School of Economics & Finance Discussion Papers

Imperfect Attention and Menu Evaluation

Paola Manzini and Marco Mariotti

School of Economics and Finance Discussion Paper No. 1319
21 Oct 2013 (revised 24 Aug 2014)
JEL Classification: D01
Keywords: bounded rationality, stochastic choice

*School of Economics & Finance
Discussion Paper Series
issn 0962-4031
http://ideas.repec.org/s/san/wpecon.html
info: econ@st-andrews.ac.uk*

Imperfect Attention and Menu Evaluation*

Paola Manzini Marco Mariotti†

This version: August 2014

Abstract

We model the choice behaviour of an agent who is vNM rational but imperfectly attentive. We define inattention axiomatically through preference over menus and *endowed alternatives*: an agent is inattentive if it is better to be endowed with an alternative a than to be allowed to pick a from a menu in which a is is the best alternative. This property and vNM rationality imply that the agent notices each alternative with a given menu-dependent probability (attention parameter) and maximises a menu independent and deterministic utility function over the alternatives he notices. Preference for flexibility restricts the model to menu independent attention parameters as in Manzini and Mariotti [26]. Our theory explains anomalies (e.g. the attraction and compromise effect) that Random Utility Maximisation cannot accommodate.

J.E.L. codes: D0.

Keywords: bounded rationality, stochastic choice

*We are grateful to Yorgos Gerasimou, Pietro Ortoleva, Kemal Ozbek and Dan Sasaki for insightful comments, to various seminar audiences (notably attendees at BRIC2014 in Barcelona), and to ESRC for financial support through grant ES/J012513/1.

†Both authors at School of Economics and Finance, University of St. Andrews, Castlecliffe, The Scores, St. Andrews KY16 9AL, Scotland, U.K. (e-mail Manzini: paola.manzini@st-andrews.ac.uk; e-mail Mariotti: marco.mariotti@st-andrews.ac.uk).

1 Introduction

In this paper we propose a model of choice with imperfect attention and a method to build such a model. By 'imperfect attention' we mean any cognitive shortcoming or bias that causes the agent to ignore, with some probability, some of the available options of choice.

For example, you have to choose a new computer, or a new house in a large city, or you are planning a complex trip. You do not examine, or even notice, all the available options. Instead, you focus your attention on a subset of them, your *consideration set*. In general, this type of attention failure is likely when the available options are numerous (in the case of computers and houses, several hundreds), or complex, or they change frequently (due e.g. to technological change), or are difficult to store in memory. As an extreme but relevant example of this last case - to which we return later - think of having to take a daily pill. Here the option set on any given day consists of exactly one element, and an attention failure means failing to take that option.

How should imperfect attention be modeled? Existing axiomatic work (such as Brady and Rehbeck [5], Caplin and Dean [6], Echenique, Saito and Tserenjigmid [11], Manzini and Mariotti [26] and Masatlioglu *et al.* [27]) extends in various ways classical revealed preference analysis by focusing directly on properties of a choice procedure that maps menus into choice data (where choice data may be highly non-standard in that they incorporate richer information than crude choices, as in Caplin and Dean [6]). Here we explore a different methodology. We infer the characteristics of the choice procedure *indirectly,* through comparisons between the values of different choice situations to an imperfectly attentive decision maker. The implied choice procedure generalises the model in Manzini and Mariotti [26] in a useful direction, addressing some prominent anomalies such as the attraction and compromise effects.

The key idea for our definition of imperfect attention is to compare the values of two types of situations:

- choosing from *menus* of alternatives
- being *endowed* with an alternative.

An endowed alternative is one that the agent simply 'has' or 'is given', without going through a process of choice. For an agent who is fully attentive *and* preference maximising there is no difference between being endowed with an alternative a and choosing from a menu x in which a is its best alternative. But if the agent may fail to consider a because of a lapse of attention, there is a gap between the value x and the value of a. This agent must be strictly better off in the situation in which he is endowed with a than in the situation in which he can pick a from x.

An endowed alternative could have, for example, the nature of a default option. Imagine someone needing a health insurance scheme. Compare two situations: (A) He is automatically enrolled in a plan a by the public authority; (B) he chooses from the market, where a is available. An imperfectly attentive agent is one who is strictly better off in A than in B, *if* plan a happens to be the best option. Having to take a daily tablet offers a second example: if you are better off with a transdermal delivery implant, then you suffer from imperfect attention by our definition.

This is a deliberately minimal and 'reduced form' definition of imperfect attention, which does not commit to any specific hypothesis on what causes attention to be imperfect, or on how the agent builds a consideration set. It ignores the details of any underlying information processing mechanism or search strategy. It ignores the specific features of the alternatives that may render them more or less salient. Yet it captures a core aspect of imperfect attention that is common to many of its possible causes. In fact, it is hard to deny that whatever it is that makes attention imperfect, it has precisely the effect of making the agent worse off when choosing from a menu than he would be by being endowed with the optimal option in that menu (below we also explain why such a preference on our domain permits a distinction from other psychological phenomena, such as temptation).

We study inattentive preferences over lotteries that are otherwise von Neumann Morgenstern (henceforth vNM) rational. We show that the implied choices *from* menus can then be represented as the outcome of deterministic, menu-independent utility maximisation over a stochastic consideration set. Moreover, we pin down four distinctive aspects of the stochastic process that generates the consideration set:

- *Menu dependence in consideration*: the probability distribution over consideration sets changes from menu to menu.

- *Stochastic independence in consideration*: the probability of considering any group of alternatives in a menu is the product of the probabilities of considering each of those alternatives individually.

- *Monotonicity in consideration*: the probability of considering an alternative in a menu is weakly decreasing with respect to the addition of a top alternative to the menu.

- *Menu independence in consideration* is equivalent to a form of *preference for flexibility* (the agent is better off with larger menus).

More in detail, we posit a preference ranking $\succsim$ of (riskless) endowed alternatives a, menus x (finite sets of alternatives), and non-trivial lotteries with as and xs as consequences. The relation $\succsim$ expresses ordinal comparisons of how well off the agent is in different situations.[1] Thus, we consider statements such as 'it is better to choose from menu x rather than from menu y'; or 'it is better to choose from menu x than being endowed with alternative a'; or 'it is better to have a fifty-fifty chance of choosing from menu x or being endowed with a, than being endowed with b'. Notably, we do not consider ranking menus of lotteries, just lotteries of menus.[2]

We introduce the axiom of *Imperfect Attention*: $a \succ x$ for any menu x that contains only alternatives not better than a (possibly including a itself). This accounts for the possibility of neglecting a (or any alternative to which a is indifferent) when choosing from x to end up with some $b \in x$ for which $a \succ b$. Observe how this axiom (and the conceptual device of endowed alternatives) helps to distinguish inattention from

[1]Let us assume for the moment that $\succsim$ is either an 'objective', normative betterness ranking, or (if a choice-based interpretation of $\succsim$ is desired) the preference of an entity whose interests are entirely aligned with those of the agent (such as a parent or a benevolent planner). Interpreting $\succsim$ as the preference over menus of the same individual that makes the choice *from* menus presents some conceptual difficulties. We address this issue in section 5.

[2]In this respect the task of ranking objects is easier than in much of the recent menu choice literature that extends Kreps [22]. See Ortoleva [32] for a recent exception.

temptation. While for non-singleton menus x the preference $a \succ x$, with a being the top alternative in x, could be the result of being tempted by inferior alternatives in x, this cannot be the case when $x = \{a\}$. But if you are inattentive you can for example fail to take the pill, so that $a \succ \{a\}$. Similarly, neither can *thought aversion* (Ortoleva [32]) or *similarity mistakes* (Payro and Ülkü [33]) explain this type of strict preference, since $\{a\}$ entails no such mistakes or thinking costs.

If $\succsim$ is also vNM, Imperfect Attention leads to an evaluation of menus in which the implied choice from menus is stochastic. Each alternative a in menu x is considered with a menu-dependent probability $\alpha(a, x)$ (the *attention parameter*), and is evaluated by a menu-independent utility value $u(a)$. Then the agent picks one of the highest utility alternatives among those which he has considered. For example, if there are three alternatives a, b and c with $u(a) > u(b) = u(c)$ the value of menu $x = \{a, b, c\}$ is

$$\begin{aligned} u(x) &= \alpha(a,x)\,u(a) \\ &\quad + (1-\alpha(a,x))\,(1-(1-\alpha(b,x))\,(1-\alpha(c,x)))\,u(b) \\ &\quad + (1-\alpha(a,x))\,(1-\alpha(b,x))\,(1-\alpha(c,x))\,u(\varnothing) \end{aligned}$$

where $u(\varnothing)$ is the utility of a default alternative or outside option which is consumed if nothing is considered (e.g. keeping your old computer, abstaining from voting, walking out of the shop).[3] Then the implicit choice from x is determined either by considering a (with the probability given by the attention parameter) and choosing it irrespective of whatever else is considered; or missing a and picking one of b or c if considered; or missing everything in x and getting the default alternative. In our main representation (theorem 1) the menu-dependence of the attention parameters is limited. Specifically, it cannot be the case in the above example that $\alpha(b, \{a, b, c\}) > \alpha(b, \{b, c\})$: that is, adding a better alternative to a menu cannot increase the attention paid to the ex-

[3]Other papers have considered the possibility of 'not choosing'. For the deterministic case, see e.g. Clark [8] and more recently Gerasimou [16]. For the stochastic case, see e.g. Corbin and Marley [9]. In the empirical IO literature using discrete choice models, it is also standad to introduce an outside option to allow for the possibility that the data do not contain all brands or models that have a positive market share (see e.g. Sovinsky Goeree [39]).

isting alternatives (as we explain later, this is a natural feature when attention derives from some types of search process).

In section 4 we relate the model to *preference for flexibility.* We introduce the axiom of *Top Expansion*, which says that adding a top alternative to a menu must make the agent better off. Top Expansion has sharp implications for the representation: it is equivalent to the property that the attention parameters can be taken to be menu independent (theorem 2). Moreover, since Top expansion is immediately seen to be equivalent to a strict version of Kreps' [22] preference for flexibility, the result also shines new light on a standard axiom of the menu preference literature: in our setting strict preference for flexibility is not related to unforeseen contingencies but rather to a property of attention.

Finally, we show in section 6 that the choice model we describe handles in a simple way observed anomalies of choice involving failures of Regularity (the property that the introduction of new alternatives in a menu reduces the probability of choosing already existing alternatives).

2 The Model

Let X be a finite set of alternatives, denoted $a, b, c, ...$ and let $\mathcal{X} = 2^X$. An element $x \in \mathcal{X}$ is called a *menu*.

Let $\Delta(X \cup \mathcal{X})$ be the set of lotteries on $X \cup \mathcal{X}$. To simplify notation we identify the degenerate lotteries in $\Delta(X \cup \mathcal{X})$ with elements of $X \cup \mathcal{X}$.

As explained, the interpretation of an $a \in X$ is that the agent is endowed with a, without having to pick it from a menu, while an $x \in \mathcal{X}$ is interpreted as the situation in which the agent has to choose an element from x. Finally, a non-degenerate element of $\Delta(X \cup \mathcal{X})$ is interpreted as a risky situation in which the agent either will have to pick an element from some menu or will be endowed with some alternative, with the identity of the menu or the alternative to be determined probabilistically.

