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

Simone Cerreia-Vioglio, Fabio Maccheroni, Massimo Marinacci
Università Bocconi and IGIER

Aldo Rustichini
University of Minnesota and IGIER

June 30, 2022

Abstract

We provide two characterizations, one axiomatic and the other neuro-computational, of the dependence of choice probabilities on deadlines, within the widely used *softmax* representation

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

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, λ is a time-dependent noise parameter measuring the unit cost of information, u is a time-independent utility function, and α is an alternative-specific bias that determines the initial choice probabilities (reflecting prior information and memory anchoring).

Our axiomatic analysis provides a behavioral foundation of softmax (also known as Multinomial Logit Model when α is constant). Our neuro-computational derivation provides a biologically inspired algorithm that may explain the emergence of softmax in choice behavior. Jointly, the two approaches provide a thorough understanding of softmaximization in terms of internal causes (neuro-physiological mechanisms) and external effects (testable implications).

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

1 Introduction

Human decisions are often made under pressing time limits that substantially affect choice. Think of a trader deciding among alternative investments in fast moving financial markets, a triage nurse screening patients in life-threatening conditions, a team player under pressure choosing an action in a split second, a civic official determining whether to evacuate an area because of an impending natural hazard.

Choice under time constraint In all of these examples, the decision maker is given an exogenously constrained deliberation time during which he can gather and process noisy information about alternatives, whose qualities are *ex ante* only imperfectly known. This binding constraint may prevent the decision maker from fully learning the values of alternatives and select with certainty the best course of action.

The stochastic nature of information acquisition translates into stochastic choice behavior: when facing the same choice problem in different occasions, the decision maker might gather different information and choose different alternatives. We study how the choice distribution over available alternatives varies as a function of deliberation time.

As an additional illustration, more amenable to laboratory analysis than the initial field examples, consider the choice of food. In many empirical studies of food decision problems, a specific product a is identified with the vector $a = (a_1, a_2, ..., a_n)$ of its attributes, which may include price, brand, origin, preparation method, nutrition facts, and so on. A few papers show that the relative importance of attributes for consumer choice depends on the time available to decide.[1] In a nutshell, when no time is available to choose, consumers resort to habit or make their decisions based on the immediately available information, say price and brand (a_1, a_2), the so called "package attributes." Instead when information acquisition is possible, their choice is informed by all of the attributes $(a_1, a_2, ..., a_n)$ and thus also less overt "label attributes" – such as origin, preparation method, and nutrition facts – play a role. As to why the resulting choice is stochastic, observe that the sampling of information from the label and from memory is typically random (e.g., Shadlen and Shohamy, 2016), thus the chosen food items may differ in different choice episodes.

Our analysis will produce a *softmax* stochastic choice in which the probability of choosing alternative a from a menu A depends on two key parameters, the slope with respect to the utility of the option and the intercept. More precisely, the representation we will identify is:

$$p_t(a, A) = \frac{e^{\frac{u(a)}{\lambda(t)} + \alpha(a)}}{\sum_{b \in A} e^{\frac{u(b)}{\lambda(t)} + \alpha(b)}} \tag{1}$$

Here $u(a)$ is the true, but ex ante unknown to the decision maker and to the analyst, subjective value of alternative a, $\lambda(t)$ is the cost of processing one unit of information in t seconds, and $\alpha(a)$ is the initial bias for alternative a. Bias captures the information that is immediately available and does not require processing. Such "prior information" is provided by two channels which anchor evaluation: memory and the immediate perception of the alternatives.[2] In our food choice example, attributes such as price and brand are evident, while origin, preparation, and nutrition facts are more difficult to detect and interpret. This suggests that α might depend only on the immediate attributes that can be effortlessly appreciated,[3] while u reasonably depends on all of them.

The functional form (1) extends well-known representations. When the initial bias is absent – i.e., when α is constant – formula (1) reduces to the *Multinomial Logit* specification. In general, under the natural assumption that the unit cost of information processing $\lambda(t)$ decreases as more time is available to analyze evidence,[4] three scenarios emerge. First, the initial bias determines choice behavior when there is no deliberation time

$$p_0(a, A) = \lim_{t \to 0} p_t(a, A) = \frac{e^{\alpha(a)}}{\sum_{b \in A} e^{\alpha(b)}}$$

Second, at the opposite extreme, with no time pressure, only the best alternatives are selected

$$p_\infty(a, A) = \lim_{t \to \infty} p_t(a, A) > 0 \iff a \in \arg\max_A u$$

as prescribed by standard microeconomic theory. Third, under constrained but nonzero deliberation time t, an intermediate stochastic behavior results. It gives, as t increases, a higher chance – in the sense

[1] See Silayoi and Speece (2004), Grunert and Wills (2007), and the more recent Huseynov and Palma (2021).

[2] See Bordalo, Gennaioli, and Shleifer (2020) and Bordalo et al. (2021) for a recent analysis of the anchoring role of memory.

[3] With the attribute "brand" acting as a memory cue, see, e.g., Warlop, Ratneshwar, and van Osselaer (2005).

[4] With $\lambda(t) \to \infty$ as $t \to 0$, and $\lambda(t) \to 0$ as $t \to \infty$.

of stochastic dominance – of choosing better alternatives: with more time to decide, more observations become available and better estimates of values are possible.

Matejka and McKay (2015) showed that softmax stochastic choice behavior arises when a decision maker optimally processes information about the unknown "state of nature" u under an entropic unit cost of information $\lambda(t)$, with α depending on the decision maker's prior. Their study provides an important optimal information acquisition foundation for softmax behavior. In this paper, we study this behavior from two different, yet complementary, external and internal viewpoints that integrate their analysis. Next we outline them.

Outside the black box: psychometric axioms We provide a framework for the external, behavioral, study of an analyst who observes the choices of a decision maker and interprets them in the "as if" mode of revealed preference analysis. We show that softmax is captured by few simple behavioral axioms, Theorems 2 and 7, and that its parameters can be elicited from behavioral data, Theorems 3 and 8. In order to achieve the softmax representation, the classic Luce's axioms are obviously necessary, as it must be the case that

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

for all deadlines t, menus A, and alternatives a in A. Yet, these axioms cannot be sufficient because formula (1) imposes the additional separable structure

$$\mathrm{v}_t(a) = \frac{u(a)}{\lambda(t)} + \alpha(a) \tag{2}$$

on the Lucean v_t's. We thus need to introduce new axioms. They involve, for each deliberation time t, a standard *preference order* $\succ_t$ that ranks alternatives a and b, as well as two binary relations $\succ_t^\natural$ and $\succ_t^*$ that account, respectively, for the *intensity of preference* and for the *ease of comparison* across pairs of distinct alternatives (a, b) and (c, d). All these relations are revealed by the choice probabilities under different deadlines t, and so they can be retrieved from observables. Our axioms connect choice behavior at different t's by requiring the consistency of the relations $(\succ_t, \succ_t^\natural, \succ_t^*)$. For instance, consistency of the intensity of preference $\succ_t^\natural$ requires that, for any two deadlines $t < s$,

$$(a, b) \succ_t^\natural (c, d) \iff (a, b) \succ_s^\natural (c, d)$$

for all distinct pairs (a, b) and (c, d). In words, if deliberation for t seconds reveals stronger evidence in favor of the hypothesis "a is preferable to b" over the hypothesis "c is preferable to d," the same happens after a longer deliberation time s. The consistency requirements for the preference order and for the ease of comparison take similar forms.

While the elicitation of these three relations is somehow standard in the absence of bias (Section 2.1.1), based on the classic analyses of Debreu (1958) and Davidson and Marschak (1959); the exercise has never been performed in the presence of bias. To carry it out, we have to introduce novel elicitation methods to reveal $(\succ_t, \succ_t^\natural, \succ_t^*)$ from choice probabilities (Section 3.1.2).

The next issue that we study axiomatically is whether the overall performance in choice improves (or degrades) with the amount of time available to decide. In our representation (1), the question can be made precise, and amounts to the characterization of decreasing (or increasing) monotonicity of $\lambda(t)$. In the study of fast choice over food items, when t is small and the time constraint is binding, both the findings of Milosavljevic et al. (2010) and Reutskaja et al. (2011) suggest that the quality of choice (measured by comparing the actual choice with preferences elicited earlier among the items) becomes worse as time pressure increases.

Inside the black box: neuro-computational process Besides the external analysis just described, we also pursue an internal, neuro-computational, approach that provides a causal analysis of the decision maker choices through a biologically inspired algorithm that may explain softmax emergence in intelligent behavior. Our algorithm has a sequential structure, motivated by the well-known limits of human working memory and supported by eye-tracking studies – from the seminal Russo and Rosen (1975) to the recent Reutskaja et al. (2011). Our neuro-computational model naturally links multialternative choice with the classic Drift Diffusion Model of binary choice (DDM, hereafter) proposed by Ratcliff (1978). Specifically, we show that a procedure consisting of:

1. Markovian exploration of the menu,

2. sequential comparisons of pairs of alternatives, each based on a DDM,

3. choice of the incumbent at the deadline,

results in a stationary distribution that approximates softmax choice probabilities, Theorems 5 and 13. When binary comparisons are performed according to the symmetric DDM, this distribution is Multinomial Logit with no initial bias (Section 2.2.1). In contrast, a bias appears when general DDMs are considered. Again, the biased case goes beyond the standard analysis (see Bogacz et al. 2006) and requires a novel, formal, understanding of how past information is encoded, behaviorally and neurally, in the DDM (Section 3.2.1). We also present physiologically-calibrated simulations that support the biological plausibility of this neuro-computational foundation (Section 2.3).

The joint achievement The two approaches just outlined provide, along with the optimality analysis of Matejka and McKay (2015), a complete perspective on softmaximization as a model of preference discovery, in terms of internal causes (neuro-physiological mechanisms) and external effects (testable behavioral implications). In particular, our two approaches show that softmax:

1. is empirically testable and its parameters can be identified by behavioral data,

2. is plausible from a neural viewpoint.

The parameters u, λ, and α of the behavioral softmax process (1) that we axiomatize have neural counterparts in the neural softmax process that our algorithm generates. The integration of the inner and outer analyses allows us to empirically identify and cross-validate these, behavioral and neural, unobservable parameters. Thus, the two analyses complement each other conceptually, through a cause-effect nexus, as well as empirically (Section 2.3).

To ease the exposition, we first introduce in Section 2 the main concepts and ideas in the special case when the initial bias is absent, which corresponds to the Multinomial Logit Model and the symmetric DDM. They are widely studied specifications of softmax (in economics) and of the DDM (in neuroscience), and so in this section we can borrow concepts and techniques from the existing literature and go quicker to the point.

The analysis of the general case is, instead, significantly more demanding since it requires the introduction of new tools that allow us to disentangle the effects of past information and of evidence accumulation on observed choice behavior. This analysis is developed in Section 3. As anticipated, in this section new techniques to elicit the relations $(\succ_t, \succ_t^\natural, \succ_t^*)$ from choice probabilities are introduced and a novel connection between the softmax initial bias and the DDM starting point bias emerges.

We close this introduction by mentioning the relevance of our research for discrete choice analysis and by discussing the related literature.

Discrete choice analysis A byproduct of our analysis is an axiomatic foundation of the Heteroscedastic Multinomial Logit Model, the workhorse of discrete choice analysis.[5] Indeed, (1) can be rewritten in terms of random utility (see Luce and Suppes, 1965, and McFadden, 1973) as

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

where $\{\epsilon(a)\}_{a \in A}$ is a collection of independent errors with type I extreme value distribution, specific mean $\alpha(a)$, and common variance $\pi^2/6$. Here $p_t(a, A)$ describes the stochastic behavior of a decision maker who is trying to maximize u but, because of time pressure, makes mistakes in evaluating the various alternatives. The standard deviation of mistakes is proportional to $\lambda(t)$ and their bias is captured by α. In discrete choice analysis, t may be time or, more in general, an index describing the experimental conditions under which data have been collected (that is, the different data sets available to the analyst).[6] Heteroscedasticity, i.e., the dependence of λ on t and the presence of α, was introduced because, while the decision makers' utility u is a stable trait to be learned, disturbances are affected by experimental conditions and alternative-specific biases.

The present paper permits to test for misspecification of the Heteroscedastic Multinomial Logit Model and provides simple techniques to directly identify its parameters from data. In return, the discrete choice analysis literature provides a number of methods to estimate the parameters of the softmax specification (1).[7]

Related theory literature This paper considers exogenous deliberation times, thus we focus our discussion on the literature dealing with this issue.[8] Our treatment is axiomatic; to the best of our knowledge, there is only one other axiomatic foundation of the softmax model based on choice frequencies, due to Matejka and McKay (2015). The main difference is that they assume that the analyst knows the state that determines the decision maker's utility,[9] while we consider the general case in which the analyst may possibly ignore this state, or even the state space. Outside the laboratory, presuming such knowledge is a quite strong assumption.

As to the behavioral (external) analysis, in a Random Expected Utility perspective Lu (2016) axiomatically captures preference learning through increasingly informative priors on the set of probabilistic beliefs of the decision maker. Fudenberg and Strzalecki (2015) axiomatize a discounted adjusted logit model. Differently from our work, these papers study 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 the general random utility extension. Saito (2017) obtains several characterizations of the Mixed Logit Model. Finally, Baldassi et al. (2020a) and Fudenberg et al. (2020) axiomatize the value-based DDM.

As to the algorithmic (internal) analysis, while we consider a sequence of binary comparisons with evidence accumulation in each comparison, the vast majority of the extensions of the DDM to choice

[5] See, e.g., the textbooks of Louviere, Hensher, and Swait (2000) and Train (2009).

