all } j=2,...,k\right)\end{aligned}$$

where $X_j=W_{a_j}(t)/\sqrt{t}$ for $j=1,2,...,k$ are a family of i.i.d. standard normal random variables. It is then easy to show that $(Y_j=X_j-X_1:j=2,...,k)$ is a $k-1$ dimensional vector of normal zero-mean random variables such that

$$\operatorname{cov}(Y_i,Y_j)=\mathbb{E}\left(X_1^2\right)=1\qquad\text{if } i\neq j$$

and $\operatorname{cov}(Y_i,Y_i)=\operatorname{var}(X_i-X_1)=2$ otherwise. ■

References

Aczel, J. (1948). Sur les opérations définies pour nombres réels. *Bulletin de la Societe Mathematique de France*, 76, 59-64.

Aczel, J. (1966). *Lectures on functional equations and their applications.* Academic Press.

Agranov, M., Caplin, A., and Tergiman, C. (2015). Naive play and the process of choice in guessing games. *Journal of the Economic Science Association*, 1, 146-157.

Anderson, S. P., Goeree, J. K., and Holt, C. A. (2004). Noisy directional learning and the logit equilibrium. *The Scandinavian Journal of Economics*, 106, 581-602.

Arrow, K. J. (1959). Rational choice functions and orderings. *Economica*, 26, 121-127.

ALQahtani, D. A., Rotgans, J. I., Mamede, S., ALAlwan, I., Magzoub, M. E. M., Altayeb, F. M., Mohamedani MA, and Schmidt, H. G. (2016). Does time pressure have a negative effect on diagnostic accuracy? *Academic Medicine*, 91, 710-716.

Baldassi, C., Cerreia-Vioglio, S., Maccheroni, F., and Marinacci, M. (2017). Simulated decision processes. Mimeo.

Baker, J. A. (1980). The stability of the cosine equation. *Proceedings of the American Mathematical Society*, 80, 411-416.

Barker, A. A. (1965). Monte Carlo calculations of the radial distribution functions for a proton-electron plasma. *Australian Journal of Physics*, 18, 119-134.

Becker, G. S. (1962). Irrational behavior and economic theory. *Journal of Political Economy*, 70, 1-13.

Ben-Akiva, M. E., and Lerman, S. R. (1985). *Discrete choice analysis: theory and application to travel demand.* MIT Press.

Ben-Akiva, M., and Morikawa, T. (1990). Estimation of switching models from revealed preferences and stated intentions. *Transportation Research*, 24A, 485-495.

Ben-Akiva, M., and Morikawa, T. (1991). Estimation of travel demand models from multiple data sources. In Koshi, M. (Ed.). *Transportation and Traffic Theory, Proceedings of the 11th ISTTT* (pp. 461-476). Elsevier.

Bhat, C. R. (1995). A heteroscedastic extreme value model of intercity travel mode choice. *Transportation Research*, 29B, 471-483.

Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, 113, 700-765.

Bogacz, R., Usher, M., Zhang, J., and McClelland, J. L. (2007). Extending a biologically inspired model of choice: multi-alternatives, nonlinearity and value-based multidimensional choice. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 362, 1655-1670.

Bradley, M., and Daly, A. (1994). Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data. *Transportation*, 21, 167-184.

Campbell, D., Boeri, M., Doherty, E., and Hutchinson, W. G. (2015). Learning, fatigue and preference formation in discrete choice experiments. *Journal of Economic Behavior and Organization*, 119, 345-363.

Caplin, A., and Dean, M. (2011). Search, choice, and revealed preference. *Theoretical Economics*, 6, 19-48.

Caplin, A., and Dean, M. (2013). Behavioral implications of rational inattention with Shannon entropy. NBER Working Paper 19318.

Cattell, J. M. (1902). The time of perception as a measure of differences in intensity. *Philosophische Studien*, 19, 63-68.

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., and Rustichini, A. (2016). Law of demand and forced choice. *IGIER Working Papers*, 593, 2016.

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., and Rustichini, A. (2018). Law of demand and stochastic choice. Mimeo.

Chang, J. T., & Pollard, D. (1997). Conditioning as disintegration. Statistica Neerlandica, 51(3), 287-317.

Chen, C., Chorus, C., Molin, E., and van Wee, B. (2016). Effects of task complexity and time pressure on activity-travel choices: heteroscedastic logit model and activity-travel simulator experiment. *Transportation*, 43, 455-472.

Davidson, D., and Marschak, J. (1959). Experimental tests of a stochastic decision theory. In Marschak, J. (Ed.) (1974). *Economic information, decision, and prediction* (pp. 133-171). Reidel Publishing Company.