A *preference* is a binary relation $\succsim$ on $\Delta(X \cup \mathcal{X})$, where $g' \succsim g$ for any $g, g' \in \Delta(X \cup \mathcal{X})$, interpreted as the agent being better off when facing situation g' than when facing situation g.

We consider the following properties for a preference $\succsim$ (with $\succ$ and $\sim$ denoting the asymmetric and symmetric parts, respectively):

A0 - Choice Is Valuable: $x \succ \varnothing$ for all $x \in \mathcal{X}$.

A1 - Order: $\succsim$ is a weak order.

A2 - Continuity: For all $g, g', g'' \in \Delta(X \cup \mathcal{X})$ such that $g'' \succsim g \succsim g'$, there exists $\alpha \in [0,1]$ such that $\alpha g'' + (1-\alpha) g' \sim g$.

A3 - Independence: For all $g, g', g'' \in \Delta(X \cup \mathcal{X})$ and $\alpha \in [0,1]$: $g \succsim g' \Leftrightarrow \alpha g + (1-\alpha) g'' \succsim \alpha g' + (1-\alpha) g''$.

A4 - Imperfect Attention: For all $x \in \mathcal{X}$ and $a \in X$: $a \succsim b$ for all $b \in x \Rightarrow a \succ x$.

A0 is just a definition of the range of choice situations we consider, namely ones in which the menus are 'opportunity sets' in the sense that they contain alternatives better than not choosing. A1-A3 are a version of the vNM axioms applied to the particular domain of menus and alternatives. Finally, A4 states that being endowed with any a is strictly preferable to choosing from a menu containing alternatives that are no better than a (including a itself). As argued in the introduction, A4 is the essence of imperfect attention.

Definition 1 *An imperfect attention preference (i.a.p.) is a preference that satisfies A0-A4.*

We shall establish a link between a i.a.p. and a numerical representation of menu values suggesting the two-stage stochastic process of choice -first consider, then choose- that we have discussed earlier.

A strict total order $\hat{\succ}$ of X *refines* $\succsim$ if $a \succ b \Rightarrow a \hat{\succ} b$. In the definition below recall that we identify degenerate lotteries on an outcome with the outcome itself.

Definition 2 *A vNM attention representation for $\succsim$ is a triple $(\hat{\succ}, u, \alpha)$, with $\hat{\succ}$ a strict total order of X that refines $\succsim$, $u : \Delta(X \cup \mathcal{X}) \to \mathcal{R}$ a vNM utility function representing $\succsim$, and*

$\alpha : X \cup \mathcal{X} \to (0,1)$, *such that, for all* $x \in \mathcal{X}$*:*[4]

$$u(x) = \sum_{a \in x} \prod_{b \in x : b \hat{\succ} a} (1 - \alpha(b,x))\, \alpha(a,x)\, u(a) + \prod_{a \in x} (1 - \alpha(a,x))\, u(\varnothing) \tag{1}$$

In this representation u is an *evaluation function*, with $u(a)$ in particular being the value for the agent of being endowed with a, not to be confused in the simplified notation with $u(\{a\})$ (the value of choosing from menu $\{a\}$, equal to $\alpha(a,\{a\})\,u(a)$). The function α is an *attention function* that assigns a value to the attention received by each alternative in each menu: the interpretation is that any alternative a has a chance $1 - \alpha(a,x)$ of being missed by the agent in menu x. The strict ordering $\hat{\succ}$ is a tie-breaking device that resolves indifferences between the alternatives that are considered. Under these interpretations, the agent maximises u on the set of alternatives that are both feasible and considered.

3 Analysis

We shall impose restrictions on the attention function. But even without any restriction, not all $\succsim$ have a vNM attention representation. It is instructive to go through a couple of examples.

Suppose first that $x = \{a, b\}$ and

$$x \succ a \succsim b \succ \varnothing$$

Then, if there were a vNM attention representation, we would have

$$\begin{aligned} u(x) &= \alpha(a,x)\,u(a) + (1-\alpha(a,x))\,\alpha(b,x)\,u(b) \\ &+ (1-\alpha(a,x))\,(1-\alpha(b,x))\,u(\varnothing) > u(a) \geq u(b) > u(\varnothing) \end{aligned}$$

a contradiction.[5] These preferences might result from attaching intrinsic value to the act of choice: you might value having the choice between Pravda and Wall Street Journal more than being 'endowed' with Pravda, even if ultimately you would choose

[4]We use the convention that the product over the empty set is equal to one.

[5]Even if we weakened the previous relations to $x \succsim a \succsim b \succ \varnothing$, insofar as these preferences deviate from those in the main text, the agent would clearly have to have perfect attention at least for one alternative, a case also excluded from the representation.

Pravda anyway. A4, which aims to isolate the attention motive, directly rules out such preferences.

A more subtle example of preference that cannot be captured by a vNM attention representation is as follows. Let $x = \{a, b\}$ and

$$\begin{aligned} a &\succ b \succ \{b\} \\ a &\succ x \succ \{b\} \\ x &\sim ka + (1-k)\, b \text{ for some } k \in (0,1) \end{aligned} \tag{2}$$

The preference in the second line is not essential and serves only to simplify the proof of the Observation below. The key preference is in the third line. The indifference between a menu and a mixture of the (endowed) alternatives it contains could be naturally explained by *correlation in consideration*. For example, suppose that when faced with the menu x the agent considers only a with probability k and considers only b with probability $(1-k)$, leading to the indifference in 2. However, this consideration pattern obviously cannot be generated independently by means of parameters $\alpha(a, x)$ and $\alpha(b, x)$. Another explanation, that does not rely on correlation, is that with probability k the agent considers both a and b (then choosing a) and with probability $(1-k)$ he considers only b. This consideration pattern can be generated independently with attention parameters $\alpha(a, x) = k$ and $\alpha(b, x) = 1$, but this type of *full attention* for an alternative is not admissible in our framework. We show that, short of assuming correlation, full attention for b is implied by the above preferences, leading to a contradiction:

Observation: The preferences in (2) have no vNM attention representation.

Proof. Suppose that a vNM attention representation exists, and let u be such a representation normalised by $u(\varnothing) = 0$. Then:

$$ku(a) + (1-k)\, u(b) = u(x) = \alpha(a, x)\, u(a) + (1 - \alpha(a, x))\, \alpha(b, x)\, u(b)$$

implying $\alpha(b, x) = 1$. Since $a \succ x \succ \{b\}$, we have $u(a) > u(x) > u(\{b\})$, so that

$$u(x) = k' u(a) + \left(1 - k'\right) u(\{b\})$$

for some $k' \in (0,1)$. Then, since $u(\{b\}) = \alpha(b,\{b\})\,u(b) + 0$ by the normalised representation, we have

$$u(x) = \alpha(a,x)\,u(a) + (1-\alpha(a,x))\,u(b) = k'u(a) + \left(1-k'\right)\alpha(b,\{b\})\,u(b)$$

implying $\alpha(b,\{b\}) = 1$. It follows that $u(\{b\}) = u(b)$, so that u cannot represent $b \succ \{b\}$, a contradiction. ■

This example highlights what kind of pattern is excluded by the lack of correlation that is implicit in a vNM representation. While no single axiom in a i.a.p. rules out the above preference configuration directly, we shall see that the axioms as a whole do so.

Indeed, Theorem 1 below says that Choice Is Valuable and Imperfect Attention together with vNM rationality are equivalent to a vNM attention representation, with an additional twist.

It turns out that i.a.p. also limits the way attention for an alternative may depend on the menu. In principle, one may imagine that adding an alternative to a menu increases the attention paid to an existing alternative. This could be the case, for example, through a similarity effect, when the new alternative is similar to the existing one. Or, a product offered by a multiproduct firm (e.g. a program of a media company) may draw attention to other products offered by the same firm.[6] But if preferences are i.a.p. this effect can be excluded when the new alternative is better than the existing ones. Preferences have a vNM attention representation if and only if they have a vNM attention representation in which the addition of a new top alternative to a menu cannot increase the attention received by any of the existing alternatives.

A vNM attention representation $(\hat{\succ}, u, \alpha)$ satisfies *monotonicity in consideration* if

$$b \hat{\succ} a \text{ for all } a \in x \Rightarrow \alpha(a,x) \geq \alpha(a, x \cup \{b\}) \text{ for all } a \in x.$$

Monotonicity in consideration (and much more) is implied for example when the attention parameters have the Luce structure $\alpha(a,x) = \frac{\lambda(a)}{\sum_b \lambda(b)}$, where λ is a strictly positive real valued function of the alternatives. Another situation where monotonicity in consideration is a natural property is when the vNM attention representation

[6] As in Eliaz and Spiegler [12].

is the reduced form of a search process. Suppose the agent searches randomly and sets a reservation value, at which he stops searching. Then adding a top alternative a should make it less likely that the other alternatives will be seen, since the agent will stop whenever he finds a. On the other hand, monotonicity in consideration excludes the effect studied in Payro and Ülkü [33], whereby the agent may consider inferior alternatives just because they are similar to the top alternative in the menu.[7]

Theorem 1 *The relation $\succsim$ is an imperfect attention preference if and only if it has a vNM attention representation $(\hat{\succ}, u, \alpha)$ that satisfies monotonicity in consideration.*

The proof of the theorem is not difficult but it is long and is thus relegated to an appendix. The logic of the proof is that the axioms enable an iterated 'peeling off' procedure that makes any menu x indifferent to a lottery over two outcomes, one of them being a sub-menu obtained by removing an alternative in x, as follows. Suppose for simplicity that preferences are strict and number the alternatives in x from best to worst as $x = \{a_1, ...a_K\}$. Suppose also that $x \succ x \setminus \{a_1\}$. Given that $a_1 \succ x$ by A4, we can construct (thanks to Continuity) a lottery $\alpha a_1 + (1-\alpha) x \setminus \{a_1\}$ for some unique $\alpha \in (0,1)$ that is indifferent to x, so that by the vNM axioms this can be represented as $u(x) = \alpha u(a_1) + (1-\alpha) u(x \setminus \{a_1\})$ for some vNM utility u. Then we can iterate the process applying the argument successively to $x \setminus \{a_1\}$, $x \setminus \{a_1, a_2\}$,..., and show that the resulting formula has the required properties. The procedure is markedly less straightforward when $x \setminus \{a_1\} \succsim x$, but it retains the same flavour.

Observe that a i.a.p. may be such that two menus x and y of the same cardinality that contain indifferent alternatives (in the sense that there exists a bijection f from x to y with $f(a) \sim a$ for all $a \in x$) are not necessarily indifferent. For example, if $a \sim b \succ c$ it can be the case that $\{a, c\} \succ \{b, c\}$. In the interpretation we are giving preferences, this is not a puzzling phenomenon: the discrepancy in the values of menus that contain indifferent alternatives can be explained by the different levels of attention received by the alternatives in the two menus - even if indifferent between a taxi and a bus, you may fail to spot bus number 38 out of many buses outside a station, while any

[7]Payro and Ülkü do not explicitly mention consideration effects, but this is a possible interpretation of their choice model.

taxi would do. In the previous example, $\alpha(a,\{a,c\}) > \alpha(b,\{b,c\})$ and $\alpha(c,\{a,c\}) = \alpha(c,\{b,c\})$ would rationalise the preference.