[6] For example, the indexes t represent locations in Train (2009, pp. 24-25), while they distinguish between stated intentions and market choices in Ben-Akiva and Morikawa (1990). In the working paper version of this paper, Cerreia-Vioglio et al. (2021a), we extend our axiomatic analysis to allow for completely general choice and index sets.

[7] The econometric study of this model, now textbook material, dates back to Ben-Akiva and Morikawa (1991), Swait and Louviere (1993), Hensher and Bradley (1993), and Bhat (1995).

[8] Models where decision time is endogenously − say, optimally − chosen are the subject of active research and we refer readers to Woodford (2014), Tajima, Drugowitsch, and Pouget (2016), Steiner, Stewart, and Matejka (2017), Fudenberg, Strack, and Strzalecki (2018), Callaway, Rangel, and Griffiths (2019), Tajima et al. (2019), Webb (2019), and Jang, Sharma, and Drugowitsch (2020) for updated perspectives.

[9] Choice situations of this kind have been studied since Saltzman and Garner (1948) and Kaufman et al. (1949). More recent contributions are Gabaix et al. (2006), Caplin and Dean (2014), Dean and Neligh (2019), and Dewan and Neligh (2020).

tasks with $|A| > 2$ alternatives considers simultaneous evidence accumulation for all the $|A|$ alternatives in the menu. In most studies, the choice task is assumed to simultaneously activate $|A|$ accumulators, each of which is primarily sensitive to one of the alternatives and integrates the evidence relative to that alternative. Choices are then based on absolute or relative evidence levels, with endogenous or exogenous stopping times. See, e.g., Roe, Busemeyer, and Townsend (2001), Anderson, Goeree, and Holt (2004), McMillen and Holmes (2006), Bogacz et al. (2007), Ditterich (2010), and Krajbich and Rangel (2011). Natenzon (2019) also belongs to this family and proposes a Multinomial Bayesian Probit model to jointly accommodate similarity, attraction, 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 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. Natenzon's paper is arguably the closest to ours. Like us, Natenzon builds on a parameterized family of random choice rules and tackles the problem of accounting for prior information and ease of comparison.[10] He also clearly spells out the interpretation of random utility on which, in the unbiased case, our identification between intensity of preference and intensity of evidence rests upon. At the same time, Natenzon provides testable implications but not a full characterization of his model, while we fully characterize ours. Moreover, he identifies the parameters of the model by assuming that the analyst observes choice from menus with phantom options, while we obtain identification by assuming that the analyst observes choice behavior at different deadlines.

Reutskaja et al. (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, select the best item encountered during the search (up to some noise). Their approach involves a "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. Although different from these models, and from those of Krajbich and Rangel (2011), our model is consistent with some of their experimental findings about the menu-exploration process and shares the reliance on the classical choice theory approach in which multialternative choice proceeds through binary comparison and elimination.

Finally, to the best of our knowledge, in the economics literature the first to showcase the fact that value based symmetric DDMs generate (binary) logit probabilities were Clithero (2018) and Webb (2019).

Organization of the paper Section 2 is dedicated to the unbiased case of multinomial logit processes. Section 3 studies the general case of softmax processes. Section 4 concludes. All proofs and additional simulations are relegated to the Online Appendices.

2 Unbiased analysis: multinomial logit processes

As previously discussed, the Multinomial Logit Model and the symmetric DDM both capture the case in which initial bias is absent, as Bogacz et al. (2006) and Matejka and McKay (2015) remark with essentially the same words from altogether different perspectives. With these models the psychometric complication of separating the effects of prior information and those of acquired evidence on observed choice is avoided, a simplification that permits a brisker exposition of the main results and ideas of our research, and allows us to borrow more from the literature. For this reason, in this section we focus on the Multinomial Logit Model (MNL, hereafter) and on the symmetric DDM.

[10] He also takes the similarity of alternatives into account, an important feature that we do not consider here.

In so doing we neglect, however, the fundamental role of memory and initial bias (and their pitfalls) in decision making with limited cognitive resources. In the next Section 3, we extend the analysis to general softmax processes to cope with these issues, and this will require us to tackle some novel, and nontrivial, conceptual and technical challenges.

2.1 Outside the black box: behavioral MNL

2.1.1 Preamble: random choice rules

Let $\mathcal{A}$ be the collection of all nonempty finite subsets A of a universal space X of possible alternatives, called *menus* (or *choice sets* or *choice problems*). We assume throughout that X is a connected topological space,[11] and we denote by $\Delta(X)$ the set of all probability measures on X with finite support.

A *random choice rule* is a function

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

such that $p(A, A) = 1$ for all A in $\mathcal{A}$. Given any alternative a in X, we interpret $p(\{a\}, A)$, simply denoted $p(a, A)$, as the probability that a decision maker chooses a when the set of available alternatives is A. More generally, if B is a subset of X, $p(B, A)$ is the probability $\sum_{b \in B} p(b, A)$ that the selected element lies in B. This probability can be viewed as the *frequency* with which an element in B is chosen. The requirement $p(A, A) = 1$ simply means that unavailable alternatives cannot be chosen. Indeed, it is equivalent to $p(x, A) = 0$ for all x that do not belong to A.

As usual, given any a and b in X, we indicate by

$$p(a, b) = p(a, \{a, b\}) \tag{3}$$

the probability with which a is chosen from the doubleton $\{a, b\}$. Luce (1959) proposes the most classical random choice model. Its main 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 Choice Axiom says that the probability of choosing an alternative a from menu A is that of first selecting B from A and then a from B. As observed by Luce, this amounts to require that $\{p(\cdot, A) : A \in \mathcal{A}\}$ is a conditional probability system in the sense of Renyi (1955).

As well known, both axioms can be expressed in terms of odds. In particular, the Choice Axiom is equivalent to the odds independence condition $p(a, b) / p(b, a) = p(a, A) / p(b, A)$, for all alternatives a and b in A and all menus A in $\mathcal{A}$, that requires the odds for a against b to be independent of the other alternatives available in the menu.[12]

A convenient continuity assumption is often automatically satisfied in applications.

Continuity *The function* $(a, b) \mapsto p(a, b)$ *is continuous on the set of all pairs of distinct alternatives in* X.

Next we state Luce's classic representation theorem.

---

[11] With at least three elements, e.g., a nonsingleton convex subset of $\mathbb{R}^n$. In the working paper version of this paper, Cerreia-Vioglio et al. (2021a), our analysis is extended to a general set X, even without a topology.

[12] This odds independence condition is often called *independence from irrelevant alternatives.* Cerreia-Vioglio et al. (2021b) characterizes the random choice rules that satisfy this condition but not necessarily Positivity.

Theorem 1 (Luce) *The following conditions are equivalent for a random choice rule* $p : \mathcal{A} \to \Delta(X)$*:*

1. *p satisfies Positivity, the Choice Axiom, and Continuity;*
2. *there exists a continuous* $\mathrm{v} : X \to \mathbb{R}$ *such that*

$$p(a, A) = \frac{e^{\mathrm{v}(a)}}{\sum_{b \in A} e^{\mathrm{v}(b)}} \tag{4}$$

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

Moreover, v *is unique up to location (i.e., up to translation by a constant).*

Under Luce's axioms, the function v is a *psychometric utility*, that is, it represents the *intensity of preference relation* $\rhd$ revealed by p. Formally, the relation $\rhd$ is defined by

$$(a, b) \rhd (c, d) \iff p(a, b) > p(c, d) \tag{5}$$

and the function v is such that

$$(a, b) \rhd (c, d) \iff \mathrm{v}(a) - \mathrm{v}(b) > \mathrm{v}(c) - \mathrm{v}(d) \tag{6}$$

for all $a \neq b$ and all $c \neq d$ in X. Discussed by Debreu (1958, p. 440), Luce (1959, p. 39), and especially Davidson and Marschak (1959, p. 237), the interpretation of $\rhd$ as the revealed intensity of preference is based on the following classic psychophysical principle, traditionally attributed to Cattell and Fullerton.

Discrimination principle *Equally often noticed differences are equal on the sensation scale, unless always or never noticed.*

In a random utility perspective, the Lucean psychometric utility $\mathrm{v}(a)$ of alternative a is unknown and a noisy signal $\mathrm{V}(a)$ is received.[13] Since $p(a, b) = \Pr\{\mathrm{V}(a) \geq \mathrm{V}(b)\}$, it follows that

$$p(a, b) > p(c, d) \iff \Pr\{\mathrm{V}(a) \geq \mathrm{V}(b)\} > \Pr\{\mathrm{V}(c) \geq \mathrm{V}(d)\} \tag{7}$$

that is, $(a, b) \rhd (c, d)$ if and only if the probability of receiving a signal in favor of a over b is higher than that of receiving a signal in favor of c over d. Intensity of preference is thus best understood in terms of intensity of evidence. The notions of this subsection are classic – e.g., the equivalence (7) dates back to Block and Marschak (1960, p. 110) and Luce and Suppes (1965, p. 338). Yet, this relationship between intensity of preference and intensity of evidence is less known and, as it will be seen later in the paper, is key for our analysis of preference discovery.

2.1.2 Axiomatic MNL

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

Definition 1 *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 a decision maker chooses alternative a from menu A if t is the *deliberation time*, that is, the maximum amount of time he is (exogenously) given to

[13] As observed by Natenzon (2019), this interpretation of random utility in which the value of options is perceived with noise is the one originally proposed by random utility models in psychology.

decide.[14] A random choice process thus describes the decision maker stochastic choice behavior under different levels of time pressure.

If each p_t satisfies Positivity, the Choice Axiom, and Continuity, then for each t in T there exists a psychometric utility $\mathrm{v}_t : X \to \mathbb{R}$ such that

$$p_t(a, A) = \frac{e^{\mathrm{v}_t(a)}}{\sum_{b \in A} e^{\mathrm{v}_t(b)}} \tag{8}$$

When this is the case, we say that $\{p_t\}_{t \in T}$ satisfies the *Psychometric Luce Axioms*. From there, to attain the MNL representation

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

we need to express all the functions v_t by means of a time-independent utility function u and an alternative-independent noise coefficient λ such that

$$\mathrm{v}_t(a) = \frac{u(a)}{\lambda(t)}$$

This is achieved through the next axiom which requires that, over deliberation times, there are no ordinal reversals in the preference intensities.

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

$$(a, b) \rhd_t (c, d) \iff (a, b) \rhd_s (c, d) \tag{9}$$

for all $a \neq b$ and $c \neq d$ in X.

In words, if the amount of evidence in favor of the hypothesis "a is preferable to b" is, after a given deliberation time, greater than that in favor of the hypothesis "c is preferable to d", the same happens after a longer deliberation time.

In terms of primitive observables, Consistency reads naturally: given any $s > t$ in T,

$$p_t(a, b) > p_t(c, d) \iff p_s(a, b) > p_s(c, d)$$

for all $a \neq b$ and $c \neq d$ in X. This is coherent with the idea that information depends on an underlying true value that does not change over time and is gradually discovered as time goes by. The next thought experiment illustrates such a concept.[15]

Example 1 Subjects are given t seconds to observe a QR code and take one of the following actions: choose b and receive a number of dollars equal to the number of *b*lack squares, choose w and receive a number of dollars equal to the number of *w*hite squares. Like in Caplin and Dean (2014, p. 59), they are also told that QR codes have been stochastically generated with an a priori equal chance of a black

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

[15] Notice that Consistency admits a formulation which does not rely on the order structure of T: *Given any s and t in T,*

$$(a, b) \rhd_t (c, d) \implies (a, b) \rhd_s (c, d) \tag{10}$$

for all $a \neq b$ and $c \neq d$ in X. This formulation allows to consider alternative interpretations of the index set T, and to extend our results accordingly.

majority or a white majority. Two different decision problems are portrayed below:

Problem 1: $\{b_1, w_1\}$ Problem 2: $\{b_2, w_2\}$

Figure 1 *Two decision problems.*

Introspection and experiments – starting with the seminal Thurstone (1929) to more recent random dot motion studies on humans, rats, and pigeons – suggest that, irrespective of t, $p_t(b_1, w_1) > p_t(b_2, w_2)$ because the true, but unknown, frequency of black squares in Problem 1 is bigger than that in Problem 2. It is from this frequency that agents sample for t seconds before choosing b or w in each problem. ▲

We are now ready to state our first representation theorem.

Theorem 2 *For a random choice process $\{p_t\}_{t\in T}$, the following conditions are equivalent:*

1. *$\{p_t\}$ satisfies the Psychometric Luce Axioms and Consistency;*
2. *there exist a continuous $u : X \to \mathbb{R}$ and a $\lambda : T \to (0, \infty)$ such that*
$$p_t(a, A) = \frac{e^{\frac{u(a)}{\lambda(t)}}}{\sum_{b\in A} e^{\frac{u(b)}{\lambda(t)}}} \tag{MNL}$$
for all $A \in \mathcal{A}$, all $a \in A$, and all $t \in T$.

In this case, u is a cardinally unique psychometric utility for p_t for all $t \in T$, and λ is unique given u, unless $\{p_t\}$ is uniform.[16]

The next result permits to elicit both the psychometric utility and the noise coefficient. Given any a and b in X and any t in T,

$$r_t(a, b) = \frac{p_t(a, b)}{p_t(b, a)} \qquad \text{and} \qquad \ell_t(a, b) = \ln r_t(a, b)$$

denote the *odds* for a against b – that is, the ratio between the number of episodes in which a is chosen and the number of episodes in which b is – and the *log-odds*, respectively. The latter are analytically convenient because they are positive if and only if odds are favorable to a.