Diederich, A., and Busemeyer, J. R. (2003). Simple matrix methods for analyzing diffusion models of choice probability, choice response time, and simple response time. Journal of Mathematical Psychology, 47, 304-322.

Ditterich, J. (2010). A comparison between mechanisms of multi-alternative perceptual decision making: ability to explain human behavior, predictions for neurophysiology, and relationship with decision theory. *Frontiers in Neuroscience*, 4.

Dovencioglu, D., Ban, H., Schofield, A. J., and Welchman, A. E. (2013). Perceptual integration for qualitatively different 3-D cues in the human brain. *Journal of Cognitive Neuroscience*, 25(9), 1527-1541.

Echenique, F., and Saito, K. (2018). General Luce model. Mimeo.

Falmagne, J. C. (1985). *Elements of psychophysical theory*. Oxford University Press.

Fehr, E., and Rangel, A. (2011). Neuroeconomic foundations of economic choice—recent advances. *The Journal of Economic Perspectives*, 25, 3-30.

Frederic, P., Di Bacco, M., and Lad, F. (2012). Combining expert probabilities using the product of odds. *Theory and Decision*, 73, 1-15.

Frick, M., Iijima, R., and Strzalecki, T. (2017). Dynamic random utility. Mimeo.

Fudenberg, D., Iijima, R., and Strzalecki, T. (2015). Stochastic choice and revealed perturbed utility. *Econometrica*, 83, 2371-2409.

Fudenberg, D., Strack, P., and Strzalecki, T. (2017). Speed, accuracy, and the optimal timing of choices. Mimeo.

Fudenberg, D., and Strzalecki, T. (2015). Dynamic logit with choice aversion. *Econometrica*, 83, 651-691.

Gold, J. I., and Shadlen, M. N. (2002). Banburismus and the brain: decoding the relationship between sensory stimuli, decisions, and reward. *Neuron*, 36, 299-308.

Gold, J. I., and Shadlen, M. N. (2007). The neural basis of decision making. *Annual review of neuroscience*, 30, 535-574.

Goeree, J. K., Holt, C. A., and Palfrey, T. R. (2016). *Quantal response equilibrium: a stochastic theory of games.* Princeton University Press.

Hensher, D. A., and Bradley, M. (1993). Using stated response choice data to enrich revealed preference discrete choice models. *Marketing Letters*, 4, 139-151.

Hensher, D., Louviere, J., and Swait, J. (1999). Combining sources of preference data. *Journal of Econometrics*, 89, 197-221.

Iverson, G., and Luce, R. D. (1998). The representational measurement approach to psychophysical and judgmental problems. In Birnbaum, M. H. (Ed.). *Measurement, judgment, and decision making* (pp. 1-79). Elsevier

Karsilar, H., Simen, P., Papadakis, S., and Balci, F. (2014). Speed accuracy trade-off under response deadlines. *Frontiers in Neuroscience*, 8, Article 248.

Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983). Optimization by simulated annealing. *Science*, 220, 671-680.

Krajbich, I., Armel, C., and Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature neuroscience*, 13, 1292-1298.

Krajbich, I., and Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences*, 108, 13852-13857.

Laming, D. R. J. (1968). *Information theory of choice-reaction times.* Wiley.

Link, S. W. (1978). The relative judgement theory analysis of response time deadline experiments. In Castellan, N. J., and Restle, F. (Eds.). *Cognitive Theory*, vol. 3 (pp. 117-138). Wiley.

Louviere, J. J., Hensher, D. A., and Swait, J. D. (2000). *Stated choice methods: analysis and applications.* Cambridge University Press.

Lu, J. (2016). Random choice and private information. *Econometrica*, 84, 1983-2027.

Luce, R. D. (1959). *Individual choice behavior: a theoretical analysis.* Wiley.

Luce, R. D., and Suppes, P. (2002). Representational measurement theory. In Yantis, S. (Ed.). *Stevens' handbook of experimental psychology* (pp. 1-41). Wiley.

Luck, S. J., and Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, 390, 279-281.

Madras, N. N. (2002). *Lectures on Monte Carlo methods.* American Mathematical Society.

Matejka, F., and McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105, 272-298.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. In Zarembka, P. (Ed.). *Frontiers in econometrics* (pp. 105-142). Academic Press.

McKelvey, R. D., and Palfrey, T. R. (1995). Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10, 6-38.

McMillen, T., and Holmes, P. (2006). The dynamics of choice among multiple alternatives. *Journal of Mathematical Psychology*, 50, 30-57.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21, 1087-1092.