A second observation concerns the uniqueness of the representation. As is obvious, the same attention parameters work with any positive affine transformation of u. In general, however, while thanks to the vNM axioms the evaluation function u is cardinally unique (given the preference), it may only be possible to restrict, but not to pin down uniquely, the attention functions α that are compatible with a given preference. Uniqueness is not guaranteed given the nonlinear way the attention parameters enter the representation. However, the monotonicity condition in the representation helps to put some bounds on the attention parameters. This point is illustrated in the following:

Example 1 *Let $X = \{a,b\}$, suppose that preferences satisfy A0-A4 and are such that $a \succ b \succ \{a,b\} \succ \{a\} \succ \{b\}$, and suppose that u represents preferences with*

$X \cup \mathcal{X}$	a	b	$\{a,b\}$	$\{a\}$	$\{b\}$	$\varnothing$
$u(\cdot)$	U	pU	qpU	rU	sU	0

where $U > 0$, $p,q,r,s \in (0,1)$ and $s < r < qp$. It is easy to see (see Appendix A.2 for details) that the representation of $u(\{a\})$, $u(\{b\})$ and $u(X)$, the restriction that $\alpha(b,X) \in (0,1)$ and monotonicity in consideration imply that

$$\alpha(a,X) \in \left[\frac{qp-s}{1-s}, qp\right)$$

Since $qp > \frac{qp-s}{1-s} \Leftrightarrow qp < 1$ the interval is non-degenerate.

There is of course a second source of non-uniqueness in the way that $\hat{\succ}$ breaks indifferences in $\succsim$, but this is less important since the *evaluation* of a menu is not impacted by the exact choice of $\hat{\succ}$, the overall attention enjoyed by alternatives in a menu that belong to the same indifference class being independent of the choice of $\hat{\succ}$. For menu evaluation purposes, regardless of how the alternatives within the same indifference class are ranked by $\hat{\succ}$ what matters is the probability that *some* alternative in a's indifference class is noted by the decision maker. There is no bonus for noticing more than one alternative in any indifference class, given that only the single alternative that is ultimately chosen determines value. So at least in this respect the lack of a complete identification of the attention parameters does not matter. Note however that together with Monotonicity in consideration the exact specification of $\hat{\succ}$ may change the admissible range of the attention parameters $\alpha(a, x)$ (this is illustrated in Appendix A.3).

Full identifiability of the attention parameters is achieved when the model is specialised, as shown next.

4 Menu independence and preference for flexibility

In this section we study the important special case in which the attention for an alternative a is independent of the menu in which a appears. In the case of brands, for example, there is some evidence that the salience of each brand is independent of which other brands are available (van Nierop *et al. [31]*).

The 'choice from menu' properties of the case with menu independent attention have already been explored in Manzini and Mariotti [26]. Here we are interested in the menu preference counterpart of this feature: what are the additional properties of preferences over menus/alternatives, beside vNM rationality and Imperfect Attention, that permit to assume menu independent attention? The answer to this questions turns out to be both simple and intriguing, and it hinges on how preferences behave with respect to *expansions* of menus.

We have seen that the addition of top alternatives is (weakly) detrimental for the chance of the existing alternatives to be considered. This fact points to a key feature of the representation in theorem 1. Suppose that $a \succsim b$. This preference is compatible with:

$$\begin{aligned}
\alpha(b, \{a, b\}) &< \alpha(b, \{b\}) \frac{1}{(1-\alpha(a, \{a, b\}))} - \frac{\alpha(a, \{a, b\})}{(1-\alpha(a, \{a, b\}))} \frac{u(a)}{u(b)} \\
&\Leftrightarrow \alpha(b, \{b\}) u(b) > \alpha(a, \{a, b\}) u(a) + (1-\alpha(a, \{a, b\})) \alpha(b, \{a, b\}) u(b) \\
&\Leftrightarrow \{b\} \succ \{a, b\}
\end{aligned}$$

In words, in a i.a.p. *adding a new top alternative to a menu may decrease the value of that menu*. This can happen when adding the new alternative reduces by a sufficient amount the attention paid to the best existing alternatives, provided that the attention paid to the new top alternative isn't too large. The following axiom excludes this situation:

A5 - Top expansion: For all $x \in \mathcal{X}$ and $a \in X$: $a \notin x$ and $a \succsim b$ for all $b \in x \Rightarrow \{a\} \cup x \succ x$.

A5 can be seen in part as a standard rationality requirement. There is however a residual component of imperfect attention in this axiom, in that it allows the possibility that a is indifferent to an existing best alternative while at the same time $\{a\} \cup x \succ x$. On the contrary, a perfectly attentive and rational agent should regard the menus $\{a\} \cup x$ and x as indifferent. Yet a rational but imperfectly attentive agent might miss a top alternative with higher probability when there are fewer of them.

In a vNM attention representation $(\hat{\succ}, u, \alpha)$, α is *menu independent* if, for all $x, y \in \mathcal{X}$ and for all $a \in x \cap y$, $\alpha(a, x) = \alpha(a, y)$. In this case we write for simplicity $\alpha(a)$ instead of $\alpha(a, x)$.

Theorem 2 *An imperfect attention preference $\succsim$ satisfies A5 if and only if it has a vNM attention representation $(\hat{\succ}, u, \alpha)$ in which α is menu independent.*

Proof. Necessity. Suppose the representation holds with α menu independent. Let $a \notin x$ and $a \succsim b$ for all $b \in x$. Then observing that $x \cup \{a\} = \{b \in x : b \hat{\succ} a\} \cup \{a\} \cup$

$\{b \in x : a \hat{\succ} b\}$, we have

$$
\begin{aligned}
u(x \cup \{a\}) - u(x) &= \\
&= \sum_{b \in x : b \hat{\succ} a} \prod_{c \in x : c \hat{\succ} b} (1 - \alpha(c)) \alpha(b) u(b) \\
&\quad + \alpha(a) u(a) \prod_{b \in x : b \hat{\succ} a} (1 - \alpha(b)) \\
&\quad + (1 - \alpha(a)) \sum_{b \in x : a \hat{\succ} b} \prod_{c \in x : c \hat{\succ} b} (1 - \alpha(c)) \alpha(b) u(b) - \\
&\quad + (1 - \alpha(a)) \prod_{b \in x} (1 - \alpha(b)) u(\varnothing) \\
&\quad - \sum_{b \in x} \prod_{c \in x : c \hat{\succ} b} (1 - \alpha(c)) \alpha(b) u(b) - \prod_{b \in x} (1 - \alpha(b)) u(\varnothing) \\
&= \alpha(a) \prod_{b \in x : b \hat{\succ} a} (1 - \alpha(b)) \left(u(a) - \sum_{b \in x : a \hat{\succ} b} \prod_{c \in x : c \hat{\succ} b} (1 - \alpha(c)) \alpha(b) u(b) + \right. \\
&\quad \left. - \prod_{b \in x : a \hat{\succ} b} (1 - \alpha(b)) u(\varnothing) \right) \\
&> 0
\end{aligned}
$$

where the last inequality follows from the fact that $u(a) \geq u(b)$ for all $b \in x$ such that $a \hat{\succ} b$ and that the sum of the coefficients on the last two terms add up to less than unity.[8]

For sufficiency, theorem 1 ensures that $\succsim$ has a vNM attention representation. To prove that α is menu independent in the representation of the theorem we begin by proving two claims.

Claim 1: For all $x \in \mathcal{X}$: $a \notin x$ and $a \succsim b$ for all $b \in x \Rightarrow \alpha(b, x) = \alpha(b, x \cup \{a\})$ for all $b \in x$.

Proof. If A5 holds, the second and third case we have examined in the proof of theorem 1 (i.e. $x \backslash \{a\} \succsim x$) are excluded. So the first case $x \succ x \backslash \{a\}$ applies and yields $\alpha(b, x) = \alpha(b, x \cup \{a\})$ for all $b \in x$. □

Claim 2: For all $x, y \in \mathcal{X}$: $a \notin x \cup y$ and $a \succsim b$ for all $b \in x \cup y \Rightarrow \alpha(a, x) = \alpha(a, y)$.

[8] This holds since

$$
\begin{aligned}
\sum_{b \in x : a \hat{\succ} b} \prod_{c \in x : c \hat{\succ} b} (1 - \alpha(c)) \alpha(b) u(b) + \prod_{b \in x : a \hat{\succ} b} (1 - \alpha(b)) &< \\
\sum_{b \in x} \prod_{c \in x : c \hat{\succ} b} (1 - \alpha(c)) \alpha(b) u(b) + \prod_{b \in x} (1 - \alpha(b)) &= 1
\end{aligned}
$$

Proof. Since by A0 and A4 $a \succ \{a\} \succ \varnothing$, by the vNM axioms there exists a unique $\alpha_{a,\{a\}} \in (0,1)$ such that

$$\{a\} \sim \alpha_{a,\{a\}} a + \left(1 - \alpha_{a,\{a\}}\right) \varnothing$$

Similarly, considering any $x \in \mathcal{X}$ such that $a \notin x$ and $a \succsim b$ for all $b \in x$, it follows by A4 and A5 that $a \succ \{a\} \cup x \succ x$. By A2 then there exist a unique $\alpha_{a,x} \in (0,1)$ such that

$$\{a\} \cup x \sim \alpha_{a,x} a + \left(1 - \alpha_{a,\{a\}}\right) x$$

By A3 it must be

$$\begin{aligned}
& k\{a\} + (1-k)\left[\{a\} \cup x\right] \\
& \sim k\left[\alpha_{a,\{a\}} a + \left(1 - \alpha_{a,\{a\}}\right) \varnothing\right] + (1-k)\left[\alpha_{a,x} a + \left(1 - \alpha_{a,\{a\}}\right) x\right] \\
& = \left(k\alpha_{a,\{a\}} + (1-k)\,\alpha_{a,x}\right) a + k\left(1 - \alpha_{a,\{a\}}\right) \varnothing + (1-k)\left(1 - \alpha_{a,\{a\}}\right) x
\end{aligned} \tag{3}$$

for any $k \in (0,1)$. Fix one such k. Since it is also the case (by A4) that $a \succ \{a\}$ and $a \succ \{a\} \cup x$, then $a \sim ka + (1-k)\,a \succ k\{a\} + (1-k)\left(\{a\} \cup x\right)$ by A3. In addition, also by A3, $k\{a\} + (1-k)\left(\{a\} \cup x\right) \succ k\varnothing + (1-k)\,x$. Therefore

$$a \succ k\{a\} + (1-k)\left(\{a\} \cup x\right) \succ k\varnothing + (1-k)\,x$$

and, by A2 there exists a unique $\gamma \in (0,1)$ such that

$$k\{a\} + (1-k)\left(\{a\} \cup x\right) \sim \gamma a + (1-\gamma)\left[k\varnothing + (1-k)\,x\right]$$

But this is simply expression (3), so that it must be

$$\left.\begin{aligned}
\left(k\alpha_{a,\{a\}} + (1-k)\,\alpha_{a,x}\right) &= \gamma \\
k\left(1 - \alpha_{a,\{a\}}\right) &= k\,(1-\gamma) \\
(1-k)\left(1 - \alpha_{a,\{a\}}\right) &= (1-\gamma)\,(1-k)
\end{aligned}\right\} \Leftrightarrow \alpha_{a,\{a\}} = \alpha_{a,x}$$

Applying the same argument to a $y \in \mathcal{X}$ such that $a \notin y$ and $a \succsim b$ for all $b \in y$ yields $\alpha_{a,y} = \alpha_{a,\{a\}} = \alpha_{a,x}$, proving the claim. □