Theorem 3 *A MNL process $\{p_t\}$ is not uniform if and only if there exist $\hat{a}, \hat{b} \in X$ and $\hat{t} \in T$ such that $p_{\hat{t}}(\hat{a}, \hat{b}) > p_{\hat{t}}(\hat{b}, \hat{a})$. In this case, the functions $\hat{u} : X \to \mathbb{R}$ and $\hat{\lambda} : T \to (0, \infty)$ given by*

$$\hat{u}(x) = \frac{\ell_{\hat{t}}(x, \hat{b})}{\ell_{\hat{t}}(\hat{a}, \hat{b})} \quad ; \quad \hat{\lambda}(t) = \frac{1}{\ell_t(\hat{a}, \hat{b})} \tag{11}$$

are well defined, with $\hat{u}$ continuous and

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

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

[16] A random choice process $\{p_t\}$ is *uniform* if $p_t(a, A) = 1/|A|$ for all $a \in A \in \mathcal{A}$ and all $t \in T$. In this case, formula (MNL) holds for any λ provided u is constant.

Together the last two results permit to test whether the MNL model is consistent with available choice data and to identify its parameters.

How can Consistency drive these results? The intuition is, in retrospect, simple. The Psychometric Luce Axioms guarantee that each v_t in (8) is a psychometric utility for the corresponding intensity of preference relation $\rhd_t$. Consistency, in turn, ensures that $\rhd_t$ coincides with $\rhd_s$. But, then, v_t has to be a positive affine transformation of v_s because of the cardinal uniqueness of psychometric utilities. This observation allows us to transform (8) into (MNL) and also ensures that the functions $\hat{u}$ and $\hat{\lambda}$ defined in (11) are "essentially independent" of the choice of $\hat{a}$, $\hat{b}$, and $\hat{t}$.

This answer, however, rises a few more questions. Why Consistency implies that alternatives are a priori homogeneous, with no initial bias? How should it be weakened to have α appear in the general softmax formula (1)? When is λ monotone/continuous? What is the limit behavior of $\{p_t\}$ as deliberation time diverges and so loses its binding constraint nature?

We will address all of these questions in the next section on general softmax processes. Here we close by characterizing the affinity of u, a case of relevance for both applications and experiments; for instance, when choice under risk is considered and X is a set of lotteries.[17] In the last lines of this subsection we assume that X is a convex subset of a topological vector space.

Stochastic Betweenness *There exists $t \in T$ such that*

$$p_t\left(a, \frac{1}{2}a + \frac{1}{2}b\right) = p_t\left(\frac{1}{2}a + \frac{1}{2}b, b\right)$$

for all $a \neq b$ in X.

In words, this axiom says that "equal probability differences are equally often noticed." It parallels its risk theory counterpart.

Proposition 4 *The following conditions are equivalent for a MNL process:*

1. *$\{p_t\}$ satisfies Stochastic Betweenness;*

2. *the psychometric utility u is affine.*

2.2 Inside the black box: neuro-computational MNL

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 MNL processes, but remain silent about what decision procedure may generate the corresponding choice probabilities. Here, we address this issue by combining DDM pairwise comparisons of alternatives and efficient Markovian search of the menu. Both assumptions, motivated by the well-known limits of working memory and by the seminal eye-tracking study of Russo and Rosen (1975), find support in recent theories about memory,[18] as well as in recent eye-tracking experiments on multialternative choice.[19]

Again, we consider a decision maker who has to choose an optimal alternative from a finite subset A of X 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 choice problem A, so it captures the overall level of time pressure.

The procedure we present can be summarized as follows. The decision marker's search:

[17] MNLs with affine utility functions are also used in discrete choice analysis, as well as in quantal response equilibrium theory (Section 4.4), and in most multicriteria decision analysis' applications. Recall that $u : X \to \mathbb{R}$ is affine when $u(\kappa x + (1-\kappa) y) = \kappa u(x) + (1-\kappa) u(y)$ for all κ in $[0,1]$ and all x, y in X.

[18] See, e.g., Luck and Vogel (1997), Vogel and Machizawa (2004), and Shadlen and Shohamy (2016).

[19] See, e.g., Krajbich and Rangel (2011) and Reutskaja et al. (2011).

1. starts from an arbitrary element b of the menu, the *incumbent*;
2. selects a candidate alternative a, the *proposal*;
3. compares them via the DDM, and makes the winner of the DDM comparison the new incumbent;
4. repeats steps 1-3 until deliberation time t comes and the current incumbent is chosen.

This is a basic variation over 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 pairwise comparisons are instantaneous and exact. In our procedure, 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.

2.2.1 Binary comparisons – the symmetric DDM

According to the DDM, when two alternatives a and b are compared, an alternative is selected as soon as the net evidence in its favor reaches a posited decision threshold β. Such a threshold represents the amount of evidence that the decision maker needs to decide and determines the speed-accuracy tradeoff: lower thresholds produce faster but less accurate responding, whereas higher thresholds produce more accurate but slower responses (see Bogacz et al., 2006). In experiments, high time pressure is known to reduce thresholds to speed up decisions.[20] In our setting, this means that the DDM decision thresholds for binary comparisons depend on the (overall) time constraint t for choice, that is, β is a function $\beta(t)$ of t.[21]

Specifically, the comparison of a and b is believed to activate two neuronal populations whose activities (firing rates) provide evidence for the two alternatives.[22] The mean activities $v(a)$ and $v(b)$ are interpreted as neural indexes of value of the alternatives. For this reason, we call $v : X \to \mathbb{R}$ the *neural utility* of the decision maker. The stochastic nature of firing rates is captured by the assumption that the actual activities of neuronal populations experience instantaneous and independent white noise fluctuations. Evidence accumulation in favor of a and b is thus represented by two uncorrelated Brownian motions with drift

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

Absence of bias corresponds to the fact that the normals $W_a(\tau)$ and $W_b(\tau)$ have zero mean (and variance τ). With these assumptions,[23] we have the *symmetric DDM* (*sDDM*), where:

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

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

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

[20] See, e.g., Milosavljevic et al. (2010) and Karsilar et al. (2014, especially p. 14).

[21] A *caveat:* $\beta(t)$ depends on the deadline t, but is constant given t within each binary comparison. Here we are not considering DDMs with "collapsing thresholds" like in Drugowitsch et al. (2012) and Fudenberg, Strack, and Strzalecki (2018). More on this in the concluding remarks (Section 4.3).

[22] See Bogacz et al. (2006) and Shadlen and Shohamy (2016) for the neuro-physiological and neuro-psychological analyses of this mechanism; and Roe, Busemeyer, and Townsend (2001), Krajbich, Armel, and Rangel (2010), Milosavljevic et al. (2010), and Rangel and Clithero (2014) for applications of the DDM to the choice of consumption goods.

[23] Which will be relaxed in Section 3.2, where more general DDM's are considered.

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

$$\mathrm{RT}_t(a,b) = \min\{\tau \in (0,\infty) : |Z_{a,b}(\tau)| = \beta(t)\}$$

- at which time, the alternative supported by evidence is selected; so, the *comparison outcome* is the random variable

$$\mathrm{CO}_t(a,b) = \begin{cases} a & \text{if } Z_{a,b}(\mathrm{RT}_t(a,b)) = \beta(t) \\ b & \text{if } Z_{a,b}(\mathrm{RT}_t(a,b)) = -\beta(t) \end{cases}$$

The presence of noise in evidence accumulation is what makes DDM comparisons time consuming and subject to error. In this regard, notice that the sDDM is an optimal Bayesian test of the hypothesis "$v(a) > v(b)$" that alternative a be preferable to alternative b.[24]

2.2.2 Proposal mechanism – efficient Metropolis exploration

The standard Markovian way to navigate a menu starts with a candidate solution b drawn from an initial distribution μ in $\Delta(A)$ and then, given an incumbent solution b, continues by considering an alternative candidate solution $a \neq b$, called proposal, with probability $Q(a \mid b)$. This exploration procedure was introduced by Metropolis et al. (1953) to calculate statistical mechanics equilibria,[25] and adopted in the *Simulated Annealing* heuristic of Kirkpatrick, Gelatt, and Vecchi (1983) to compute extrema in high dimensional combinatorial optimization problems.

The only formal requirements of Metropolis et al. (1953) on the probability transition matrix Q, called *exploration matrix*, are symmetry and irreducibility. At the same time, starting with the seminal Kirkpatrick, Gelatt, and Vecchi (1983), the adequate selection of the exploration matrix Q has been the object of fine tuning for the efficiency of optimization algorithms. The natural exploration matrix for a "traveling salesman" might not be equally compelling to describe our "spouse sent to buy wine 10 minutes before the arrival of the in-laws." In this case, time pressure compels towards comparisons of category "white or red?" rather than more sophisticated distinctions "Champagne or Crémant?"

The sDDM provides a specific suggestion for exploration under time pressure. As observed by Fudenberg, Strack, and Strzalecki (2018), the Bayesian decision problem that the sDDM solves features correct expectations on the time that it will take to compare two alternatives, *before* engaging in evidence accumulation to try and find the superior one. Easy choices will produce fast and accurate responses, while difficult ones will be time consuming and poorly efficient, absorbing a lot of time in exchange for a small payoff difference and presenting high error probabilities. These considerations suggest the adoption of an exploration matrix of the form

$$Q_t(a \mid b) \propto \frac{1}{\mathbb{E}[\mathrm{RT}_t(a,b)]} \qquad \forall a \neq b$$

where the probability with which a is proposed if b is the incumbent is inversely proportional to the expected response time. This choice adapts exploration to time constraints by giving priority to fast and accurate comparisons.[26]

Finally, the absence of initial biases assumed in this section suggests a uniform initial distribution μ.[27]

[24] See Gold and Shadlen (2002, 2007), Bogacz et al. (2006), Shadlen and Shohamy (2016), and especially Fudenberg, Strack, and Strzalecki (2018, Section II.B).

[25] See also Barker (1965).

[26] See Baldassi et al. (2020b) for a semi-Markov analysis of general exploration matrices.

[27] Mathematically, we can use any initial distribution. Yet, we find this specification conceptually compelling.

2.2.3 The Metropolis-DDM algorithm

We now combine sDDM pairwise comparisons and efficient exploration. The resulting procedure describes a decision maker who, given time t to decide, first automatically adjusts his evidence threshold $\beta(t)$ and then compares alternatives according to the sDDM. His search, starting with a random initial solution, is driven by speed and accuracy concerns, and it continues until time t is reached, at which point the incumbent solution is chosen.

Metropolis-DDM Algorithm

Input: *Given* $t > 0$.

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* $Q_t(\cdot \mid b_n)$ *and compare it to* b_n *via the sDDM, so:*

- *set* $\tau_{n+1} = \tau_n + \mathrm{RT}_t(a_{n+1}, b_n)$,
- *set* $b_{n+1} = \mathrm{CO}_t(a_{n+1}, b_n)$.

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

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

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

Notice that Q_t, RT_t, and CO_t all depend on both the neural utility v and the threshold level $\beta(t)$, which in turn depends on the deadline t. 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 by the stopping rule $\tau_{n+1} > t$ at the chosen alternative $b^* = b_n$.

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

$$\mathbb{P}_t(a, b) = \mathbb{P}[\mathrm{CO}_t(a, b) = a]$$

while a is rejected and the old incumbent b is maintained with the complementary probability $1 - \mathbb{P}_t(a, b)$. Therefore, the resulting probability of selecting a as a new incumbent given old incumbent b is

$$M_t(a \mid b) = Q_t(a \mid b)\, \mathbb{P}_t(a, b) \qquad \forall a \neq b \text{ in } A$$

This transition probability combines the stochasticity of the proposal mechanism and that of the acceptance/rejection rule. Specifically, if the algorithm stops at the $(n+1)$-th iteration, the resulting choice probabilities are given by the vector $M_t^n \mu$ in $\Delta(A)$.[28]

[28] Here $M_t = [M_t(a \mid b)]_{a,b \in A}$ is regarded as the probability transition matrix of a Markov chain with initial distribution μ, and M_t^n as its n-th power. Moreover, we write $M_t(A)$ and $Q_t(A)$ since these matrices depend on the menu A under consideration.

Theorem 5 *Let $v : X \to \mathbb{R}$ and $\beta : T \to (0, \infty)$. Given any $A \in \mathcal{A}$ and any $t \in T$, the exploration matrix $Q_t = Q_t(A)$ is irreducible and symmetric and the incumbents' transition matrix $M_t = M_t(A)$ is aperiodic, irreducible, and reversible, with stationary distribution*

$$m_t(a, A) = \frac{e^{\beta(t)v(a)}}{\sum_{b \in A} e^{\beta(t)v(b)}} \qquad \forall a \in A \tag{12}$$

In particular, $\lim_{n\to\infty} M_t^n(A)\mu = m_t(\cdot, A)$ for all $A \in \mathcal{A}$ and all $\mu \in \Delta(A)$.

What does this theorem say about the output of the Metropolis-DDM algorithm? Since the average duration of each iteration is bounded above by $\beta^2(t)/2$, then the average number of iterations is

$$\bar{n}_t \geq \frac{2t}{\beta^2(t)} - 1 \tag{13}$$

and the "average" output of the algorithm is

$$M_t^{\bar{n}_t}\mu \tag{14}$$

The distance between this output and the stationary distribution (12) decreases exponentially in $\bar{n}_t$.[29] Therefore, if $\beta^2(t)$ is small relative to t, then $\bar{n}_t$ is large and Theorem 5 guarantees that the choice probabilities produced by the Metropolis-DDM algorithm are essentially given by (12), which is a MNL distribution under the following identification

Inner		**Outer**	
Neural utility	v	Psychometric utility	u
Threshold	β	Inverse noise	$1/\lambda$

(15)

This important table connects our inside and outside analyses, as we discuss next.