Milosavljevic, M., Malmaud, J., Huth, A., Koch, C., and Rangel, A. (2010). The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judegment and Decision Making*, 5, 437-449.

Milosavljevic, M., Navalpakkam, V., Koch, C., and Rangel, A. (2012). Relative visual saliency differences induce sizable bias in consumer choice. *Journal of Consumer Psychology*, 22, 67-74.

Mosteller, F., and Nogee, P. (1951). An experimental measurement of utility. *Journal of Political Economy*, 59, 371-404.

Natenzon, P. (2017). Random choice and learning. *Journal of Political Economy*, forthcoming.

Ortega, P, and Stocker, A. A. (2016). Human decision-making under limited time. Proceedings of the *NIPS 2016 Conference.* MIT Press.

Plott, C. R. (1996). Rational individual behavior in markets and social choice processes: the discovered preference hypothesis. In Arrow, K., Colombatto, E., Perlaman, M., and Schmidt, C. (Eds.). *The rational foundations of economic behavior* (pp. 225-250). Macmillan.

Proto, E., Rustichini, A., and Sofianos, A. (2017). Intelligence, personality and gains from cooperation in repeated interactions. *Journal of Political Economy*, forthcoming.

Rangel, A., and Clithero, J. A. (2014). The computation of stimulus values in simple choice. In Glimcher, P. W., and Fehr, E. (Eds.). *Neuroeconomics* (pp. 125-148). Academic Press.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59-108.

Ratcliff, R., and Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological review*, 111, 333-367.

Ratcliff, R., Smith, P. L., Brown, S. D., and McKoon, G. (2016). Diffusion decision model: current issues and history. *Trends in Cognitive Sciences*, 20, 260-281.

Renyi, A. (1956). On conditional probability spaces generated by a dimensionally ordered set of measures. *Theory of Probability and Its Applications*, 1, 55-64.

Resnick, S. I. (2002). *Adventures in stochastic processes.* Springer.

Reutskaja, E., Nagel, R., Camerer, C. F., and Rangel, A. (2011). Search dynamics in consumer choice under time pressure: An eye-tracking study. *The American Economic Review*, 101, 900-926.

Rieskamp, J., Busemeyer, J. R., and Mellers, B. A. (2006). Extending the bounds of rationality: evidence and theories of preferential choice. *Journal of Economic Literature*, 44, 631-661.

Roe, R. M., Busemeyer, J. R., and Townsend, J. T. (2001). Multialternative decision field theory: A dynamic connectionst model of decision making. *Psychological Review*, 108, 370.

Ross, S. M. (1970). *Applied probability models with optimization applications.* Dover.

Russo, J. E., and Rosen, L. D. (1975). An eye fixation analysis of multialternative choice. *Memory & Cognition*, 3, 267-276.

Rustichini, A., and Padoa-Schioppa, C. (2015). A neuro-computational model of economic decisions. *Journal of Neurophysiology*, 114, 1382-1398.

Saito, K. (2017). Axiomatizations of the Mixed Logit Model. Mimeo.

Savage, S. J., and Waldman, D. M. (2008). Learning and fatigue during choice experiments: a comparison of online and mail survey modes. *Journal of Applied Econometrics*, 23, 351-371.

Shadlen, M. N., and Shohamy, D. (2016). Decision making and sequential sampling from memory. *Neuron*, 90, 927-939.

Steiner, J., Stewart, C., and Matejka, F. (2017). Rational inattention dynamics: inertia and delay in decision-making. *Econometrica*, 85, 521-553.

Swait, J., and Louviere, J. (1993). The role of the scale parameter in the estimation and comparison of multinomial logit models. *Journal of marketing research*, 305-314.

Train, K. E. (2009). *Discrete choice methods with simulation.* Cambridge University Press.

Tsunada, J., Liu, A. S., Gold, J. I., and Cohen, Y. E. (2016). Causal contribution of primate auditory cortex to auditory perceptual decision-making. *Nature Neuroscience*, 19, 135.

Vass, C. M., Wright, S., Burton, M., and Payne, K. (2018). Scale Heterogeneity in Healthcare Discrete Choice Experiments: A Primer. *The Patient-Patient-Centered Outcomes Research*, 11, 167-173.

Vogel, E. K., and Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, 428, 748-751.

Webb, R. (2018). The (neural) dynamics of stochastic choice, *Management Science*, forthcoming.

Woodford, M. (2014). Stochastic choice: an optimizing neuroeconomic model. *American Economic Review*, 104, 495-500.