To prove sufficiency, take $x, y \in \mathcal{X}$ and $a \in x \cap y$ (if $a \notin x \cap y$ for all a then there is nothing to prove). Let $x_L = \{b \in x : a \succ b\}$, enumerate arbitrarily the elements other than a in $x \backslash x_L$, that is let $x \backslash x_L = \{a, b_1, ...b_n\}$, and let $x_i = x_L \cup \{b_1, ...b_i\}$ for all $i = 1, ...n$ where $n = |x \backslash x_L| - 1$. Similarly, let $y_L = \{c \in y : a \succ c\}$, $y \backslash y_L = \{a, c_1, ...c_m\}$ and let $y_i = y_L \cup \{c_1, ...c_j\}$ for all $j = 1, ...m$ where $m = |y \backslash y_L| - 1$. Claim 2 implies that $\alpha(a, \{a\} \cup x_L) = \alpha(a, \{a\} \cup y_L)$. By Claim 1 we have that $\alpha(a, \{a\} \cup x_L) = \alpha(a, \{a\} \cup x_1)$, and by induction $\alpha(a, \{a\} \cup x_i) = \alpha(a, \{a\} \cup x_{i+1})$ for all $i \leq n - 1$, where of course $x = \{a\} \cup x_n$, so that

$$\alpha(a, x) = \alpha(a, \{a\} \cup x_L) = \alpha(a, \{a\} \cup y_L) \tag{4}$$

A similar reasoning applied to $\alpha(a, \{a\} \cup y_L)$ yields

$$\alpha(a, y) = \alpha(a, \{a\} \cup y_L) = \alpha(a, \{a\} \cup x_L) \tag{5}$$

and then by (4) and (5) we conclude $\alpha(a, x) = \alpha(a, y)$. ■

Preference for flexibility. We can also relate the menu independence of the attention parameters to a version of a classical axiom of the menu choice literature, 'preference for flexibility', which states that the agent is better off when the menu expands (irrespective of whether the expansion is by means of top alternatives or not). When the attention parameters are menu independent, it is easily verified (by a calculation analogous to the proof of necessity in theorem 2) that for any x and $a \notin x$, we have $u(x \cup \{a\}) > u(x)$. Therefore the following axiom is necessary in the vNM attention representation with menu independent attention parameters:

A6 (Strict preference for flexibility): $y \subset x \Rightarrow x \succ y$.

Since it is also true that A6 directly implies A5 as a special case, we conclude that it can replace it in the characterisation of theorem 2.[9] Strict preference for flexibility is, for a vNM rational but inattentive agent, the preference counterpart of menu independence in attention:

[9]Necessity of A6 is straightforward. Rather than providing a full proof, we show the argument with an example. Consider sets $y = \{a\}$ and $x = \{a, b\}$ so that $u(y) = \alpha(a) u(a) + (1 - \alpha(a)) u(\varnothing)$. Then either $u(x) = \alpha(a) u(a) + (1 - \alpha(a)) \alpha(b) u(b) + (1 - \alpha(a))(1 - \alpha(b)) u(\varnothing)$ (if $a \hat{\succ} b$, so that $u(a) \geq u(b)$), or $u(x) = \alpha(b) u(b) + (1 - \alpha(b)) \alpha(a) u(a) + (1 - \alpha(a))(1 - \alpha(b)) u(\varnothing)$ (if $b \hat{\succ} a$ and $u(b) \geq u(a)$). Either way, $u(x) > u(y)$ so that $x \succ y$.

Corollary 1 *An imperfect attention preference satisfies A6 if and only if it has a vNM attention representation* $(\hat{\succ}, u, \alpha)$ *in which* α *is menu independent.*

On the one hand, this result highlights the distinction from other models. For example, if the agent were *thought averse* in the sense of Ortoleva [32], he would never strictly prefer larger menus. On the other hand, Corollary 1 also highlights how the same menu preference may have different interpretations according to the context. While the classical interpretation of preference for flexibility is in terms of uncertainty about future tastes, it now turns out to be also indicative of a specially disciplined form of inattention. This is an instance of how it is impossible to get to the 'right' representation - that is, to the true model of the cognitive process underlying choice - only on the basis of preferences: extraneous information of some kind is necessary.

A formula for the attention parameters. With menu independent attention for all $x \in \mathcal{X}$ and with $a \hat{\succ} b$ for all $b \in x \setminus \{a\}$:

$$\begin{aligned}
u(x) &= \sum_{b \in x} \prod_{c \in x : c \hat{\succ} b} (1-\alpha(c))\,\alpha(b)\,u(b) + \prod_{b \in x} (1-\alpha(b))\,u(\varnothing) \\
&= \alpha(a)\,u(a) + \\
&\quad + (1-\alpha(a)) \left(\sum_{b \in x \setminus \{a\}} \prod_{c \in x \setminus \{a\} : c \hat{\succ} b} (1-\alpha(c))\,\alpha(b)\,u(b) + \prod_{b \in x \setminus \{a\}} (1-\alpha(b))\,u(\varnothing) \right) \\
&= \alpha(a)\,u(a) + (1-\alpha(a))\,u(x \setminus \{a\})
\end{aligned}$$

and therefore

$$\alpha(a) = \frac{u(x) - u(x \setminus \{a\})}{u(a) - u(x \setminus \{a\})}.$$

This formula is interesting in two respects. First, it shows that in this case the $\alpha(a)$ are *uniquely defined*, since they are invariant to any positive affine transformation of u, and u (as a vNM utility) is unique precisely up to such transformations:. Observe that for any $a \in x$ there exists an x with $a \in x$ and $a \hat{\succ} b$ for all $b \in x \setminus \{a\}$, for example $x = \{a\}$.

Secondly, the formula provides an interpretation of the attention parameters in terms of utility. The attention paid to a measures the ratio between the incremental utility of *offering the agent the opportunity of choosing a* (from a menu in which a is best), and the incremental utility of *endowing* the agent with a instead.

5 A discussion on interpretation and related literature

5.1 Interpretation of $\succsim$

We return to the interpretation of the menu/alternatives ranking in our model. As we mentioned in the introduction, the interpretation of this ranking as held by the same agent who is supposed to make choices from menus is not straightforward. If the agent may neglect some alternatives in the menu, his ranking of menus will not express his genuine preference over them, but rather a preference that is biased by his incomplete powers of attention. On the other hand, if we assumed that somehow the agent perceives all alternatives at the stage of evaluating menus, it is not clear why the later stage of a choice from menu may be vitiated by the lack of consideration of some alternatives. For these reasons, we have assumed so far that $\succsim$ expresses either an objective normative ranking or the ranking of a second decision maker who, while not making choices from menus, is in a 'paternalistic' relation with the agent in the sense that he (1) is responsible for selecting a menu from which the agent would choose, and could force the agent to consume certain alternatives; (2) knows both the preferences and the cognitive abilities of the agent; and (3) internalises the agent's preferences over alternatives. Such a decision maker could be a policy maker, a doctor, or a guardian setting constraints on the agent's behaviour. For example, a regulator might regulate more or less stringently the markets for financial products, houses, health and so on. The agent might be forced to subscribe to a given pension or health plan, or be offered the opportunity to pick from plans available in the market.

It is, however, also possible to interpret $\succsim$ as being held by the same agent who chooses from menus. A first interpretation is that the inattentive agent looks at *past* choices from menus and evaluates their results. At some point after having made choices from menus, he becomes aware of the exact composition of the menus. He can thus make 'hindsight' statements of the type 'I've been better off choosing from menu x rather than from menu y', 'I ended up better off when having a than when I faced a choice from x', and so on.

A second possibility, this time ex-ante, that suggests itself is imperfect *memory*.

Some alternatives may escape the agent's grasp after he has contemplated the menu and before the moment of choice. In this interpretation, when evaluating menus he is aware of his memory imperfection. Endowed a means immediate choice and $a \in x$ means that a will be available for choice later on provided it hasn't escaped attention. This interpretation is analogous to the *random availability* assumption in Barberá and Grodal [2]. These authors study decision makers who rank menus, have von Neuman-Morgenstern utility functions over alternatives, and attach subjective probabilities to each subset of alternatives surviving to the stage when the choice *from* the menu has to be made. For instance, it may be that at the time when the choice from a menu x is implemented the stocks of the preferred alternatives have run out. Thus the problem facing the agent is that of ranking menus ex ante (so that the choice from the menu is actually not carried out) based on the expected utility calculated taking into account the survival probabilities. In our context, availability can be interpreted psychologically: it expresses the ability of the agent to hold in mind the necessary information about the content of the menu. In this interpretation, too, the distinction between a and $\{a\}$ arises quite naturally, as a must, unlike $\{a\}$, be available later, as the agent has been endowed with it (Barberá and Grodal [2] do not make such a distinction, and their framework is in the vein of Kreps [22])

A third interpretation, also ex-ante, is in terms of absent-mindedness or 'implementation errors' (with self-awareness). In choosing exits from motorways, you know that it is better to take a given exit, but you are aware that you may end up missing it and being forced to take an inferior second exit. In this case you might for instance have to compare situations such as 'driving and having to choose from {first motorway exit, second motorway exit}' (with the risk of choosing suboptimally) and 'getting someone to drive you to destination' (the endowed alternative). Absent-mindedness as a behavioral phenomenon has been examined by Piccione and Rubinstein [34] (in a game theoretic context) and the subsequent literature. Implementation errors in choice are studied by Mattson and Weibull [28] as a foundation for the logit model. Finally, the possibility of making choices over menus while contemplating the possibility of making mistakes at the time of choice from the menu is a possible interpretation of frameworks such as Ahn and Sarver [1] and Koida [20].[10] While in that literature the reaction to such mistakes is modeled as a preference for commitment in the sense of preference for a singleton menu, we admit the possibility of missing an alternative even in a singleton menu.

5.2 Related literature

Two papers consider menu-dependence in attention in a different way from the one modeled here: Echenique, Saito and Tserenjigmid [11]'s (EST) and Brady and Rehbeck [5] (BR).

EST's 'Perception Adjusted Luce Model' (PALM) shares with ours some of the formalism of the choice from menu stage, though with a very different interpretation. A PALM has two primitives, a utility function $u : X \cup \mathcal{X} \to \mathcal{R}$ and a weak order $\succsim_p$ (the *perception order*) encapsulating the order with which alternatives are perceived. $a \succ_p b$ signifies that a is perceived sooner than b. For any $x \in \mathcal{X}$, $u(x)$ denotes the probability of not choosing any alternative from x. It is similar to $u(\varnothing)$ in a vNM attention representation, with the important difference that in a PALM the utility of not choosing can vary across choice sets. Letting $x \backslash \succsim_p$ denote the indifference classes induced by $\succsim_p$ on x, in a PALM the probability $p_{\succsim_p}(a,x)$ that alternative a is chosen from x is

$$p_{\succsim_p}(a,x) = \lambda(a,x) \left(\prod_{\tau \in x \backslash \succsim_p : \tau \succ_p a} \left(1 - \sum_{b \in x : b \in \tau} \lambda(b,x) \right) \right)$$

where, for all $a \in x$:

$$\lambda(a,x) = \frac{u(a)}{\sum_{a' \in x} u(a') + u(x)}$$

That is, $\lambda(a,x)$ is the standard Luce choice probability adjusted for a menu dependent outside option (as captured by $u(x)$). The structure of a PALM and of our representation

$$p_{iap}(a,x) = \alpha(a,x) \prod_{b \in x : b \hat{\succ} a} (1 - \alpha(b,x))$$

[10] E.g. in Koida: "Namely, she prefers the restaurant that serves only chicken to the one that serves both chicken and fish, to avoid “mistakenly” choosing suboptimal alternatives ex post."

are formally similar, since when the perception ordering $\succ_p$ in a PALM is strict we have

$$p_{\succ_p}(a,x) = \lambda(a,x)\left(\prod_{b\in x: b\succ_p a}(1-\lambda(b,x))\right)$$

A PALM thus offers an alternative way of interpreting our own implied process of choice from menus. Our attention function is replaced by a specific functional form (Luce probabilities) that, depending on the values u takes on menus, can satisfy the Monotonicity in consideration condition: if $b \succ_p a$ for all $a \in x$, then $\lambda(a,x) \geq \lambda(a, x \cup \{b\})$ as long as $u(b) + u(x \cup \{b\}) \geq u(x)$, which is compatible with a variety of u functions. Monotonicity in consideration generalises for example the condition of regularity that EST use in the characterisation,[11] and for menus of two alternatives it is implied by regularity. However a vNM attention representation does not generalise a PALM, as there are admissible u functions that would imply a contradiction of the monotonicity condition in theorem 1.