2.3 Inside and outside the black box

Remarkably, the neural utility v and threshold parameter β appearing in the neuro-computational specification (12) of the MNL are the ones governing the sDDM pairwise comparisons. Thus, table (15) allows us to combine Theorems 2, 3, and 5 in order to *identify* and *cross-validate* the unobservable parameters of internal and external processes.

Specifically, if the multialternative choice data available to the analyst do not reject the hypotheses of Theorem 2, she can use Theorem 3 and identify the psychometric parameters u and λ. At the same time, she can obtain sDDM data from binary comparisons under different levels of time pressure and estimate the neural parameters v and β.[30] Theorem 5 and, again, Theorem 2 say that, if the Metropolis-DDM algorithm converges, then the neural utility v must be a cardinal transformation of the psychometric utility u, and the threshold β must be an inverse transformation of the noise parameter λ. In sum, inside and outside analyses cross-validate each other (see 2.3.2 below for the details of the identification and cross-validation procedure).

2.3.1 Simulations

The crucial hypothesis behind this cross-validation procedure is the numerical convergence of the Metropolis-DDM algorithm to its stationary distribution in the finite number of iterations allowed by the time limit t. For a large t, thanks to inequality (13) numerical convergence is guaranteed whenever $\beta(t)$ is $o\left(\sqrt{t}\right)$.

[29] This follows from Doeblin Theorem. See, e.g., Stroock (2005, p. 28) for an explicit bound on the variation distance between $M_t^{\bar{n}_t}\mu$ and m_t.

[30] This has been done, e.g., by Milosavljevic et al. (2010).

But, often t is small. It is therefore important to understand whether numerical convergence takes place, with laboratory calibrated parameters v and $\beta(t)$, when t is in the order of a few seconds.

In the following simulations, we consider choice from two different menus of eight snacks $A = \{a_0, a_1, ..., a_7\}$ and $B = \{b_0, b_1, ..., b_7\}$, under two different deadlines of $t = 4$ and $t = 12$ seconds that induce high and low time pressure on binary choices. Our sDDM parameters v, $\beta(4)$, and $\beta(12)$ are derived from the estimates of Milosavljevic et al. (2010).

Menu A consists of alternatives that are equally spaced in utility, and no two alternatives are indifferent. Menu B instead consists of four pairs of indifferent alternatives that have the same utility.

For each menu, in each time pressure condition, we run the Metropolis-DDM algorithm ten thousand times. This procedure simulates ten thousand identical subjects who choose according to the Metropolis-DDM algorithm and delivers the empirical choice probabilities that an analyst would observe. These probabilities are plotted in orange below. Instead, in blue we plotted the stationary distribution of the algorithm. The indistinguishability of the orange and blue plots says that, even if the algorithm only performs a few iterations because of the imposed deadline t, numerical convergence to the stationary MNL distribution (12) is achieved. The average number of iterations – that is, of binary comparisons – performed in each simulation is also reported. Its small size makes the observed convergence results even more surprising.[31]

Simulation 1 (high time pressure, menu A, no indifferent alternatives) Choose in 4 seconds with $v(a_i) = i - 3.5$ (for $i = 0, ..., 7$) and $\beta(4) = 0.849$. In the picture below, the horizontal axis represents the alternatives. Specifically, the 8 ticks on the horizontal axis correspond to the alternatives, with the number below each tick representing the neural utility of the corresponding alternative. The choice probability of each alternative is reported on the vertical axis: again, the orange plot describes the empirical choice probabilities obtained by simulating the Metropolis-DDM algorithm, the blue plot describes the theoretical MNL choice probabilities. The average number of binary comparisons in this simulation is 18.

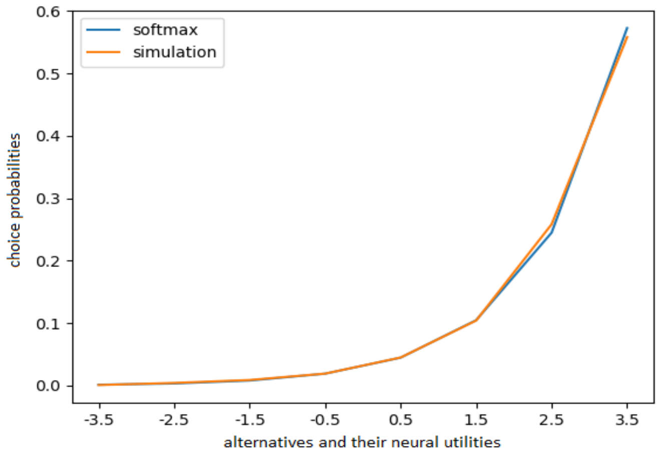


Figure 2 *Simulation 1.*

Simulation 2 (high time pressure, menu B, four pairs of indifferent alternatives) Choose in 4 seconds with $v(b_i) = |i - 3.5|$ (for $i = 0, ..., 7$) and $\beta(4) = 0.849$. The average number of binary comparisons in this simulation is 12.

[31] As anticipated, and in line with Krajbich and Rangel (2011) and Reutskaja et al. (2011), we use a uniform initial distribution μ in the simulations below; but the observed convergence, a number of alternative simulations, and the Doeblin Theorem (on the exponential convergence of Markov chains) tell us that any other μ would lead to similar results.

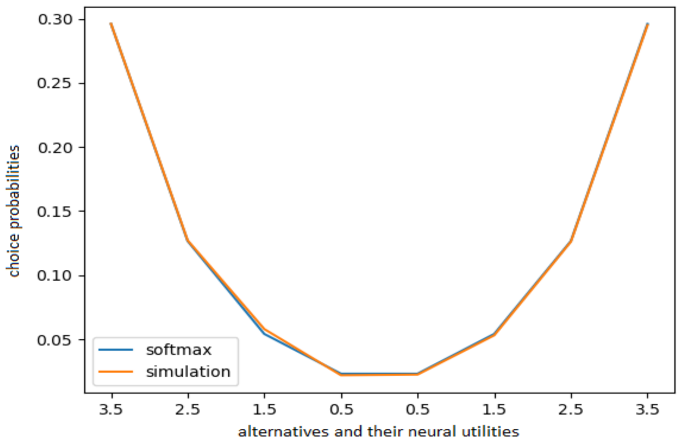


Figure 3 *Simulation 2.*

Simulation 3 (low time pressure, menu A, no indifferent alternatives) Choose in 12 seconds with $v(a_i) = i - 3.5$ (for $i = 0, ..., 7$) and $\beta(12) = 1.442$. The average number of binary comparisons in this simulation is 32.

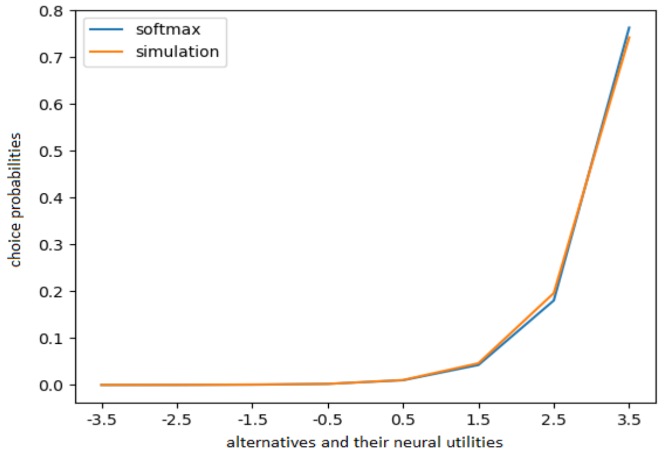


Figure 4 *Simulation 3.*

Simulation 4 (low time pressure, menu B, four pairs of indifferent alternatives) Choose in 12 seconds with $v(b_i) = |i - 3.5|$ (for $i = 0, ..., 7$) and $\beta(12) = 1.442$. The average number of binary comparisons in this simulation is 19.

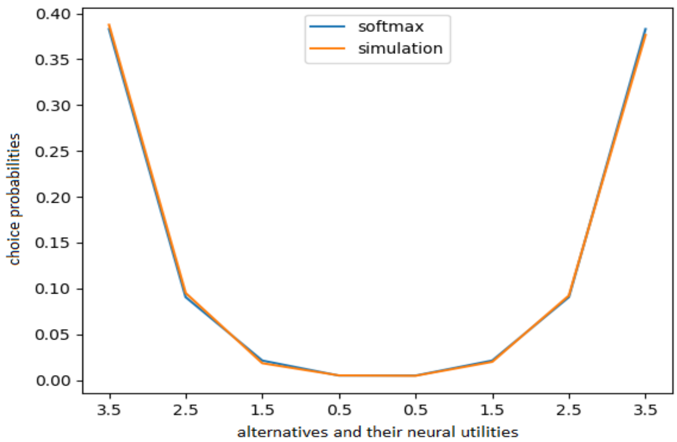


Figure 5 *Simulation 4.*

Summing up, the Metropolis-DDM algorithm seems to numerically converge when calibrated with physiological data.[32] This substantiates the identification and cross-validation techniques connecting the axiomatic and neuro-computational approaches that we described at the beginning of this section, and that we further discuss below.

2.3.2 Taking stock

Our axiomatic characterization of the MNL, Theorem 2, permits to test the Metropolis-DDM algorithm, which in turn, as shown by Theorem 5, approximates Matejka and McKay's optimal information acquisition strategy in a biologically feasible way. Our identification result, Theorem 3, allows the analyst to discover the "unknown state" that the decision maker is trying to determine, and reveals that this state simultaneously coincides with the psychometric utility (a behavioral measurement) and the neural utility (a physiological parameter). The same result also relates the unit cost of information $\lambda(t)$ of the MNL (a behavioral measurement) with the evidence threshold $\beta(t)$ of the DDM (a physiological parameter).

Our inside-outside analysis thus provides a procedure enabling the analyst to combine behavioral and physiological measurements in studying human choices. We can summarize it as follows.

Identification and cross-validation procedure

Neural MNL hypothesis For all A in $\mathcal{A}$ and all t in T, the Metropolis-DDM approximates its MNL distribution

$$m_t(a, A) = \frac{e^{\beta(t)v(a)}}{\sum_{b \in A} e^{\beta(t)v(b)}} \qquad \forall a \in A$$

with unknown neural components v and β.

Behavioral data The analyst observes a random choice process $\{p_t\}$, describing the frequencies of choice.

Behavioral test The analyst checks whether $\{p_t\}$ satisfies the axioms of Theorem 2. If this is the case, the neural MNL hypothesis is not rejected, and she posits that $m = p$, that is,

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

for all A in $\mathcal{A}$ and all t in T, with unknown behavioral components u and λ.

Identification If not uniform, the MNL choice process $\{p_t\}$ reveals, by Theorem 3, to the analyst the values $\hat{u}$ and $\hat{\lambda}$ of the behavioral components of $\{p_t\}$. By the uniqueness properties of MNL representations (Theorem 2), we have

$$v = u = \hat{u} \qquad \text{and} \qquad \beta = \frac{1}{\lambda} = \frac{1}{\hat{\lambda}}$$

up to cardinal transformations.[33] The neural and behavioral parameters are thus identified.

Cross-validation If available physiological data permit to identify the neural components v and β, the analyst can cross-validate the values previously obtained.

---

[32] See also Online Appendix D for some stress testing.

[33] That is, there exist $k > 0$ and $h \in \mathbb{R}$ such that $v = ku + h$ and $\beta = 1/k\lambda$, and the same holds for $\hat{u}$ and $\hat{\lambda}$ (with possibly different constants $\hat{k}$ and $\hat{h}$).

This procedure elucidates the interrelation between the inner and outer perspectives on stochastic choice studied in this paper. Far from being disconnected, these two perspectives complement each other conceptually – by providing external (behavioral) verification and internal (causal) explanation of MNL stochastic choice – as well as empirically – by permitting to identify and cross-validate the components of MNL specifications.

3 General analysis: softmax processes

We now present a general analysis of softmax processes able to account for the role of memory and initial biases.

3.1 Outside the black box: behavioral softmax

3.1.1 Preamble: psychometric utilities

The previous MNL analysis presents the advantages of going beyond the traditional ordinal framework – in which preferences only rank alternatives – by introducing a richer setting where preference intensities and their utility representations play a crucial role.

To extend the analysis, we consider three asymmetric and negatively transitive relations $\succ$, $\succ^{\natural}$, and $\succ^{*}$ on X.[34] The first, $\succ$, is a standard *preference order* that ranks alternatives, in the sense of Debreu (1954, 1964). The second, $\succ^{\natural}$, ranks pairs of alternatives in terms of *intensity of preference*, in the sense of Shapley (1975). The third, $\succ^{*}$, ranks binary choice problems in terms of *ease of comparison*, in the sense of Suppes and Winet (1955).

Formally, the relation $\succ$ is defined on the set of alternatives X and ranks them

$$a \succ b$$

In words, "a is preferred to b." The relation $\succ^{\natural}$ is defined on the set of pairs of distinct alternatives $X^2_{\neq} = \{(a,b) : a \neq b \text{ in } X\}$ and ranks them

$$(a,b) \succ^{\natural} (c,d)$$

In words, "the strength of preference for a over b is higher than that for c over d." Finally, the relation $\succ^{*}$ is defined on the set of binary choice sets $\mathcal{A}_2 = \{\{a,b\} : a \neq b \text{ in } X\}$ and ranks them

$$\{a,b\} \succ^{*} \{c,d\}$$

In words, "choosing between a and b is easier than choosing between c and d."

Next we introduce a joint numerical representation of these three binary relations that extends the traditional ordinal representation.

Definition 2 *A function $u : X \to \mathbb{R}$ is a* psychometric utility *for the triplet $(\succ, \succ^{\natural}, \succ^{*})$ if, for each pair of alternatives $a, b \in X$,*

$$a \succ b \iff u(a) > u(b) \tag{16}$$

and if, for each quadruple of alternatives $a \neq b$ and $c \neq d$ in X,

$$(a,b) \succ^{\natural} (c,d) \iff u(a) - u(b) > u(c) - u(d) \tag{17}$$

as well as

$$\{a,b\} \succ^{*} \{c,d\} \iff |u(a) - u(b)| > |u(c) - u(d)| \tag{18}$$

[34] Asymmetry and negative transitivity are the standard assumptions for a strict preference (see Definition 2.2 of Kreps, 1988).

A psychometric utility not only represents the preference order $\succ$ in the standard fashion, but also accounts for the intensity of preferences, quantified via utility differences, as well as for the ease of comparison, quantified via absolute values of utility differences.

Psychometric utilities are cardinal, as the next routine lemma shows.

Lemma 6 *Continuous psychometric utilities are cardinally unique.*

The function v of Theorem 1 is, aptly, a psychometric utility for the triplet $(\succ, \succ^\natural, \succ^*)$ defined by

$$\begin{aligned} a \succ b &\iff p(a,b) > p(b,a) \\ (a,b) \succ^\natural (c,d) &\iff p(a,b) > p(c,d) \\ \{a,b\} \succ^* \{c,d\} &\iff \mathrm{ER}\,(a,b) < \mathrm{ER}\,(c,d) \end{aligned} \tag{19}$$

for all quadruples of alternatives $a \neq b$ and $c \neq d$ in X. The error rate

$$\mathrm{ER}\,(a,b) = \min\{p(a,b), p(b,a)\} = \begin{cases} p(a,b) & \text{if } b \succ a \\ p(b,a) & \text{if } a \succ b \\ 1/2 & \text{otherwise} \end{cases}$$

is the probability of choosing an alternative from $\{a,b\}$ which is inferior under $\succ$. To check that v is a psychometric utility for this triplet $(\succ, \succ^\natural, \succ^*)$, first observe that

$$(a,b) \succ^\natural (c,d) \iff (a,b) \rhd (c,d)$$

i.e., $\succ^\natural$ reveals intensity of preference (and of evidence) as discussed in Section 2. Second, $\succ$ reveals ordinal preference according to the standard (unbiased) stochastic notion

$$a \succ b \iff a \neq b \text{ and } p\,(a,b) > 1/2$$

which informed economics and psychology since the 1950s.[35] Finally, the idea that error rates measure difficulty of comparison, which dates back to Fechnerian psychophysics,[36] is often spelled out via the following classic psychophysical principle.

Psychometric Principle *Easier choice problems are more likely to elicit correct responses than harder ones.*[37]

We conclude that v is, indeed, a psychometric utility for the triplet $(\succ, \succ^\natural, \succ^*)$ defined in (19).

3.1.2 Measurement and revelations

The MNL analysis of Section 2 showed that – in the absence of initial bias – stochastic choice behavior under time pressure can be faithfully described by means of preference intensities and psychometric utilities. Our general softmax representation theorem, Theorem 7 below, will show that this continues to be the case also when initial bias is present. To take advantage of psychometric techniques, however, we will have to extract a triplet $(\succ_t, \succ_t^\natural, \succ_t^*)$ from each component p_t of a general random choice process. But, how can

[35] See, e.g., Georgescu-Roegen (1936, 1958), Mosteller and Nogee (1951), Papandreou (1953) and Papandreou et al. (1957), Quandt (1956), Debreu (1958), and Davidson and Marschack (1959).

[36] See, e.g., Falmagne (1985).

[37] This principle is often discussed under the name "Psychometric Function." See, e.g., Alos-Ferrer, Fehr, and Netzer (2021).

an analyst detect and measure the preference order and its intensity, or ease of comparison, if the decision maker's behavior is initially biased?

To address this question, we first augment the set T of strictly positive deliberation times to $T_0 = T \cup \{0\}$ and, given any two distinct alternatives a and b, we interpret the *initial probability*

$$p_0(a,b)$$

as the frequency with which a is chosen over b, when no evidence-based deliberation is possible. With this, we can introduce a key notion.

Definition 3 *Alternatives a and b are* a priori homogeneous *if $p_0(a,b) = 1/2$*

In words, alternatives are a priori homogeneous when there is no initial bias for one over the other. The interpretation of

$$p_t(a,b)$$

for $t > 0$ is still that discussed in Section 2 as the frequency with which a is chosen over b after the decision maker had the possibility to deliberate for t seconds, during which information about the two alternatives is gathered and processed. Depending on the obtained evidence, be it from environment or memory (or both),[38] the choice probability $p_t(a,b)$ at deliberation time t may well be different from the initial $p_0(a,b)$.[39] We interpret this change in light of the following basic principle.

Measurement Principle *Prior behavior gets transformed into posterior behavior through consideration of evidence, and the transformation itself represents the amount of evidence processed during deliberation.*

This principle is best formalized through a change in odds as:

$$\underbrace{r_t(a,b)}_{\text{posterior odds}} = \underbrace{f}_{\text{strength of evidence}} \times \underbrace{r_0(a,b)}_{\text{prior odds}} \tag{20}$$

The ratio

$$f = f_t(a,b) = \frac{r_t(a,b)}{r_0(a,b)}$$

represents the *strength of evidence*, processed in t seconds, in favor of the hypothesis "a is preferable to b."

That said, in both statistics and neuroscience, additive measurements are preferred.[40] This is routinely achieved by taking logarithms on both sides of (20):

$$\underbrace{\ell_t(a,b)}_{\text{posterior log-odds}} = \underbrace{\ln f_t(a,b)}_{\text{weight of evidence}} + \underbrace{\ell_0(a,b)}_{\text{prior log-odds}}$$

The difference

$$w_t(a,b) = \ln f_t(a,b) = \ell_t(a,b) - \ell_0(a,b)$$

is the additive version of $f_t(a,b)$, called *weight of evidence*, a convenient logarithmic rescaling of strength of evidence.

Summing up, the strength of evidence is the change in odds for a against b induced by evidence accumulation before the deadline. Next we show how this important notion permits to reveal preference

[38] See, e.g., Bogacz et al. (2006), Gold and Shadlen (2007), Shadlen and Shohamy (2016), and Bordalo, Gennaioli, and Shleifer (2020).

[39] See, again, Silayoi and Speece (2004) and Huseynov and Palma (2021).

[40] See, again, Gold and Shadlen (2007).

order, intensity of preference, and ease of comparison – in symbols, to reveal the triple $(\succ_t, \succ_t^\natural, \succ_t^*)$ for all t in T.

We begin with the traditional ordinal notion. In the following, as usual, the term "revealed" is short for "revealed to an analyst."

Definition 4 *After a deliberation time t, an alternative a is* revealed preferred *to b, written $a \succ_t b$, if $p_t(a,b) > p_0(a,b)$.*

In words, a is revealed preferred to b if deliberation favors a over b. In particular, when alternatives are a priori homogeneous, i.e., $p_0(a,b) = 1/2$, this definition coincides with the standard unbiased notion of stochastically revealed preference

$$a \succ_t b \iff p_t(a,b) > p_t(b,a)$$

introduced in (19). In general, when alternatives are not necessarily a priori homogeneous, preference for a over b is equivalently revealed by an increase in the odds for a against b after deliberation. Indeed,

$$a \succ_t b \iff w_t(a,b) > 0 \iff f_t(a,b) > 1$$

Starting from this observation, Luce (1957, pp. 17-19) observes that, while the preference order is determined by the sign of $w_t(a,b)$, the intensity of preference is determined by its value. This motivates the next definition.

Definition 5 *After a deliberation time t, the preference for a over b is* revealed to be stronger than *that for c over d, written $(a,b) \succ_t^\natural (c,d)$, if $w_t(a,b) > w_t(c,d)$.*

In words, the preference for a over b is stronger than that for c over d if deliberation provides stronger evidence in favor of a against b than in favor of c against d. Formally,

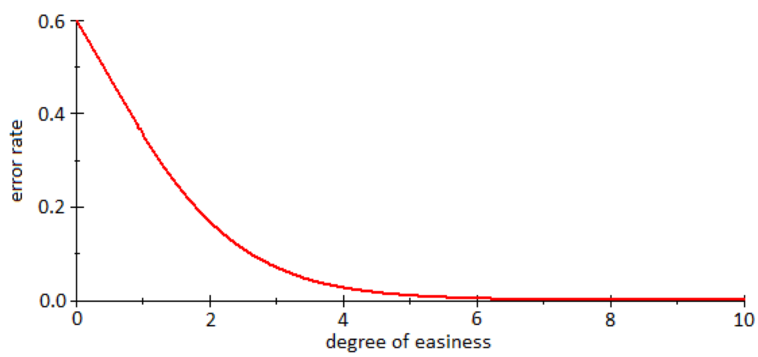
$$(a,b) \succ_t^\natural (c,d) \iff w_t(a,b) > w_t(c,d) \iff f_t(a,b) > f_t(c,d)$$

This definition thus extends the identification between intensity of preference and intensity of evidence that we discussed in Section 2. Clearly, in the a unbiased case, the relation $\succ_t^\natural$ reduces to the (unbiased) intensity of preference relation $\rhd_t$ adopted there.

In a series of papers,[41] Georg Rasch calls *degree of easiness* of a decision problem $\{a,b\}$ the quantity

$$|w_t(a,b)|$$

The reason why this quantity captures the Psychometric Principle is immediately seen by drawing the error rate – the probability of choosing the inferior alternative – in the decision problem $\{a,b\}$ as a function of the degree of easiness.



[41] See Rasch (1960, 1961, 1980).

Figure 6 *Error rate as a function of the degree of easiness, with an initial bias of* 10% *in favor of the inferior alternative.*[42]

When the degree of easiness is zero, the error rate is maximal and coincides with the initial probability of choosing the inferior alternative. It then decreases exponentially as the degree of easiness increases, and eventually vanishes. Analytically, this is seen by observing that, if $w_t(a,b) > 0$, that is, $a \succ_t b$, then the error rate is the probability

$$\underbrace{\mathrm{ER}_t(a,b)}_{\text{error rate}} = p_t(b,a) = \frac{1}{1 + \exp\left(\ell_0(a,b) + \underbrace{|w_t(a,b)|}_{\text{degree of easiness}}\right)}$$

of choosing b from $\{a,b\}$, and similar considerations apply if $w_t(a,b) < 0$.

Since $w_t(a,b) = -w_t(b,a)$, the evidence in favor of a coincides with that against b. The degree of easiness $|w_t(a,b)|$ thus represents the *total amount of evidence* that can be obtained by comparing a and b for t seconds. A decision problem is difficult when this quantity is small, say because sensory evidence or memory do not provide information to the decision maker about the alternatives. All this leads to the following definition.

Definition 6 *After a deliberation time* t*, a decision problem* $\{a,b\}$ *is* revealed to be easier than *a decision problem* $\{c,d\}$*, written* $\{a,b\} \succ_t^* \{c,d\}$*, if* $|w_t(a,b)| > |w_t(c,d)|$*.*

At this point, it should not surprise the reader that also this definition reduces to (19) when alternatives are a priori homogeneous.

Summing up, weight of evidence – or, equivalently, strength of evidence – can be elicited from choice data by looking at the variation of choice probabilities *before* and *after* deliberation. It reveals three relations: preference order, preference intensity, and ease of comparison – in symbols $(\succ_t, \succ_t^\natural, \succ_t^*)$.

3.1.3 Axiomatic softmax

The natural extension of Consistency to the general, possibly biased case, is:

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

$$(a,b) \succ_t^\natural (c,d) \iff (a,b) \succ_s^\natural (c,d)$$

for all $a \neq b$ *and* $c \neq d$ *in* X*.*

This axiom says that if the weight of evidence in favor of the hypothesis "a is preferable to b" is, after a given deliberation time t, greater than that in favor of the hypothesis "c is preferable to d", the same happens after a longer deliberation time s. Formally,

$$w_t(a,b) > w_t(c,d) \iff w_s(a,b) > w_s(c,d)$$

for all $s > t$. In the unbiased case, $\succ_t^\natural$ reduces to $\rhd_t$ and so Intensity Consistency reduces to Consistency.[43]

The next representation theorem will show that – together with the Psychometric Luce Axioms – Intensity Consistency characterizes general softmax processes. Yet, an alternative characterization is obtained by using non-reversal conditions for the preference order $\succ_t$ and the ease of comparison relation

[42] For instance, when $a \succ_t b$ the inferior alternative is b.

[43] Like Consistency, Intensity Consistency admits a formulation which is viable for general index sets T, see Footnote 15.

$\succ_t^*$. Interestingly, these conditions have a one-way form, weaker than the two-way form of Intensity Consistency.

Preference (Order) Consistency *Given any $s > t$ in T,*

$$a \succ_t b \implies a \succ_s b$$

for all $a, b \in X$.

Ease (of Comparison) Consistency *Given any $s > t$ in T,*

$$\{a, b\} \succ_s^* \{c, d\} \implies \{a, b\} \succ_t^* \{c, d\}$$

for all $a \neq b$ and $c \neq d$ in X.

In terms of primitives, Preference Consistency is equivalent to

$$p_t(a, b) > p_0(a, b) \implies p_s(a, b) > p_0(a, b)$$

for all $s > t$. It says that preferences are stable: as time passes, they are not reverted. This is in accord with the idea that, during deliberation, statistically correct (yet noisy) evidence is gathered and analyzed by the decision maker to inform his choice between the two alternatives.