BR strictly generalise Manzini and Mariotti [26] in a different way (and with a different methodology) from ours. In their model, the probability $P(C,A)$ of a consideration set C obtaining when the agent faces menu A is generated by a fixed probability distribution π over 2^X, by

$$P(C,A) = \frac{\pi(C)}{\sum_{B\subseteq A}\pi(B)}$$

The fact that π is fixed (menu-independent) disciplines the consideration probabilities in a way that is not required in a vNM attention representation. On the other hand, in this way BR capture both menu-dependence and correlation in the formation of the consideration set, while a vNM attention representation implies stochastic independence.

In these two papers as well as in ours the interpretation of the concept of inattention is in line with those of Caplin and Dean [6] (CD); Caplin, Dean and Martin [7] (CDM); Eliaz and Spiegler ([12], [13]) (ES); Manzini and Mariotti [26] (MM); Masatlioglu, Nakajma and Ozbay [27] (MNO); and Sovinsky Goeree [39] (SG). In all these papers inattention means failing to consider alternatives in a menu. While ES explore

[11]See EST for details.

in detail the consequences of imperfect attention in a strategic setting, MNO and MM give an abstract characterisation of consideration sets models based on a standard revealed preference method. They consider the agent's choices from a set of menus and state conditions under which the agent's choices could be interpreted as deriving from a certain type of imperfect attention (the same applies to EST and BR discussed above). MNO focus on deterministic choices and look at a special restriction of the dependence of attention on the menu. MM study stochastic choices and characterise the menu independent version of the choice procedure implied by theorem 1. CD instead innovate the revealed preference method by using an enriched set of non-standard data, and assuming that the analyst can observe provisional choices and contemplation times. CDM put in practice this methodology in an experimental setting, validating the CD search-satisficing model of choice. While the innovative techniques used by CD and CDM allow in a sense the consideration to be observed directly by the analyst, SG uses careful econometric techniques of a more standard kind to infer the existence of non-trivial consideration sets in the purchase of personal computers.

The 'rational inattention' literature started by Sims [38] takes a different view of inattention: the agent faces uncertainty about the true state of the world and must select, at a cost, an optimal signalling structure (a joint distribution on signals and states), on the basis of which ex-post choices from menus are made. The work by de Oliveira, Denti, Mihm and Ozbek [10], however, is in a broadly similar vein to ours, in that it relates inattention to preferences over menus. By developing the menu choice approach (see also Ergin and Sarver [14], who laid out a related formalism to axiomatise contemplation costs), this is the first work to provide a such a type of foundation for rational inattention. Like that of Ergin and Sarver [14], it exploits much subtler nuances of preferences over lotteries than we do in this paper.

Finally we should mention Machina [25], perhaps the first to argue for an explanation of stochastic behaviour by means of a deterministic preference over lotteries rather than a stochastic preference as in RUM. We note that this distinction may sometimes be a matter of appearance: the vNM attention model with menu-independent parameters illustrates this by using a deterministic preference while being formally part of the

RUM family (as shown in MM). Even our general model, while not a RUM, admits a preference based interpretation: $\alpha(a, x)$ is the probability of being 'in the mood' for a when the menu is x. Nevertheless, Machina's remains an important intuition.[12]

6 Failures of Regularity and the Attraction Effect

Many models of stochastic choice satisfy the property of Regularity, according to which the probability of choosing an alternative does not increase as the menu gets larger. Regularity is satisfied by the very general Random Utility Model (RUM)[13], and consequently by the version of our model with menu independent attention parameters (which, as shown in Manzini and Mariotti [26], is a particular case of RUM), as well as by Luce's [23] classical model, its Gumbel error and multinomial logit versions popular in econometrics,[14] and also by its recent generalisation by Gul, Natenzon and Pesendorfer [17].[15] However, some prominent anomalies contradict Regularity, in particular the *attraction effect* and the *compromise* effect. These anomalies are both reasonably well documented in experimental findings and introspectively plausible. We consider it as an important feature of a stochastic theory of choice that it offers a pathway to representing such anomalies.

Marketers use a number of strategies to manipulate the attractiveness or otherwise of alternatives. The *attraction effect* (also known as the 'asymmetric dominance' effect, see Huber and Puto [18], [19]) refers to the fact that the choice frequency of a target alternative t increases when a new decoy alternative d is introduced in a menu, with the property that the d is markedly worse than the target t, while incomparable to a third ('other') alternative, o. This ranking is generally induced by presenting alternatives

[12] See Fudenberg, Iijima and Strzalecki [15] for a recent and very general development of Machina's approach.

[13] In a RUM model (Block and Marschack [4]) the agent picks the top element of a ranking extracted at random according to a known probability distribution.

[14] See Holman and Marley (as attributed in [24]), Mc Fadden [29] and Yellot [40] for the Gumbel error interpretation of the Luce/Logit model.

[15] EST and BR discussed in the previous section are notable instances of models that can explain violations of Regularity.

as described in two desirable attributes/dimensions: while the ranking between t and o in one dimension is reversed in the other, d is Pareto dominated by t but Pareto incomparable to o. The *compromise effect* instead refers to the introduction of a different type of decoy, which has the highest degree of one attribute and the lowest of another in such a way that t is now 'middle ranking'.

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

All these effects are easy to accommodate in our setup. From the formula $p(a, A) = \alpha(a, A) \prod_{b \in A : b \hat{\succ} a} (1 - \alpha(b, A))$ it follows immediately that $p(a, A)$ increases with $\alpha(a, A)$ and decreases with $\alpha(b, A)$ for any $b \in A$ such that $b \hat{\succ} a$. Now consider adding a decoy d to a menu. Then, as long as $a \hat{\succ} d$, we have

$$\begin{aligned} p(a, A \cup \{d\}) &= \alpha(a, A \cup \{d\}) \prod_{b \in A \cup \{d\} : b \hat{\succ} a} (1 - \alpha(b, A \cup \{d\})) \\ &= \alpha(a, A \cup \{d\}) \prod_{b \in A : b \hat{\succ} a} (1 - \alpha(b, A \cup \{d\})) \end{aligned}$$

As long as the introduction of d increases the attention parameter of alternative a without affecting the attention parameters of the other alternatives (or at least without increasing them too much), we have $p(a, A \cup \{d\}) > p(a, A)$, while $p(b, A \cup \{d\}) < p(b, A)$ for all b such that $a \hat{\succ} b$. Even if $b \hat{\succ} a$ we can still have a decrease in the probability that b is chosen after the introduction of the decoy alternative, provided that the attention paid to it is decreased by this event. Indeed, adding a dominated alternative imposes in our model no constraints on the attention parameters of the existing alternatives. The advantage of this way of modelling the phenomenon is that the previous reasoning holds regardless of the type of decoy that is introduced, whether it is a phantom alternative, or one that induces compromise, or a symmetrically dominated alternative, and so on. What conforms to the intentions of the manipulator and accords with the structure of our model is that the target alternative is made more appealing not by improvements to it, but simply by *framing* - what we call 'attention' is whatever it is that the manipulator/marketer strives to influence. For example, in the compromise effect it is hard to tell whether what is *behaviourally* a compromise-seeking attitude really reflects a compromise seeking *psychology*. Indeed, Mochon and Frederick [30] find experimentally that an *order* effect could be a more plausible explanation for the 'compromise effect': the alternative presented as second seems more salient in the choice between three items, regardless of its attributes. We would go even further and regard 'attention' as a catch-all concept that gathers the factors, psychological or of other nature, that affect choice through consideration and separately from preferences.

7 Concluding remarks

The most appealing aspect of the methodology of this paper is that it yields a specific representation of imperfect attention using only a very broad, and we hope uncontroversial, definition of the welfare consequences of imperfect attention (our axiom A4), without committing to any particular assumption on what causes it. Thus, our implied choice model can be seen as a 'reduced form' of several more detailed stories that may lie behind the agent's failure to consider all alternatives. These stories may range from search theories to models of brand loyalty. In addition to cognitive reasons, there may also be ethical or ideological reasons why some available options are deliberately not considered.

Because we have assumed vNM rationality, one can see our representation as capturing deviations from rational behaviour that can be imputed *exclusively* to imperfect attention and not to other cognitive imperfections. For example, since we have shown that vNM rationality implies a lack of correlation in consideration, we can interpret any evidence of correlated consideration as a separate departure from full rationality, distinct form imperfect attention *per se*.[16]

It remains to be explored whether our approach may prove useful, as we hope, to model aspects of bounded rationality different from imperfect attention. A hint of a possible development in this direction is as follows. Each alternative may present itself in two or more 'modes'. So the full description of an alternative a is (a, m), spec-

[16]See Brady and Rehbeck [5] for intuitively plausible examples of correlations in consideration.

ifying that a comes in mode m.[17] In this paper $m \in \{e, \bar{e}\}$, where e ='endowed' and $\bar{e}$ ='not endowed'. Specific restrictions on preferences over such enriched alternatives and menus of them may capture different interpretations and psychological attitudes. Examples of what a mode might indicate are: the procedure with which a is made available; whether the acts of choice and of consumption of a are simultaneous or not; whether a is the status quo; whether a is a social norm; whether a is chosen privately or in front of an audience.

References

[1] Ahn D. and T. Sarver (2013) "Preference for Flexibility and Random Choice", *Econometrica* 81: 341–361.

[2] Barberá, S. and B. Grodal (2011) "Preference for Flexibility and the Opportunities of Choice", *Journal of Mathematical Economics* 47: 272-278.

[3] Beshears, J., Choi, J. J., Laibson, D., & Madrian, B. C. (2009) "The importance of default options for retirement saving outcomes: Evidence from the United States", in *Social security policy in a changing environment* (pp. 167-195), University of Chicago Press.

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

[5] Brady, R. and J. Rehbeck (2014) "Menu-Dependent Stochastic Consideration", working paper, University of California San Diego.

[6] Caplin, A. and M. Dean "Search, Choice and Revealed Preference", *Theoretical Economics* 6: 19-48.

[17] We thank Ariel Rubinstein for suggesting this generalisation.

[7] Caplin, A., M. Dean and D. Martin (2011) “Search and Satisficing”, *American Economic Review* 101(7): 2899-2922.

[8] Clark, S.A. (1995) “Indecisive choice theory”, *Mathematical Social Sciences* 30: 155-170.

[9] Corbin, R. and A.A.J Marley (1974) “Random Utility Models with Equality: An Apparent, but Not Actual, Generalization of Random Utility Models", *Journal of Mathematical Psychology*, 11: 274-293

[10] de Oliveira, H., T. Denti, M. Mihm and M.K. Ozbek (2013) "Rationally Inattentive Preferences", working paper, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2274286.

[11] Echenique, F., K. Saito and Tserenjigmid (2013) “The Perception-Adjusted Luce Model”, mimeo, California Institute of Techonology.

[12] Eliaz, K. and R. Spiegler (2011) “Consideration Sets and Competitive Marketing”, *Review of Economic Studies* 78: 235-262.

[13] Eliaz, K. and R. Spiegler (2011) “On the strategic use of attention grabbers”, *Theoretical Economics* 6: 127–155.

[14] Ergin H. and T. Sarver (2010) "A Unique Costly Contemplation Representation", *Econometrica* 78 (4): 1285-1339.

[15] Fudenberg, D., R. Iijima and T. Strzalecki (2014) “Stochastic Choice and Revealed Perturbed Utility”, working paper, Harvard University.

[16] Gerasimou, G. (2014) “Asymmetric Dominance, Deferral and Status Quo Bias in a Behavioral Model of Choice”, mimeo, University of St Andrews.

[17] Gul, F., P. Natenzon and W. Pesendorfer (2014) “Random Choice as Behavioral Optimization”, *Econometrica*, forthcoming.

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

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

[20] Koida, N. (2010) “Anticipated Stochastic Choice”, mimeo, Iwate Prefectural University.

[21] Krajbich, I. and A. Rangel (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.

[22] Kreps, D. M. (1979) "A Representation Theorem for ‘Preference for Flexibility’", *Econometrica*, 47: 565-577.

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

[24] Luce, R.D. and Suppes, P. (1965) “Preference, Utility, and Subjective Probability”. In R.D. Luce, R.R. Bush, and E. Galanter (Eds.), *Handbook of Mathematical Psychology* (Vol. III, pp. 235–06). New York: Wiley.

[25] Machina, M. (1985) “Stochastic Choice Functions Generated by Deterministic Preferences Over Lotteries”, *Economic Journal 95*: 575-94.

[26] Manzini, P. and M. Mariotti (2014) “Stochastic Choice and Consideration Sets", *Econometrica* 82(3): 1153-76.

[27] Masatlioglu, Y., D. Nakajima and E. Ozbay (2012): “Revealed Attention”, *American Economic Review* 102(5): 2183–2205.

[28] Mattsson, L.-G. and J. W. Weibull (2002) “Probabilistic choice and procedurally bounded rationality”, *Games and Economic Behavior*, 41(1): 61-78.

[29] McFadden, D.L. (1974) “Conditional Logit Analysis of Qualitative Choice Behavior”, in P. Zarembka (ed.), Frontiers in Econometrics, 105-142, Academic Press: New York.

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

[31] van Nierop, E., B. Bronnenberg, R. Paap, M. Wedel & P.H. Franses (2010) “Retrieving Unobserved Consideration Sets from Household Panel Data", *Journal of Marketing Research* XLVII: 63–74.

[32] Ortoleva, P. (2013) “The Price of Flexibility: Towards a Theory of Thinking Aversion", *Journal of Economic Theory* 148: 903–934.

[33] Payro, F. and L. Ülkü (2014) “Similarity-Based Mistakes in Choice”, mimeo, ITAM.

[34] Piccione M. and A. Rubinstein (1997) “On the Interpretation of Decision Problems with Imperfect Recall", *Games and Economic Behavior*, 20: 3-24.

[35] Regenwetter, M., Marley, A. A. J. and Joe, H. (1998) “Random Utility Threshold Models of Subset Choice", Australian Journal of Psychology, 50: 175–185

[36] Reutskaya, E., R. Nagel, C. F. Camerer, and A. Rangel (2011) “Search Dynamics in Consumer Choice under Time Pressure: An Eye-Tracking Study”, *The American Economic Review*, 101:900-926.

[37] Simon, H. A. (1971), "Designing Organizations for an Information-Rich World", in Martin Greenberger, Computers, Communication, and the Public Interest, Baltimore, MD: The Johns Hopkins Press,

[38] Sims, C.A. (2003) "Implications of rational inattention", *Journal of Monetary Economics*, 50: 665–690.

[39] Sovinsky Goeree, M. (2008) “Limited Information and Advertising in the US Personal Computer Industry”, *Econometrica* 76: 1017-1074.

[40] Yellot, J.I., JR. (1977) "The relationship between Luce's choice axiom, Thurstone's theory of comparative judgment, and the double exponential distribution", *Journal of Mathematical Psychology*, 15: 109-144.

A Appendix

A.1 Proof of theorem 1

Necessity. Suppose $\succsim$ on $\Delta(X \cup \mathcal{X})$ has a vNM attention representation $(\hat{\succ}, u, \alpha)$ in which α satisfies the monotonicity condition. We show that it satisfies A0-A4. Since $\alpha(a, x) \in (0,1)$ for all $x \in \mathcal{X}$ and $a \in x$, it follows

$$u(x) = \sum_{a \in x} \prod_{b \in x: b \hat{\succ} a} (1 - \alpha(b, x)) \alpha(a, x) u(a) + \prod_{a \in x} (1 - \alpha(a, x)) u(\varnothing) > u(\varnothing)$$

so that A0 holds. The necessity of A1-A3 is standard and thus omitted. Finally, let $a \succsim b$ for all $b \in x$. Then

$$u(x) = \sum_{c \in x} \prod_{c \in x: c \hat{\succ} b} (1 - \alpha(c, x)) \alpha(b, x) u(b) < u(a)$$

since the left hand side is a convex combination of values which do not exceed $u(a)$ and the sum of the weights on $\max_{b \in x} u(b) \leq u(a)$ is strictly less than unity (given that $\prod_{c \in x} (1 - \alpha(c, x))$, the weight on $u(\varnothing)$, is strictly positive), so that A4 holds.

For sufficiency, let $\succ'$ be an arbitrary linear order of X and define $\hat{\succ}$ lexicographically as follows: $a \hat{\succ} b$ iff $a \succ b$ or $a \sim b$ and $a \succ' b$. Denote menus by numbering the alternatives in them according to $\hat{\succ}$, as $x = \{a_1, ..., a_K\}$ with $a_i \hat{\succ} a_{i+1}$ for all $i = 1, ..., K-1$. We will show that for all $x \in \mathcal{X}$ there exist numbers $\alpha(a_1, x), ..., \alpha(a_K, x) \in (0,1)$ such that

$$\begin{aligned} x \quad \sim \quad & \alpha(a_1, x) a_1 + (1 - \alpha(a_1, x)) \alpha(a_2, x) a_2 + ... + \prod_{i=1}^{K-1} (1 - \alpha(a_i, x)) \alpha(a_K, x) a_K \\ & + \prod_{i=1}^{K} (1 - \alpha(a_i, x)) \varnothing \end{aligned} \tag{6}$$

Then by the vNM theorem and A1-A3 there exists a vNM utility u on $\Delta(X \cup \mathcal{X})$ representing $\succsim$ such that

$$\begin{aligned} u(x) &= u\left(\alpha_1(a_1,x)a_1+...+\prod_{i=1}^{K-1}(1-\alpha(a_i,x))\alpha(a_K,x)a_K+\prod_{i=1}^{K}(1-\alpha(a_i,x))\varnothing\right) \\ &= \sum_{i=1}^{K}\prod_{j=1}^{i-1}(1-\alpha(a_j,x))\alpha(a_i,x)u(a_i)+\prod_{i=1}^{K}(1-\alpha(a_i,x))u(\varnothing) \end{aligned}$$

(we use the convention that $\prod_{i=m}^{n} f(i) = 1$ and $\sum_{i=m}^{n} f(i) = 0$ for all functions $f : \mathbb{N} \to (0,1)$ whenever $m > n$).

If x consists of only one element, then by A0 and A4 $a_1 \succ \{a_1\} \succ \varnothing$. By the vNM axioms and textbook arguments,

$$\{a_1\} \sim \alpha(a_1,\{a_1\})a_1+(1-\alpha(a_1,\{a_1\}))\varnothing$$

for some unique $\alpha(a_1,\{a_1\}) \in (0,1)$ and so the result holds. Suppose then that x consists of two or more elements. There are three cases to consider. In all cases we argue by induction on the cardinality of the menu, supposing that the assertion is true for all menus with fewer than K elements and letting $x = \{a_1,...,a_K\}$ (where recall that $a_i \hat{\succ} a_{i+1}$ for all $i = 1,...,K-1$).

Case 1: $x \succ x\setminus\{a_1\}$. Then $a_1 \succ x \succ x\setminus\{a_1\}$ by A4, and by the vNM axioms there exists a unique $\alpha(a_1,x) \in (0,1)$ such that

$$x \sim \alpha(a_1,x)a_1+(1-\alpha(a_1,x))x\setminus\{a_1\}.$$

By the inductive hypothesis, there exist $\alpha(a_2,x\setminus\{a_1\}),...,\alpha(a_K,x\setminus\{a_1\}) \in (0,1)$ such that

$$\begin{aligned} x\setminus\{a_1\} &\sim \alpha(a_2,x\setminus\{a_1\})a_2+...+\prod_{i=2}^{K-1}(1-\alpha(a_i,x\setminus\{a_1\}))\alpha(a_K,x\setminus\{a_1\})a_K \\ &\quad +\prod_{i=2}^{K}(1-\alpha(a_i,x\setminus\{a_1\}))\varnothing \end{aligned}$$

and so by Independence the desired conclusion follows by setting $\alpha(a_i,x) = \alpha(a_i,x\setminus\{a_1\})$ for all $i = 2,...,K$. Note that in this case monotonicity in consideration is satisfied with equality.

Case 2: $x\backslash\{a_1\} \succ x$. Together with A0 this implies $x\backslash\{a_1\} \succ x \succ \varnothing$, and by the vNM axioms there exists a unique $\beta \in (0,1)$ with

$$x \sim \beta x\backslash\{a_1\} + (1-\beta)\varnothing. \tag{7}$$

Moreover by A4 $a_1 \succ x\backslash\{a_1\} \succ x$, so that there exists a unique $\alpha \in (0,1)$ such that

$$x\backslash\{a_1\} \sim \alpha a_1 + (1-\alpha)x.$$

Having defined α and β in this way, we claim that equation (6) (with the stated properties on the coefficients) holds by setting the coefficients recursively as follows:

$$\begin{aligned} \alpha(a_1, x) &= \alpha\beta \\ \alpha(a_k, x) &= \gamma\alpha(a_k, x\backslash\{a_1\}) \text{ with} \\ \gamma &= \frac{(1-\alpha)\beta^2\prod_{i=2}^{k-1}(1-\alpha(a_i, x\backslash\{a_1\}))}{1-\alpha\beta-(1-\alpha)\beta^2\left(\sum_{i=2}^{k-1}\alpha(a_i, x\backslash\{a_1\})\prod_{j=2}^{i-1}\left(1-\alpha\left(a_j, x\backslash\{a_1\}\right)\right)\right)} \end{aligned} \tag{8}$$

for all $k = 2, ...K$.