Ease Consistency, instead, says that the difficulty of decision problem $\{a, b\}$ relative to decision problem $\{c, d\}$ is inherent to the alternatives involved and independent of deliberation times. If the comparison between a and b is not easier than that between c and d, given deliberation time t, then the passage of time does not make a and b easier to compare than c and d. In terms of degree of easiness, Ease Consistency is equivalent to

$$|w_t(a, b)| \leq |w_t(c, d)| \implies |w_s(a, b)| \leq |w_s(c, d)|$$

for all $s > t$.

We can now state the general softmax representation theorem. We adopt the convention $\lambda(0) = \infty$.

Theorem 7 *For a random choice process $\{p_t\}_{t \in T_0}$, the following conditions are equivalent:*

1. *$\{p_t\}$ satisfies the Psychometric Luce Axioms and Intensity Consistency;*
2. *$\{p_t\}$ satisfies the Psychometric Luce Axioms, Preference Consistency, and Ease Consistency;*
3. *there exist two continuous functions $u : X \to \mathbb{R}$ and $\alpha : X \to \mathbb{R}$, and a $\lambda : T \to (0, \infty)$ such that*

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

for all $A \in \mathcal{A}$, all $a \in A$, and all $t \in T_0$.

In this case, u is a psychometric utility for $(\succ_t, \succ_t^\natural, \succ_t^*)$ for all t in T, α is unique up to location, and λ is unique given u, unless $\{p_t\}$ is constant.

Moreover, the process $\{p_t\}_{t \in T}$ is MNL if and only if all distinct alternatives are a priori homogeneous.

An analyst, who observes that the decision maker's behavior does not contradict the axioms of the theorem, can thus interpret this behavior in terms of preference discovery, that is, *as if* the decision maker were trying to learn the value that alternatives have for him and choose the best one.

Like in the unbiased case, our analyst can identify from the probabilistic choices of the decision maker the softmax parameters u, α, and λ.[44] Specifically, u is a psychometric utility for each triplet $(\succ_t, \succ_t^\natural, \succ_t^*)$, and as such it is cardinally unique (Lemma 6). Then, if the process is constant, u must be constant, $\alpha(a) - \alpha(b) = \ell_0(a, b)$ for all $a, b \in X$, and λ is undefined; else there exist at least a pair of alternatives $\hat{a}$ and $\hat{b}$ and a deliberation time $\hat{t}$ such that the preference $\hat{a} \succ_{\hat{t}} \hat{b}$ is revealed, and the next result provides the explicit expression of the parameters.

Theorem 8 *A softmax process $\{p_t\}$ is not constant if and only if there exist $\hat{a}, \hat{b} \in X$ and $\hat{t} \in T$ such that $p_{\hat{t}}(\hat{a}, \hat{b}) > p_0(\hat{a}, \hat{b})$. In this case, the functions $\hat{u}, \hat{\alpha} : X \to \mathbb{R}$ and $\hat{\lambda} : T \to (0, \infty)$ given by*

$$\hat{u}(x) = \frac{w_{\hat{t}}(x, \hat{b})}{w_{\hat{t}}(\hat{a}, \hat{b})} \quad ; \quad \hat{\alpha}(x) = \ell_0(x, \hat{b}) \quad ; \quad \hat{\lambda}(t) = \frac{1}{w_t(\hat{a}, \hat{b})} \tag{21}$$

are well defined, with $\hat{u}$ and $\hat{\alpha}$ continuous, and

$$p_t(a, A) = \frac{e^{\frac{\hat{u}(a)}{\hat{\lambda}(t)} + \hat{\alpha}(a)}}{\sum_{b \in A} e^{\frac{\hat{u}(b)}{\hat{\lambda}(t)} + \hat{\alpha}(b)}} \tag{22}$$

for all $A \in \mathcal{A}$, all $a \in A$, and all $t \in T_0$.

The last two results extend the "unbiased" Theorems 2 and 3 to the general case. So, they enable the analyst to interpret the stochastic choice behavior of the decision maker in terms of softmax preference discovery and to empirically identify the softmax parameters.

3.1.4 Ordinality and learning

We conclude this section by characterizing the case in which, as deliberation time increases, the stochastic choice behavior of a decision maker improves and becomes less prone to errors.[45]

Decreasing Error Rates *Given any $s > t$ in T,*

$$p_t(a, b) > p_0(a, b) \implies p_s(a, b) \geq p_t(a, b)$$

for all $a, b \in X$.

Under Preference Consistency, this axiom requires the frequency of mistakes to decrease over deliberation time. Indeed, $p_t(a, b) > p_0(a, b)$ is equivalent to the fact that $a \succ_t b$, Preference Consistency implies $a \succ_s b$, and $p_s(a, b) \geq p_t(a, b)$ is equivalent to a positive reduction of the error rate. In words, longer deliberation times decrease the chance of selecting an inferior alternative. To appreciate the consequences of this axiom, we need an additional one.

Stochastic Dominance Improvements *Given any $s > t$ in T,*

$$p_s(\{a \in A : a \succ_s b\}, A) \geq p_t(\{a \in A : a \succ_t b\}, A) \tag{23}$$

[44] Their estimation 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.

[45] For instance, in medical decision making under severe time pressure, the longer the time a doctor has to process information, the lower the chance of selecting a suboptimal treatment seems to be (see, e.g., ALQathani et al., 2016). Yet, other evidence from psychology suggests that overly slack deadlines leave room to procrastination, distractions, and fatigue that may deteriorate choice performance (see, e.g., Ariely and Wertenbroch, 2002). Clearly, the study of these time-increasing error rate situations mirrors the one we consider here.

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

Stochastic Dominance Improvements requires that, for any given benchmark alternative b, the probability of choosing a superior alternative a is higher after deliberating for a longer amount of time. This notion thus records, in preferential terms, a probabilistic improvement of the decision maker stochastic choice behavior as deliberation times increase.

The next proposition shows that Decreasing Error Rates and Stochastic Dominance Improvements are equivalent axioms for softmax processes. Moreover, they characterize the choice behavior of a decision maker who takes better and better decisions, according to stochastic dominance in payoffs, as deliberation time increases. In terms of the softmax specification, each of these axioms corresponds to a decreasing noise function λ. In terms of rational inattention, to a time-decreasing unit cost of information processing (e.g., the attention cost of reading and understanding a given paragraph decreases with the time available to do so).

Proposition 9 *Let $\{p_t\}$ be a nonconstant softmax process with utility u, bias α, and noise λ. The following conditions are equivalent:*

1. *$\{p_t\}$ satisfies Decreasing Error Rates;*
2. *$\{p_t\}$ satisfies Stochastic Dominance Improvements;*
3. *given any $s > t$ in T,*

$$p_s\left(\{a \in A : u(a) > \bar{u}\}, A\right) \geq p_t\left(\{a \in A : u(a) > \bar{u}\}, A\right)$$

   for all $\bar{u} \in \mathbb{R}$ and all $A \in \mathcal{A}$;

4. *λ is decreasing.*

In view of this result, it is natural to wonder whether, for longer and longer deliberation times, the decision maker eventually learns his ranking over alternatives, that is, his preference order over them. In other words, is the preference discovery interpretation of softmax processes true to its name?

To address this question, assume for simplicity that $T = (0, \infty)$.[46] By the last result, under Decreasing Error Rates, the noise function λ is decreasing on $(0, \infty)$. This permits to define a limit random choice rule $p_\infty : \mathcal{A} \to \Delta(X)$ by

$$p_\infty(a, A) = \lim_{t \to \infty} p_t(a, A)$$

for all A in $\mathcal{A}$ and all a in A. This random choice rule describes behavior in the absence of time pressure.

Asymptotic Tie-breaking *Given any $a, b \in X$,*

$$p_\infty(a, b) \neq 0, 1 \implies p_\infty(a, b) = p_0(a, b)$$

In conjunction with the previous ones, this axiom says that the decision maker is unable to make up his mind between alternatives a and b, irrespective of deliberation time, only if it is impossible for him to retrieve new information about them.

Proposition 10 *Let $\{p_t\}$ be a nonconstant softmax process with utility u, bias α, and noise λ. If $\{p_t\}$ satisfies Decreasing Error Rates and Asymptotic Tie-breaking, then*

$$p_\infty(a, A) = \delta_a(\arg\max_A u) \frac{e^{\alpha(a)}}{\sum_{b \in \arg\max_A u} e^{\alpha(b)}}$$

[46] Otherwise, the role of ∞ is played by the supremum of T.

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

$$u(a) > u(b) \Longleftrightarrow p_{\infty}(a, b) = 1$$

for all $a \neq b$ in X.

According to this proposition, the choice rule p_{∞} reveals a preference $\succ$ on X defined by

$$a \succ b \Longleftrightarrow p_{\infty}(a, b) = 1$$

and represented by u. This preference permits to interpret the non-stochastic limit choice behavior in a traditional ordinal way, as if carried out by a decision maker who learned his preference – so, his psychometric utility u up to an ordinal transformation – and accordingly selects the best alternatives.[47]

Standard ordinal analysis thus emerges as the limit version, as deliberation time becomes arbitrarily large, of our cardinal analysis. Alternatively, one can regard standard theory as assuming deliberation time to be virtual; in real time, decision makers act as if they know their preferences.

3.2 Inside the black box: neuro-computational softmax

Symmetry of the DDM presented in Section 2.2.1 permits the description of the net evidence obtained by the decision maker in differential form as

$$\begin{aligned} \mathrm{d}Z_{a,b}(\tau) &= [v(a) - v(b)]\,\mathrm{d}\tau + \sqrt{2}\mathrm{d}W(\tau) \\ Z_{a,b}(0) &= 0 \end{aligned} \tag{24}$$

where the initial condition $Z_{a,b}(0) = 0$ captures the lack of prior information/initial bias in favor or against either a or b. As extensively discussed by Bogacz et al. (2006) and Ratcliff et al. (2016), prior information/initial bias are captured in the DDM by replacing the null initial condition (24) with

$$Z_{a,b}(0) = \zeta_t(a, b)$$

where $\zeta_t(a, b) > 0$ means that a is initially favored, and $\zeta_t(a, b) < 0$ that b is. In fact, with starting point position $\zeta_t(a, b) \neq 0$, the evidence needed to select a and b becomes $\beta(t) - \zeta_t(a, b)$ and $\beta(t) + \zeta_t(a, b)$, respectively.

In this section, we study what happens when a general DDM replaces the symmetric DDM in the Metropolis-DDM algorithm. The additional ingredient, relative to the neural utility $v : X \to \mathbb{R}$ and threshold function $\beta : T \to (0, \infty)$ of Section 2.2.1, is the *starting point function*

$$\begin{array}{rccl} \zeta : & X^2_{\neq} \times T & \to & \mathbb{R} \\ & (a, b, t) & \mapsto & \zeta_t(a, b) \end{array}$$

such that

$$-\beta(t) < \zeta_t(a, b) = -\zeta_t(b, a) < \beta(t)$$

for all (a, b, t). The unbiased case of Section 2.2.1 corresponds to $\zeta = 0$. In what follows, we denote by

$$\mathrm{DDM}(v, \beta, \zeta)$$

a DDM with parameters v, β, and ζ.

[47] In contrast, an analyst learns u up to a cardinal transformation by observing the decision maker softmax stochastic behavior (Theorem 8).

3.2.1 Gibbs transitions

In this subsection we relate starting point positions and initial probabilities as equivalent ways to express initial biases. To this end, consider a choice between two alternatives a and b to be carried out through a DDM featuring neural utility v, threshold β, and starting point ζ. As we just observed, the values

$$\zeta_t(a,b) = -\zeta_t(b,a)$$

at (a,b) and (b,a), determine whether the DDM comparison favors either a or b. However, it should be possible to express the same bias with the values

$$\pi_t(a,b) = 1 - \pi_t(b,a)$$

of an *ex ante (binary preference) probability* $\pi_t : X^2_{\neq} \to \mathbb{R}$ at (a,b) that describes the chances of choosing either a or b before the DDM comparison takes place, that is, before new evidence is gathered and processed.[48] Both ζ_t and π_t incorporate the decision maker past information, in particular his past memories. Both lack – differently from $\mathrm{d}Z$ and β – an obvious physiological counterpart and their values are posited by the analyst to better interpret the model and fit the data.[49]

How do they relate? Which ex ante probability π_t corresponds to a starting point function ζ_t? To address these natural questions, observe that ζ induces an *ex post (binary preference) probability*

$$\mathbb{P}_t^\zeta(a,b) = \mathbb{P}\left[\mathrm{CO}_t^\zeta(a,b) = a\right]$$

of choosing a over b, after the DDM that starts at position $\zeta_t(a,b)$ has accumulated neural evidence $\beta(t)$.

Definition 7 *Given a DDM*(v,β,ζ)*, the* Gibbs transition *of* ζ *is*

$$\pi_t^\zeta(a,b) = \frac{e^{-\beta(t)v(a)}\mathbb{P}_t^\zeta(a,b)}{e^{-\beta(t)v(a)}\mathbb{P}_t^\zeta(a,b) + e^{-\beta(t)v(b)}\mathbb{P}_t^\zeta(b,a)} \tag{25}$$

for all $a \neq b$ *in* X *and all* t *in* T.