Step 1: the $\alpha(a_k, x)$ defined in equation (8) satisfy expression (6). By Independence applied to formula (7), given the definition of α, it must be:

$$x \sim \beta\alpha a_1 + \beta(1-\alpha)x + (1-\beta)\varnothing. \tag{9}$$

In turn, using the expression for x from condition (7) and Independence in expression (9) we have:

$$x \sim \alpha\beta a_1 + (1-\alpha)\beta^2 x\backslash\{a_1\} + (1-\beta)(1+\beta(1-\alpha))\varnothing$$

so that by the inductive hypothesis and Independence:

$$\begin{aligned} x \sim\ & \alpha\beta a_1 + (1-\alpha)\beta^2\left(\alpha(a_2, x\backslash\{a_1\})a_2 + ... + \prod_{i=2}^{K-1}(1-\alpha(a_i, x\backslash\{a_1\}))\alpha(a_K, x\backslash\{a_1\})a_K\right) \\ & + (1-\beta)(1+\beta(1-\alpha))\varnothing + (1-\alpha)\beta^2\prod_{i=2}^{K}(1-\alpha(a_i, x\backslash\{a_1\}))\varnothing \end{aligned} \tag{10}$$

Using (8) for any $k \geq 2$, we have

$$1-\alpha\left(a_k,x\right)=1-\frac{(1-\alpha)\beta^2\prod_{i=2}^{k-1}(1-\alpha(a_i,x\backslash\{a_1\}))}{1-\alpha\beta-(1-\alpha)\beta^2\left(\sum_{i=2}^{k-1}\alpha(a_i,x\backslash\{a_1\})\prod_{j=2}^{i-1}\left(1-\alpha\left(a_j,x\backslash\{a_1\}\right)\right)\right)}\alpha\left(a_k,x\backslash\left\{a_1\right\}\right)=$$

$$=\frac{1-\alpha\beta-(1-\alpha)\beta^2\left(\sum_{i=2}^{k-1}\alpha(a_i,x\backslash\{a_1\})\prod_{j=2}^{i-1}\left(1-\alpha\left(a_j,x\backslash\{a_1\}\right)\right)\right)-(1-\alpha)\beta^2\left(\prod_{i=2}^{k-1}(1-\alpha(a_i,x\backslash\{a_1\}))\right)\alpha(a_k,x\backslash\{a_1\})}{1-\alpha\beta-(1-\alpha)\beta^2\left(\sum_{i=2}^{k-1}\alpha(a_i,x\backslash\{a_1\})\prod_{j=2}^{i-1}\left(1-\alpha\left(a_j,x\backslash\{a_1\}\right)\right)\right)}=$$

$$=\frac{1-\alpha\beta-(1-\alpha)\beta^2\left(\sum_{i=2}^{k-1}\alpha(a_i,x\backslash\{a_1\})\prod_{j=2}^{i-1}\left(1-\alpha\left(a_j,x\backslash\{a_1\}\right)\right)+\alpha(a_k,x\backslash\{a_1\})\prod_{i=2}^{k-1}(1-\alpha(a_i,x\backslash\{a_1\}))\right)}{1-\alpha\beta-(1-\alpha)\beta^2\left(\sum_{i=2}^{k-1}\alpha(a_i,x\backslash\{a_1\})\prod_{j=2}^{i-1}\left(1-\alpha\left(a_j,x\backslash\{a_1\}\right)\right)\right)}=$$

$$=\frac{1-\alpha\beta-(1-\alpha)\beta^2\sum_{i=2}^{k}\alpha(a_i,x\backslash\{a_1\})\prod_{j=2}^{i-1}\left(1-\alpha\left(a_j,x\backslash\{a_1\}\right)\right)}{1-\alpha\beta-(1-\alpha)\beta^2\left(\sum_{i=2}^{k-1}\alpha(a_i,x\backslash\{a_1\})\prod_{j=2}^{i-1}\left(1-\alpha\left(a_j,x\backslash\{a_1\}\right)\right)\right)}$$

so that the numerator of $1-\alpha\left(a_k,x\right)$ is equal to the denominators of $1-\alpha\left(a_{k+1},x\right)$ and $\alpha\left(a_{k+1},x\right)$. Consequently, the product $\alpha\left(a_k,x\right)\prod_{i=1}^{k-1}\left(1-\alpha\left(a_i,x\right)\right)$ is a telescoping product, yielding

$$\alpha\left(a_k,x\right)\prod_{i=1}^{k-1}\left(1-\alpha\left(a_i,x\right)\right)=\left(1-\alpha\right)\beta^2\prod_{i=2}^{k-1}\left(1-\alpha\left(a_i,x\backslash\left\{a_1\right\}\right)\right)$$

which is precisely the coefficient of a_k in the lottery on the right hand side of (10). Note (see Step 3 below) that in this case monotonicity in consideration is satisfied with inequality.

Step 2: $\alpha\left(a_k,x\right)>0$ for all $k=2,...K$. It is obvious that $\alpha\left(a_1,x\right)>0$ given the admissible values of α and β. For $k=2,...K$, note that the numerator is positive, and that (given the admissible values of α and β) we have $0<\left(1-\alpha\right)\beta^2<\left(1-\alpha\beta\right)$. So the denominator is positive given that

$$\sum_{i=2}^{k-1}\alpha\left(a_i,x\backslash\left\{a_1\right\}\right)\prod_{j=2}^{i-1}\left(1-\alpha\left(a_j,x\backslash\left\{a_1\right\}\right)\right)<1$$

To prove this last inequality, observe that (keeping an eye on the summation and product indexes):

$$
\begin{aligned}
& \sum_{i=2}^{k-1} \alpha\left(a_i, x \backslash \left\{a_1\right\}\right) \prod_{j=2}^{i-1}\left(1-\alpha\left(a_j, x \backslash \left\{a_1\right\}\right)\right)<1 \\
\Leftrightarrow\ & \alpha\left(a_2, x \backslash \left\{a_1\right\}\right)+\sum_{i=3}^{k-1} \alpha\left(a_i, x \backslash \left\{a_1\right\}\right) \prod_{j=2}^{i-1}\left(1-\alpha\left(a_j, x \backslash \left\{a_1\right\}\right)\right)<1 \\
\Leftrightarrow\ & \sum_{i=3}^{k-1} \alpha\left(a_i, x \backslash \left\{a_1\right\}\right) \prod_{j=2}^{i-1}\left(1-\alpha\left(a_j, x \backslash \left\{a_1\right\}\right)\right)<1-\alpha\left(a_2, x \backslash \left\{a_1\right\}\right) \\
\Leftrightarrow\ & \sum_{i=3}^{k-1} \alpha\left(a_i, x \backslash \left\{a_1\right\}\right) \prod_{j=3}^{i-1}\left(1-\alpha\left(a_j, x \backslash \left\{a_1\right\}\right)\right)<1 \\
\Leftrightarrow\ & \alpha\left(a_3, x \backslash \left\{a_1\right\}\right)+\sum_{i=4}^{k-1} \alpha\left(a_i, x \backslash \left\{a_1\right\}\right) \prod_{j=3}^{i-1}\left(1-\alpha\left(a_j, x \backslash \left\{a_1\right\}\right)\right)<1 \\
\Leftrightarrow\ & \sum_{i=4}^{k-1} \alpha\left(a_i, x \backslash \left\{a_1\right\}\right) \prod_{j=3}^{i-1}\left(1-\alpha\left(a_j, x \backslash \left\{a_1\right\}\right)\right)<1-\alpha\left(a_3, x \backslash \left\{a_1\right\}\right) \\
\Leftrightarrow\ & \sum_{i=4}^{k-1} \alpha\left(a_i, x \backslash \left\{a_1\right\}\right) \prod_{j=4}^{i-1}\left(1-\alpha\left(a_j, x \backslash \left\{a_1\right\}\right)\right)<1 \\
\Leftrightarrow\ & \ldots \\
\Leftrightarrow\ & \alpha\left(a_{k-1}, x \backslash \left\{a_1\right\}\right)\left(1-\alpha\left(a_{k-1}, x \backslash \left\{a_1\right\}\right)\right)<1
\end{aligned}
$$

where the last inequality holds true by the inductive hypothesis, since $|x \backslash \{a_1\}| = K-1$).

Step 3: $\alpha\left(a_k, x\right)<1$. It is obvious that $\alpha\left(a_1, x\right)<1$ given the admissible values of α and β. For the other coefficients we show that

$$
\frac{\alpha\left(a_k, x\right)}{\alpha\left(a_k, x \backslash \left\{a_1\right\}\right)}<1 \text{ for all } k \leq K, \tag{11}
$$

which implies the result (since $\alpha\left(a_k, x \backslash \left\{a_1\right\}\right)<1$ by the inductive hypothesis on the cardinality of x). We proceed by induction on k (given K). If $k=2$, then from the second line in (8) we have

$$
\frac{\alpha\left(a_2, x\right)}{\alpha\left(a_2, x \backslash \left\{a_1\right\}\right)}=\frac{(1-\alpha) \beta^2}{1-\alpha \beta}<1.
$$

Now suppose that $\frac{\alpha\left(a_k, x\right)}{\alpha\left(a_k, x \backslash \left\{a_1\right\}\right)}<1$ for all k for which $2 \leq k \leq k^{\prime}-1<K$, and consider

$k = k'$. Then

$$\frac{\alpha\left(a_{k'},x\right)}{\alpha\left(a_{k'},x\backslash\{a_1\}\right)} = \frac{(1-\alpha)\beta^2\prod_{i=2}^{k'-1}(1-\alpha(a_i,x\backslash\{a_1\}))}{1-\alpha\beta-(1-\alpha)\beta^2\left(\sum_{i=2}^{k'-1}\alpha(a_i,x\backslash\{a_1\})\prod_{j=2}^{i-1}\left(1-\alpha\left(a_j,x\backslash\{a_1\}\right)\right)\right)} < 1$$

$$\Leftrightarrow (1-\alpha)\,\beta^2\prod_{i=2}^{k'-1}\left(1-\alpha\left(a_i,x\backslash\{a_1\}\right)\right)$$

$$< 1-\alpha\beta-(1-\alpha)\,\beta^2\left(\sum_{i=2}^{k'-1}\alpha\left(a_i,x\backslash\{a_1\}\right)\prod_{j=2}^{i-1}\left(1-\alpha\left(a_j,x\backslash\{a_1\}\right)\right)\right)$$

$$\Leftrightarrow (1-\alpha)\,\beta^2\left(1-\alpha\left(a_{k'-1},x\backslash\{a_1\}\right)\right)\prod_{i=2}^{k'-2}\left(1-\alpha\left(a_i,x\backslash\{a_1\}\right)\right)$$

$$< 1-\alpha\beta-(1-\alpha)\,\beta^2\left(\begin{array}{c}\sum_{i=2}^{k'-2}\alpha\left(a_i,x\backslash\{a_1\}\right)\prod_{j=2}^{i-1}\left(1-\alpha\left(a_j,x\backslash\{a_1\}\right)\right)\\ +\alpha\left(a_{k'-1},x\backslash\{a_1\}\right)\prod_{j=2}^{i-1}\left(1-\alpha\left(a_j,x\backslash\{a_1\}\right)\right)\end{array}\right)$$

$$\Leftrightarrow (1-\alpha)\,\beta^2\prod_{i=2}^{k'-2}\left(1-\alpha\left(a_i,x\backslash\{a_1\}\right)\right)$$

$$< 1-\alpha\beta-(1-\alpha)\,\beta^2\left(\sum_{i=2}^{k'-2}\alpha\left(a_i,x\backslash\{a_1\}\right)\prod_{j=2}^{i-1}\left(1-\alpha\left(a_j,x\backslash\{a_1\}\right)\right)\right)$$

$$\Leftrightarrow \frac{(1-\alpha)\beta^2\prod_{i=2}^{k'-2}(1-\alpha(a_i,x\backslash\{a_1\}))}{1-\alpha\beta-(1-\alpha)\beta^2\left(\sum_{i=2}^{k'-2}\alpha(a_i,x\backslash\{a_1\})\prod_{j=2}^{i-1}\left(1-\alpha\left(a_j,x\backslash\{a_1\}\right)\right)\right)} = \frac{\alpha\left(a_{k'-1},x\right)}{\alpha\left(a_{k'-1},x\backslash\{a_1\}\right)} < 1$$

where $\frac{\alpha\left(a_{k'-1},x\right)}{\alpha\left(a_{k'-1},x\backslash\{a_1\}\right)} < 1$ holds by the inductive hypothesis on k. Thus, condition (11) holds.

Case 3: $x\backslash\{a_1\} \sim x$. Then $a_1 \succ x \succ \varnothing$ and A2 imply that there exists a unique $\alpha \in (0,1)$ with

$$x \sim \alpha a_1 + (1-\alpha)\,\varnothing.$$

Applying Independence repeatedly, the above and $x\backslash\{a_1\} \sim x$ imply that, for all $\beta \in [0,1]$,

$$x \sim \beta\left(\alpha a_1 + (1-\alpha)\,\varnothing\right) + (1-\beta)\,x\backslash\{a_1\}$$

so that by the inductive hypothesis

$$\begin{aligned} x \;\sim\;& \alpha\beta a_1 + (1-\beta)\left(\alpha\left(a_2,x\backslash\{a_1\}\right)a_2 + \ldots + \prod_{i=2}^{K-1}\left(1-\alpha\left(a_i,x\backslash\{a_1\}\right)\right)\alpha\left(a_K,x\backslash\{a_1\}\right)a_K\right)\\ &+\beta\,(1-\alpha)\,\varnothing + (1-\beta)\prod_{i=2}^{K}\left(1-\alpha\left(a_i,x\backslash\{a_1\}\right)\right)\varnothing \end{aligned} \qquad (12)$$

Fix β so that $\beta \in (0,1)$. Then, similarly to case 2, with α and β so defined condition (6) (with the stated properties on the coefficients) holds by setting recursively

$$\begin{aligned} \alpha(a_1, x) &= \alpha\beta \qquad (13) \\ \alpha(a_k, x) &= \frac{(1-\beta)\prod_{i=2}^{k-1}(1-\alpha(a_i, x\setminus\{a_1\}))}{1-\alpha\beta-(1-\beta)\left(\sum_{i=2}^{k-1}\alpha(a_i, x\setminus\{a_1\})\prod_{j=2}^{i-1}(1-\alpha(a_j, x\setminus\{a_1\}))\right)}\alpha(a_k, x\setminus\{a_1\}) \end{aligned}$$

A straightforward adaptation of Step 2 and Step 3 in the proof of case 2 shows that $\alpha(a_k, x) \in (0,1)$ for all $k = 1, ...K$. To see that (13) retrieves the coefficients in (6) correctly, again a straightforward adaptation of the proof of Step 1 in case 2 shows that the product $\alpha(a_k, x)\prod_{i=1}^{k-1}(1-\alpha(a_i, x))$ is a telescoping product, yielding

$$\alpha(a_k, x)\prod_{i=1}^{k-1}(1-\alpha(a_i, x)) = (1-\beta)\prod_{i=2}^{k-1}(1-\alpha(a_i, x\setminus\{a_1\}))$$

namely the coefficient of a_k in the lottery on the right hand side of (6). ■

A.2 Example 1

Let $X = \{a, b\}$, suppose that preferences satisfy A0-A4 and are such that $a \succ b \succ \{a, b\} \succ \{a\} \succ \{b\}$, and suppose that u represents preferences with

$X \cup \mathcal{X}$	a	b	$\{a,b\}$	$\{a\}$	$\{b\}$	$\varnothing$
$u(\cdot)$	U	pU	qpU	rU	sU	0

where $U > 0$, $p, q, r, s \in (0,1)$ and $s < r < qp$. Since $\alpha(a, \{a\})U = u(\{a\}) = rU$ and $\alpha(b, \{b\})pU = u(\{b\}) = sU$ we determine the parameters

$$\begin{aligned} \alpha(a, \{a\}) &= r \\ \alpha(b, \{b\}) &= \frac{s}{p} \end{aligned}$$

The other constraint is

$$\begin{aligned} & \alpha(a, X)U + (1-\alpha(a, X))\alpha(b, X)pU = u(X) = qpU \\ \Leftrightarrow \quad & \alpha(b, X) = \frac{qp-\alpha(a, X)}{(1-\alpha(a, X))p} \end{aligned}$$

Since $\alpha(b, X) \in (0,1)$, it must be that

$$\frac{qp - \alpha(a, X)}{(1 - \alpha(a, X))\, p} \in (0,1) \Leftrightarrow \alpha(a, X) < qp$$

(observing that the numerator is less than the denominator if and only if $(q-1)\, p < (1-p)\, \alpha(a, X)$, which holds true always). Moreover, since the monotonicity condition on α imposes that $\alpha(b, X) \leq \alpha(b, \{b\})$, it must also be

$$\frac{qp - \alpha(a, X)}{(1 - \alpha(a, X))\, p} \leq \frac{s}{p} \Leftrightarrow \alpha(a, X) \geq \frac{qp - s}{1 - s}$$

In short, then, we have the restriction

$$\alpha(a, X) \in \left[\frac{qp - s}{1 - s}, qp\right)$$

A.3 An example of the effect of $\hat{\succ}$ on the attention function

Let $X = \{a, b\}$, $a \sim b \succ \{a, b\} \succ \{a\} \sim \{b\}$, with u representing these preferences and defined as

$X \cup \mathcal{X}$	a	b	$\{a, b\}$	$\{a\}$	$\{b\}$	$\varnothing$
$u(\cdot)$	U	U	pU	pqU	pqU	0

with $U > 0$, $p, q \in (0,1)$.

Since $\alpha(a, \{a\})\, U = u(\{a\}) = u(\{b\}) = \alpha(b, \{b\})\, U = pqU$ we determine the parameters $\alpha(a, \{a\}) = pq = \alpha(b, \{b\})$. The other constraint is

$$\begin{aligned} & \alpha(a, X)\, U + (1 - \alpha(a, X))\, \alpha(b, X)\, U = u(X) = pU \\ \Leftrightarrow \quad & \alpha(b, X) = \frac{p - \alpha(a, X)}{1 - \alpha(a, X)} \end{aligned} \tag{14}$$

with $\alpha(a, X) < p$ to ensure $\alpha(b, X) > 0$.

Suppose first that $a \hat{\succ} b$. As in example 1, since the monotonicity condition on α requires that $\alpha(b, X) \leq \alpha(b, \{b\})$, it must also be

$$\frac{p - \alpha(a, X)}{1 - \alpha(a, X)} \leq pq \Leftrightarrow \alpha(a, \{a, b\}) \geq \frac{p(1-q)}{1 - pq}$$

In short, then, we have $\alpha(a, \{a, b\}) \in \left[\frac{p(1-q)}{1-pq}, p\right) \neq \varnothing$ and $\alpha(b, X) = \frac{p - \alpha(a,X)}{1 - \alpha(a,X)}$.

Now consider the alternative case $b \hat{\succ} a$. As the two attention parameters $\alpha(a, X)$ and $\alpha(b, X)$ are completely symmetric, we obtain $\alpha(b, X) \in \left[\frac{p(1-q)}{1-pq}, p\right)$ and $\alpha(a, X) =$

$\frac{p-\alpha(b,X)}{1-\alpha(b,X)} \Leftrightarrow \alpha\left(b,X\right) = \frac{p-\alpha(a,X)}{1-\alpha(a,X)}$. That is, while equation (14) establishes the same condition regardless of whether $a\hat{\succ}b$ or $b\hat{\succ}a$, the monotonicity condition imposes a different range of values for the attention parameters. For instance, setting $p = 0.6$ and $q = 0.8$, $\alpha\left(a,X\right) = 0.2$ and $\alpha\left(b,X\right) = 0.5$, the requirements for the case $a\hat{\succ}b$ fail (since for that case $\alpha\left(a,X\right) \in [0.23, 0.6)$) while those for $b\hat{\succ}a$ hold (since $0.5 = \alpha\left(b,X\right) \in [0.23, 0.6)$ and $\alpha\left(a,X\right) = \frac{0.6-0.5}{1-0.5} = 0.2$).

More in general, for any $\varepsilon \in \left(0, \min\left\{1-p, p\left(1-q\right)\right\}\right)$, the case $a\hat{\succ}b$ allows for $\alpha\left(a,X\right) = p-\varepsilon$ and $\alpha\left(b,X\right) = \frac{p-(p-\varepsilon)}{1-(p-\varepsilon)} = \frac{\varepsilon}{1-p+\varepsilon} > 0$; since however $\frac{\varepsilon}{1-p+\varepsilon} < \frac{p(1-q)}{1-pq}$ given our condition on ε, this value falls outside the range for $\alpha\left(b,X\right)$ when $b\hat{\succ}a$.

Finally, we note that taking e.g. the case $a\hat{\succ}b$, if we also required $\alpha\left(a,X\right) \leq \alpha\left(a,\{a\}\right) = pq$, we would have

$$\begin{aligned}\frac{p\left(1-q\right)}{1-pq} &< pq \Leftrightarrow \frac{1-q}{1-pq} < q \Leftrightarrow 1-2q+pq^2 < 0 \\ &\Leftrightarrow p < \frac{2q-1}{q^2}\end{aligned}$$

If q is sufficiently small, the rhs in the last inequality is negative, so that the condition cannot hold. For instance with $U = 12$, $p = \frac{1}{4}$ and $q = \frac{1}{3}$ so that utilities are $u\left(\varnothing\right) = 0$, $u\left(a\right) = u\left(b\right) = 12 > 0$, $u\left(\{a,b\}\right) = 3$, $u\left(\{a\}\right) = 1 = u\left(\{b\}\right)$. Then $\alpha\left(a,\{a\}\right) = \alpha\left(b,\{b\}\right) = \frac{1}{12}$, while from

$$u\left(X\right) = 3 = \alpha\left(a,X\right)12 + \left(1-\alpha\left(a,X\right)\right)\alpha\left(b,X\right)12$$

we obtain

$$\alpha\left(b,X\right) = \frac{1-4\alpha\left(a,X\right)}{4\left(1-\alpha\left(a,X\right)\right)}$$

which is positive since by monotonicity it must be $\alpha\left(a,X\right) \leq \alpha\left(a,\{a\}\right) = \frac{1}{12} < \frac{1}{4}$. On the other hand, since by monotonicity we must also have $\alpha\left(b,X\right) \leq \frac{1}{12}$, it follows that

$$\frac{1-4\alpha\left(a,X\right)}{4\left(1-\alpha\left(a,X\right)\right)} \leq \frac{1}{12} \Leftrightarrow \alpha\left(a,X\right) \geq \frac{2}{11} > \frac{1}{12}$$