Formula (25) associates an ex ante probability π_t to each section ζ_t at t of the starting point function ζ.[50] The Gibbs transition of ζ can be equivalently expressed in terms of odds as follows:

$$\underbrace{\frac{\mathbb{P}_t^\zeta(a,b)}{\mathbb{P}_t^\zeta(b,a)}}_{\text{odds post DDM}} = \underbrace{e^{\beta(t)[v(a)-v(b)]}}_{\text{strength of evidence}} \times \underbrace{\frac{\pi_t^\zeta(a,b)}{\pi_t^\zeta(b,a)}}_{\text{odds ante DDM}} \tag{26}$$

Therefore, this formula can be interpreted according to the Measurement Principle of Section 3.1.2. With one caveat: in equation (20) of that section, the analyst observes the ex ante and the ex post odds, and aims to measure the unknown strength of evidence. Here, in contrast, the analyst observes the ex post odds and the evidence threshold, and aims to measure the unknown ex ante odds that are implied by a posited bias ζ. Yet, the underlying measurement principle is the same: the change in odds for a against b resulting from the DDM is proportional, via an exponential factor, to the accumulated neural evidence $\beta(t)$ weighted by the neural utility difference $v(a) - v(b)$. The quantity $\beta(t)[v(a) - v(b)]$ is thus the weight of evidence for a against b that makes the neural system move from the ex ante to the ex post probability of choosing a over b.

Observe that, for $\zeta_t(a,b) = 0$, formula (26) is easily seen to imply $\pi_t^\zeta(a,b) = 1/2$, as the intuition for the unbiased sDDM suggests. In words, a null starting point position corresponds to the a priori homogeneity of alternatives.

The definition of Gibbs transition can also be justified through large deviations arguments.[51] But

[48] The "binary preference probability" terminology is due to Luce and Suppes (1965). Like ζ_t, also π_t may in principle depend on t.

[49] See Bogacz et al. (2006).

[50] It is, *mutatis mutandis*, the analog of the Gibbs prior of Zhang (2006a, 2006b).

[51] See, e.g., Dupuis and Ellis (1997, p. 27), and the working paper version of this paper, Cerreia-Vioglio et al. (2021a).

perhaps more importantly, the relation between starting point positions and ex ante probabilities, defined via (25), features some remarkable properties.

Proposition 11 *Given a DDM(v, β, ζ), for any $a \neq b$ in X we have*

$$\zeta_t(a, b) \geq 0 \Longleftrightarrow \pi_t^{\zeta}(a, b) \geq \frac{1}{2} \tag{27}$$

and

$$\left|\mathbb{P}_t^{\zeta}(a, b) - \pi_t^{\zeta}(a, b)\right| \leq \frac{\beta(t)}{4}|v(a) - v(b)| \tag{28}$$

for all t in T.

The monotonicity formula (27) ensures that a positive bias $\zeta_t(a, b)$ in favor of a against b corresponds to a higher ex ante probability $\pi_t^{\zeta}(a, b)$ of selecting a over b. It implies, *inter alia*, that a null ζ corresponds to a uniform π, as we previously checked in a direct way.

The monotonicity formula thus substantiates the claim that π_t^{ζ} is the ex ante probability naturally associated to the initial condition ζ_t. Inequality (28) further corroborates this role of π_t^{ζ} by showing that it actually governs the DDM's probabilistic choices when the accumulated evidence β is small or alternatives have similar neural utilities, thus initial bias prevails.

This leads to a crucial observation: since it is not affected by evidence accumulation, the ex ante choice probability π_t should not depend on the level of time pressure.

Gibbs Transition Consistency *Given any $s > t$ in T,*

$$\pi_t^{\zeta}(a, b) = \pi_s^{\zeta}(a, b)$$

for all $a \neq b$ in X.

If this consistency condition is violated, there is a pair (a, b) of alternatives whose ex ante choice probability differ, $\pi_t^{\zeta}(a, b) \neq \pi_s^{\zeta}(a, b)$, depending on the level of time pressure, t or s. But, this contradicts the intuition that ζ and π^{ζ} describe ex ante information only. In other words, if a starting point function does not satisfy Gibbs Transition Consistency, then it captures something more than (or something different from) memory anchoring.

3.2.2 Transitive DDMs and stationarity

When considering a Metropolis-DDM algorithm with general DDM(v, β, ζ) binary comparisons, the first observation is that, even if DDM(v, β, ζ) satisfies Gibbs Transition Consistency, the transition matrix $M_t(A)$ of the algorithm may fail to be reversible for some A in $\mathcal{A}$. But reversibility is the Markovian condition which convergence of MCMC algorithms to a distribution that satisfy Luce's Choice Axiom rests upon (see Baldassi et al., 2020b and Valkanova, 2021).

The next proposition will show that reversibility has the following meaningful behavioral counterpart.

(Stochastic) Transitivity *Given any $t \in T$,*

$$\mathbb{P}_t^{\zeta}(b, a)\,\mathbb{P}_t^{\zeta}(c, b)\,\mathbb{P}_t^{\zeta}(a, c) = \mathbb{P}_t^{\zeta}(c, a)\,\mathbb{P}_t^{\zeta}(b, c)\,\mathbb{P}_t^{\zeta}(a, b) \tag{29}$$

for all distinct alternatives $a, b, c \in A$.

In words, DDM(v, β, ζ) is transitive when the violations of transitivity that it determines are due only to the presence of noise. Indeed, condition (29) amounts to require that the intransitive cycles

$$a \to b \to c \to a \quad \text{and} \quad a \to c \to b \to a$$

be equally likely. The symmetric (unbiased, value-based) DDMs, considered in Section 2.2, are an important example of transitive DDM's that satisfy Gibbs Transition Consistency. Biased value-based DDMs, instead, might well not be transitive, so may result in choices between alternatives that feature systematic intransitivities, violating a basic rationality tenet. Transitivity ensures that this is not the case.

The next result, which builds upon Kolmogorov (1936) and Luce and Suppes (1965),[52] shows the importance of transitive DDMs in our setting, and a key feature of the corresponding ex ante binary probabilities.

Proposition 12 *The following conditions are equivalent for a DDM(v, β, ζ) which satisfies Gibbs Transition Consistency:*

1. *DDM(v, β, ζ) is transitive;*

2. *the incumbents' transition matrix $M_t(A)$ is reversible for all $A \in \mathcal{A}$ and all $t \in T$;*

3. *there exists $\varsigma : X \to \mathbb{R}$ such that,*
$$\pi_t^{\zeta}(a, b) = \frac{e^{\varsigma(a)}}{e^{\varsigma(a)} + e^{\varsigma(b)}}$$
for all $(a, b, t) \in X^2_{\neq} \times T$.

In this case, ζ is uniquely determined by ς, and vice versa (up to location).

Remarkably, this proposition connects properties of altogether different nature:

1. transitivity of the DDM, a behavioral property which ensures that violations of transitivity in the probabilistic choices that it determines are due only to the presence of noise;

2. reversibility of M, an algorithmic property which is an important sufficient condition for the existence of stationary distributions of Markov chains,[53] which is necessary to guarantee that the generated distributions satisfy Renyi conditioning as the set of Markov states changes;[54]

3. existence of a universal *Gibbs density* $\varsigma : X \to \mathbb{R}$ for ζ that, via conditioning, determines all Gibbs transitions.

Because of the one-to-one correspondence between ζ and ς (up to location) established by Proposition 12, transitive DDMs that satisfy Gibbs Transition Consistency can be simply denoted by DDM(v, β, ς).[55] With this, we can establish the general version of Theorem 5.

Theorem 13 *Let DDM(v, β, ς) be a transitive DDM which satisfies Gibbs Transition Consistency. Given any $A \in \mathcal{A}$ and any $t \in T$, the exploration matrix $Q_t = Q_t(A)$ is irreducible and symmetric and the incumbents' transition matrix $M_t = M_t(A)$ is aperiodic, irreducible, and reversible, with stationary distribution*
$$m_t(a, A) = \frac{e^{\beta(t)v(a)+\varsigma(a)}}{\sum_{b \in A} e^{\beta(t)v(b)+\varsigma(b)}} \qquad \forall a \in A$$
In particular, $\lim_{n \to \infty} M_t^n(A)\mu = m_t(\cdot, A)$ *for all $A \in \mathcal{A}$ and all $\mu \in \Delta(A)$.*

[52] On reversibility, see, e.g., Kelly (2011).

[53] As Geyer (2011, p. 6) writes "All known methods for constructing transition probability mechanisms that preserve a specified equilibrium distribution in non-toy problems are ... reversible."

[54] See, again, Baldassi et al. (2020b) and Valkanova (2021).

[55] In practice, once v, β, and ζ are estimated in the lab, then checking whether DDM(v, β, ζ) satisfies Gibbs Transition Consistency and transitivity is a matter of hypothesis testing. For example, Milosavljevic et al. (2010) cannot reject the nullity of ζ, and *a fortiori* these weaker assumptions.

This result concludes our analysis by extending the identification and cross-validation procedure that we described in Section 2.3 to the general, possibly biased, case. Table (15) becomes

Inner		**Outer**	
Neural utility	v	Psychometric utility	u
Threshold	β	Inverse noise	$1/\lambda$
Starting point bias	ς	Behavioral bias	α

where the Gibbs density ς corresponding to the starting point function ζ is shown to be the counterpart of initial bias α. As mentioned, the neuroscience literature and the rational inattention literature interpreted these parameters in the same way, but the formal connection between them has been so far elusive.

Theorem 13 reveals the connection and shows how the Metropolis-DDM algorithm (approximately) generates softmax probabilities, thus implementing the optimal information acquisition strategy of Matejka and McKay (2015). In turn, the axiomatic representation Theorem 7 characterizing the latter, provides testable implications for the former.[56] For brevity, we omit the, by now, routine details of identification and cross-validation that parallel those spelled out in Section 2.3.

4 Concluding remarks

In a few concluding remarks, we explore limitations and possible future extensions of our analysis.

4.1 Latency

Our main interpretation of t is as the amount of time available to *gather and process new additional evidence* when facing a decision problem. In Section 3, at the "limit deadline" $t = 0$ only the *immediately available prior information* can then be used by the agent to make a decision.

Going back to the first example in the introduction of a trader deciding among alternative investments, $t = 0$ is an abstraction for the fact that he has to choose "immediately," that is, within the smallest possible amount of time without looking at his terminal; instead, $t = 10$ means that the trader has 10 seconds to look at the terminal and decide.

The "immediately" available prior information about the alternative investments resides in the trader's memory and builds on previous experience. As such, it must be quickly retrieved, a process which however – in reality – takes some time $\varepsilon > 0$ called *latency*. This ε is the smallest possible amount of time after which a decision maker who has to choose immediately is able to answer. Thus, $p_0(a, A)$ corresponds, more precisely, to $p_\varepsilon(a, A)$, while $p_t(a, A)$ corresponds to $p_{\varepsilon+t}(a, A)$. Moreover, ε might depend on A, with larger sets commanding larger latency. This leads to a model where $p_t(a, A)$ is replaced by $p_{\varepsilon(A)+t}(a, A)$, so that

$$p_{\varepsilon(A)+t}(a, A) = \frac{e^{\frac{u(a)}{\lambda(t)}+\alpha(a)}}{\sum_{b\in A} e^{\frac{u(b)}{\lambda(t)}+\alpha(b)}}$$

and, for $t = 0$,

$$p_{\varepsilon(A)}(a, A) = \frac{e^{\alpha(a)}}{\sum_{b\in A} e^{\alpha(b)}}$$

All of our results continue to hold in this more sophisticated setting. Thus, our "abuse" in Section 3 consists in writing $p_t(a, A)$ instead of $p_{\varepsilon(A)+t}(a, A)$. This abuse is practically negligible when the minimum element of T is large relative to ε (it is safe to assume that ε is smaller than a second). What happens when the experimenter gives the subject deadlines that are smaller than ε is an interesting research question that goes beyond the scope of this paper.

[56] If numerical convergence is achieved, see the discussion in Section 2.3.

4.2 Beyond softmax

To put our softmax revealed preference analysis (Section 3) in a wider perspective we briefly discuss a general specification of a random choice process. In particular, in the next definition we generalize to our deliberation context the definition of utility for random choice rules introduced by Debreu (1958) and Davidson and Marschak (1959).

Definition 8 *A psychometric utility function u and an initial bias α on X rationalize a random choice process $\{p_t\}$ if, at each deliberation time t and all alternatives $a \neq b$ and $c \neq d$ in X:*

$$u(a) - u(b) \leq u(c) - u(d) \quad \text{and} \quad \alpha(a) - \alpha(b) \leq \alpha(c) - \alpha(d) \Longrightarrow p_t(a,b) \leq p_t(c,d).$$

It is easy to see that this amounts to requiring the existence, at each deliberation time t, of a time-dependent function ϕ_t, increasing in both arguments, such that

$$p_t(a,b) = \phi_t(u(a) - u(b), \alpha(a) - \alpha(b))$$

for all distinct alternatives $a, b \in X$.[57] Softmax is the special case

$$\phi_t(x,y) = \frac{1}{1 + e^{-\frac{x}{\lambda(t)} - y}}$$

In fact, in the binary case the softmax formula becomes:

$$p_t(a,b) = \frac{1}{1 + e^{-\frac{u(a)-u(b)}{\lambda(t)} - [\alpha(a) - \alpha(b)]}}$$

In the presence of bias, a pair (u, α) is thus needed to understand random choice processes, u alone is no longer enough – as it was, instead, in the analyses of Debreu, Davidson, and Marschak of random choice rules. The noise λ is peculiar to the softmax case, where it parameterizes the exponential form of ϕ_t.

A natural question, which may be explored in future research, is how the analysis of Debreu (1958) may generalize in this setup, determining which conditions on a random choice process ensure the existence of the pair (u, α). A result along these lines would be in relation to Debreu (1958) what our softmax representation theorem (Theorem 7) is to Luce (1959).

4.3 Testing of the Metropolis-DDM algorithm and possible extensions

A novel object such as the Metropolis-DDM algorithm calls for accurate experimental tests, both based on existing datasets and on ad hoc experiments focused to trace its specific moving parts. This research is ongoing, it rests on the dialogue between axioms and mechanisms introduced in this paper, but goes well beyond the scope of this work. More in general, the hypothesis that sequential binary comparisons form the basic structure of multialternative choice remains an open challenge that we plan to explore.

A natural extension of the Metropolis-DDM is obtained by considering either more general DDMs – like the ones with random starting points *à la* Ratcliff et al. (2016), or with collapsing barriers *à la* Churchland, Kiani, and Shadlen (2008), or with attentional weights *à la* Krajbich, Armel, and Rangel (2010) – or recent non-DDM models – like the optimal coding models of Rustichini and Padoa Schioppa (2015) and Rustichini et al. (2017), or the extrema detection models of Stine et al. (2020). In all these cases, very weak assumptions guarantee that the resulting algorithm has a stationary distribution which approximates the resulting stochastic choice behavior; but, since these models do not have closed form binary choice probabilities, the chances of finding a closed form for such a stationary distribution do not seem to be high; but see Baldassi et al. (2020b).

[57] The real-valued function ϕ_t has domain $(\operatorname{Im} u - \operatorname{Im} u) \times (\operatorname{Im} \alpha - \operatorname{Im} \alpha)$.

4.4 Quantal response equilibrium

The softmax functional form can be regarded as a formalization of the Discovered Preference Hypothesis (DPH) outlined in Plott (1996). According to this hypothesis, decision makers learn how their basic needs are satisfied by the different alternatives in the choice environment through a process of reflection and practice that, in the long run, leads to optimizing behavior.

Reflection is readily captured in our model by the deliberation time. If one considers applications to repeated choice situations, the DPH points to a different natural interpretation of the set T. Instead of a deadline, each t of T may represent the number of times that the decision maker has been facing choice problem A. Under this interpretation, softmax can be seen as capturing preference discovery through practice.

MNL, that is, unbiased softmaximization is the form that preference discovery takes in the *quantal response equilibrium* (*QRE*) theory of McKelvey and Palfrey (1995). In their theory, t is the number of times an agent played the game, and thus measures his experience level, $u(a)$ is the expected payoff of action a, and $\lambda(t)$ indexes the agent'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 decreases as time passes and the decision making environment is better understood.[58]

Our axiomatic and neuro-computational characterizations of MNL can thus be seen as 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 developed in this paper. Goeree, Holt and Palfrey (2016) give a broad perspective of its different applications.[59]

5 Acknowledgements

We thank Jerome Adda, Pierpaolo Battigalli, Patrick Beissner, Andrei Borodin, Roberto Corrao, Federico Echenique, Agata Farina, Marcelo Fernandez, Nicola Gennaioli, Salvatore Greco, Loic Grenie, Philip Holmes, Ryota Iijima, Michael Konig, Giacomo Lanzani, Shuige Liu, Jay Lu, Laura Maspero, Filip Matejka, Paulo Natenzon, Thomas Palfrey, Charles Plott, Antonio Rangel, Roger Ratcliff, Giorgia Romagnoli, Kota Saito, Larry Samuelson, Vaibhav Srivastava, Jakub Steiner, Tomasz Strzalecki, Adam Szeidl (the editor), four anonymous referees, as well as several seminar audiences for stimulating comments. We especially thank Carlo Baldassi and Giulio Principi for very useful discussions and Marco Pirazzini for coding assistance. Simone Cerreia-Vioglio and Massimo Marinacci gratefully acknowledge the financial support of ERC (grants SDDM-TEA and INDIMACRO, respectively), Fabio Maccheroni acknowledges the financial support of PRIN (grant 2017CY2NCA).

6 Data Availability Statement

This paper does not involve analysis of external data. The code generating all simulations' plots is available at `https://doi.org/10.5281/zenodo.6599226`

[58] 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 alternatives in A before choosing.

[59] For an epistemic characterization of QRE, see the recent Liu and Maccheroni (2021).

7 References

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.

Alos-Ferrer, C., Fehr, E., and Netzer, N. (2021). Time will tell: recovering preferences when choices are noisy. *Journal of Political Economy*, 29, 1828-1877.

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

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.

Ariely, D., and Wertenbroch, K. (2002). Procrastination, deadlines, and performance: Self-control by precommitment. *Psychological science*, 13, 219-224.

Baldassi, C., Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., and Pirazzini, M. (2020a). A behavioral characterization of the Drift Diffusion Model and its multi-alternative extension for choice under time pressure. *Management Science*, 66, 5075-5093.

Baldassi, C., Maccheroni, F., Marinacci, M., and Pirazzini, M. (2020b). Time-constrained sequential search in multialternative choice problems. arXiv:2005.01081

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

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

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.

Block, H. D., and Marschak, J. (1960). Random ordering and stochastic theories of responses. In Olkin, I., Ghurye, S., Hoeffding, W., Madow, W., and Mann, H. (eds.). *Contributions to probability and statistics* (97-132). Stanford University Press, Stanford

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.

Bordalo, P., Gennaioli, N., and Shleifer, A. (2020). Memory, attention, and choice. *Quarterly Journal of Economics*, 135, 1399-1442.

Bordalo, P., Conlon, J., Gennaioli, N., Kwon, S., and Shleifer, A. (2021). Memory and probability. Mimeo.

Callaway, F., Rangel, A., and Griffiths, T. (2019). Fixation patterns in simple choice are consistent with optimal use of cognitive resources. PsyArXiv, 4 Mar. 2019.

Caplin, A., and Dean, M. (2014). Revealed preference, rational inattention, and costly information acquisition. NBER WP 19876

Cerreia-Vioglio, S., Maccheroni, F., Marinacci, M., and Rustichini, A. (2021a). Multinomial logit processes and preference discovery: inside and outside the black box. arXiv:2004.13376.

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

Churchland, A. K., Kiani, R., and Shadlen, M. N. (2008). Decision-making with multiple alternatives. *Nature Neuroscience*, 11, 693-702.

Clithero, J. A. (2018). Improving out-of-sample predictions using response times and a model of the decision process. *Journal of Economic Behavior & Organization*, 148, 344-375.

Davidson, D., and Marschak, J. (1959). Experimental tests of a stochastic decision theory. In *Measurement: Definitions and Theories* (C. W. Churchman, ed.), Wiley, New York.

Dean, M., and Neligh, N. L. (2019). Experimental tests of rational inattention. Mimeo.

Dewan, A., and Neligh, N. (2020). Estimating information cost functions in models of rational inattention. *Journal of Economic Theory*, 187, 105011.

Debreu G. (1954). Representation of a preference ordering by a numerical function. In *Decision Processes* (R. M. Thrall, C. H. Coombs and R. L. Davis, eds.), Wiley, New York.

Debreu G. (1958). Stochastic choice and cardinal utility, *Econometrica*, 26, 440-444.

Debreu G. (1964). Continuity properties of Paretian utility, *International Economic Review*, 5, 285-293.

de Palma, A., Fosgerau, M., Melo, E., Shum, M. (2019). Discrete choice and rational inattention: A general equivalence result. Mimeo.

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.

Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N., and Pouget, A. (2012). The cost of accumulating evidence in perceptual decision making. *Journal of Neuroscience*, 32, 3612-3628.

Dupuis, P., and Ellis, R. E. (1997). *A weak convergence approach to the theory of large deviations.* Wiley, New York.

Falmagne, J. C. (1985). *Elements of Psychophysical Theory.* Oxford University Press, Oxford.

Frick, M., Iijima, R., and Strzalecki, T. (2017). Dynamic random utility. *Econometrica*, 87, 1941-2002.

Fudenberg, D., Newey, W. K., Strack, P., and Strzalecki, T. (2020). Testing the drift-diffusion model. *Proceedings of the National Academy of Sciences*, 117, 33141-33148.

Fudenberg, D., Strack, P., and Strzalecki, T. (2018). Speed, accuracy, and the optimal timing of choices. *American Economic Review*, 108, 3651-3684.

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

Gabaix, X., Laibson, D., Moloche, G., and Weinberg, S. (2006). Costly information acquisition: Experimental analysis of a boundedly rational model. *American Economic Review*, 96, 1043-1068.

Geyer, C. (2011). Introduction to Markov Chain Monte Carlo. In *Handbook of Markov Chain Monte Carlo* (S. Brooks, A. Gelman, G. L. Jones, X. Meng, eds.), CRC Press, Boca Raton.

Georgescu-Roegen, N. (1936). The pure theory of consumers behavior. *Quarterly Journal of Economics*, 50, 545-593.

Georgescu-Roegen, N. (1958). Threshold in Choice and the Theory of Demand. *Econometrica*, 26, 157-168.

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.

Grunert, K. G., and Wills, J. M. (2007). A review of European research on consumer response to nutrition information on food labels. *Journal of Public Health*, 15, 385-399.

Hardy, G. H., Littlewood, J. E., and Polya, G. (1934). *Inequalities.* Cambridge University Press, Cambridge.

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

Huseynov, S., and Palma, M. A. (2021). Food decision-making under time pressure. *Food Quality and Preference*, 88, 104072.

Jang, A., Sharma, R., and Drugowitsch, J. (2020). Optimal policy for attention-modulated decisions explains human fixation behavior. bioRxiv 2020.08.04.237057

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

Kaufman, E.L., Lord, M.W., Reese, T.W., and Volkmann, J. (1949). The discrimination of visual number. *American Journal of Psychology*, 62, 498-525.

Kelly, F. P. (2011). *Reversibility and stochastic networks.* Cambridge University Press, Cambridge.

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

Kolmogorov, A. (1936). Zur theorie der Markoffschen ketten. *Mathematische Annalen*, 112, 155-160.

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.

Kreps, D. M. (1988). *Notes on the theory of choice.* Westview, Boulder.

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

Liu, S., and Maccheroni, F. (2021). Rationalization, quantal response equilibrium, and robust outcomes in large populations. arXiv:2106.16081

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

Luce, R. D. (1957). A theory of individual choice behavior. Mimeo.

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

Luce, R. D., and Suppes, P. (1965). Preference, utility and subjective probability. In Luce, R. D., Bush, R. R., and Galanter, E. (eds.). *Handbook of mathematical psychology*, vol. 3 (pp. 249-410). 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, Providence.

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, New York.

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.

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

Natenzon, P. (2019). Random choice and learning. *Journal of Political Economy*, 127, 419-457.

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

Papandreou, A. G. (1953). An experimental test of an axiom in the theory of choice. *Econometrica*, 21, 477.

Papandreou, A. G., Sauerlender, O. H., Brownlee, O. H., Hurwicz, L., and Franklin, W. (1957). A test of a stochastic theory of choice. *University of California Publications in Economics*, 16, 1-18.

Pinsky, M., and Karlin, S. (2011). *An introduction to stochastic modeling.* Academic Press, New York.

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, London.

Quandt, R. E. (1956). A probabilistic theory of consumer behavior. *Quarterly Journal of Economics*, 70, 507-536.

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, New York.

Rasch, G. (1961). On general laws and the meaning of measurement in psychology. In Neyman, J. (ed.). *Proceedings of the fourth Berkeley symposium on mathematical statistics and probability*, v. 4 (pp. 321-333). University of California Press.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests.* Danish Institute for Educational Research. Expanded edition (1980), University of Chicago Press, Chicago.

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

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. (1955). On a new axiomatic theory of probability. *Acta Mathematica Hungarica*, 6, 285-335.

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

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.

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.

Rustichini, A., Conen, K.E., Cai, X., Padoa-Schioppa, C. (2017). Optimal coding and neuronal adaptation in economic decisions. *Nature Communications*, 8, 1208.

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

Saltzman, I.J., and Garner, W.R. (1948). Reaction time as a measure of span of attention. *Journal of Psychology*, 25, 227-241.

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

Shapley, L. S. (1975). Cardinal utility from intensity comparisons. RAND WP R-1683-PR.

Silayoi, P., and Speece, M. (2004). Packaging and purchase decisions: An exploratory study on the impact of involvement level and time pressure. *British food journal*, 106, 607-628.

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

Stine, G. M., Zylberberg, A., Ditterich, J., and Shadlen, M. N. (2020). Differentiating between integration and non-integration strategies in perceptual decision making. *Elife*, 9, e55365.

Stroock, D. (2005). *An introduction to Markov processes.* Springer, Berlin.

Suppes, P., and Winet, M. (1955). An axiomatization of utility based on the notion of utility differences. *Management science*, 1, 259-270.

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*, 30, 305-314.

Tajima, S., Drugowitsch, J., and Pouget, A. (2016). Optimal policy for value-based decision-making. *Nature Communications*, 7, 12400.

Tajima, S., Drugowitsch, J., Patel, N., and Pouget, A. (2019). Optimal policy for multi-alternative decisions. *Nature Neuroscience*, 22, 1503-1511.

Thurstone, L. L. (1929). Theory of attitude measurement. *Psychological Review*, 36, 222-241.

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

Valkanova, K. (2021). Markov stochastic choice. Mimeo.

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

Wakker, P. P. (1989). *Additive representations of preferences: A new foundation of decision analysis.* Springer.

Warlop, L., Ratneshwar, S., and van Osselaer, S. M. (2005). Distinctive brand cues and memory for product consumption experiences. *International Journal of Research in Marketing*, 22, 27-44.

Webb, R. (2019). The (neural) dynamics of stochastic choice. *Management Science*, 65, 230-255.

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

Zhang, T. (2006a). From ε-entropy to KL-entropy: Analysis of minimum information complexity density estimation. *Annals of Statistics*, 34, 2180-2210.

Zhang, T. (2006b). Information-theoretic upper and lower bounds for statistical estimation. *IEEE Transactions on Information Theory*, 52, 1307-1321.