Repetition Aversion and Dynamic Consideration

Paola Manzini
University of Bristol

Marco Mariotti
Queen Mary University of London

Evgenii Safonov
Queen Mary University of London

Levent Ülkü
ITAM*

This version: March 2026

Abstract

We model an agent who holds a stable intrinsic preference over alternatives but also experiences *repetition aversion*—a dislike for choosing the same object again "too soon", which leads the agent to temporarily exclude it from consideration. A long sequence of the agent's choices is observed. Repetition aversion fundamentally disrupts the usual connection between preference and choice: even when x is *both* strictly intrinsically preferred to y and less repetition aversive, it may fail to be chosen more frequently than y. We provide a complete characterisation of repetition-averse choice sequences and establish how intrinsic preference can be disentangled from repetition aversion.

J.E.L. codes: D0.

Keywords: Dynamic choice, Consideration sets.

*Levent Ülkü acknowledges financial support from the Asociaciòn Mexicana de Cultura. Manzini and Mariotti are grateful to ITAM for their generous hospitality during several visits. We thank audiences at Royal Holloway University of London, the University of Glasgow, Tel Aviv University, the BRIC workshop in Mexico City and the D-TEA workshop in Paris.

1 Introduction

It is unlikely that those who regard *Anna Karenina* as the greatest novel ever written feel compelled to reread it every month. Even experiences of the highest quality may generate extended periods of satiation. Economic decisions are often shaped by such history dependence: consumers avoid repeating meals, viewers rotate entertainment choices, and workers diversify tasks. This psychological dimension of preference lies outside the scope of standard timeless models of discrete choice. We propose a dynamic discrete choice framework that captures this dimension in a parsimonious way. We study how repetition aversion distorts conventional preference revelation, how intrinsic preferences over alternatives can nevertheless be identified, and how the model can be characterised exactly by properties of observable choice behaviour. We also show, through examples, that long-run behaviour can be highly sensitive to changes in the environment and thus, in principle, susceptible to manipulation.

To illustrate the central ideas, consider Monsieur Arthens, who dines out every Saturday. Asked to rank his favourite restaurants, he responds without hesitation: S (a seafood restaurant) is the best, followed by V (a vegan restaurant), and next B (a burger restaurant). This ranking constitutes Arthens' *intrinsic preference*. However, he dislikes consuming seafood in two consecutive weeks, prefers not to eat vegan food more often than once every three weeks, and follows medical advice to limit beef consumption to no more than once a month. These constraints define Arthens' *repetition aversion*.

Each week, he chooses among the restaurants offering a type of food not consumed within these respective temporal bounds, and selects the most preferred among them. The resulting routine unfolds as summarised in Table 1:

Weeks	1	2	3	4	5	6	7	8	...
Restaurant	S	V	S	B	S	V	S	B	...

Table 1: Arthens' routine

Arthens first dines at S, switches to V the following week to avoid repeated seafood consumption, returns to S, then opts for B since both seafood and vegan have been consumed too recently. This pattern then repeats indefinitely.

A striking feature of this routine concerns preference revelation. The burger restaurant is *strictly* worse than the vegan restaurant along both dimensions: it ranks lower in intrinsic preference and it is associated with a longer aversion phase. Yet B is chosen just as frequently as V, namely once every four periods. Strict inferiority along both preference dimensions is therefore not reflected in choice frequencies. An external observer relying solely on choice frequency data might entirely fail to uncover Arthens' inclination for a healthier diet.

In this simple example, if the observer accepts our proposed behavioural mechanism, there is a straightforward route to partial preference revelation. The order of appearance of the restaurants in the sequence must reflect the intrinsic preference ranking: on the first Saturday Arthens chooses his favourite restaurant, then as soon as he cannot repeat the experience he switches to the second favourite, and so on. Unfortunately, this methodology is not always portable to more general situations, in which menus may vary and the satiation state at the time of the first observation may be unknown. We therefore develop more robust ways in which choice patterns reveal preference, by identifying periods in which an alternative is demonstrably not subject to aversion, yet another alternative is chosen.

A second question concerns how an observer could learn from the data whether or not Arthens' behaviour is compatible with the mechanism we propose, namely one driven by the interaction of intrinsic preference and repetition aversion. Which features of observed choice sequences render such an explanation possible? The answer is closely connected to the identification problem highlighted in the preceding discussion. One approach is to examine the consistency between the deeper form of preference revelation and the simpler order of appearance argument when the latter is available. Another is to check the internal coherence of the deeper revelation mechanism directly.

The restaurant example above can be interpreted through the lens of choice with *dynamic consideration sets*: every Saturday, the agent selects the best restaurant from those currently under consideration, where the consideration set itself depends on the history of past choices. The extensive literature on static consideration sets has shown that preference revelation may be distorted when attention or other cognitive constraints limit the set of considered alternatives. But, in such cases, discrepancies between preference and choice are due to a misalignment between attention/salience and preference parameters: an inferior but more salient alternative may be chosen as often as a better one. By contrast, no such misalignment is explicitly present in the example, in which vegan is superior in both aspects. In this case the tension has deeper roots than the fact that highly regarded alternatives may also cause more aversion to repeated consumption.

We posit that the analyst observes an idealised, open-ended dataset consisting of an infinite sequence of choices made by an agent over time over finite menus of alternatives. At each point in time, the agent chooses the most preferred option among those toward which they are not currently averse. In our baseline setting—with a fixed aversion phase and a fixed menu—we identify a central behavioural property, stated as a novel *Worse Follows Faster* condition: if x is intrinsically (revealed) preferred to y, then whenever y follows x in the sequence before x is chosen again, it must do so *faster* than x ever follows itself. Worse Follows Faster admits an equivalent formulation in terms of a dynamic analogue of the standard Weak Axiom of Revealed Preference.

We show that the core insights of the baseline model are robust to several extensions, most notably to environments with unknown initial aversion and menu variation (Section 3) and to settings with history-dependent forms of repetition aversion (Section 4). In each case, the core observable properties take the form of variants of Worse Follows Faster or dynamic analogues of the Strong Axiom of Revealed Preference. We also demonstrate that long-run behaviour can be highly sensitive to seemingly minor features of the environment. Small changes in initial aversion or in the sequence of menus faced by the agent may induce regime shifts in the resulting pattern of choice.

In section 5 we study to what extent the parameters of repetition aversion can also be identified. In the concluding Section 6, beside other remarks, we sketch how the model can serve as a foundation for stochastic choice, where long-run choice frequencies are interpreted as choice probabilities. Viewed from this perspective, the model falls outside of the Random Utility Maximisation family.

1.1 Related Literature

A psychological feature related to what we term repetition aversion is described in the literature as a *taste for variety*. In a static setting, for example, Kaiser and Schwabe [11] axiomatise a model of preferences over menus in which the agent prefers to choose from menus with more varied alternatives. The realised sequence of choices in our model will often exhibit variety, though strictly speaking our framework does not impute to the agent any positive "taste" for variety, nor are intrinsic preferences modified by previous choices.

Structurally, our model is more closely related to a class of *dynamic* choice models in which current preferences depend on past consumption. Traditionally, the link between past and present choices has been studied through the lens of *habit formation* or *addiction*, as in the classical work by Becker and Murphy [4] and Pollak [18], and more recently Rozen [20].

Turansick [21] develops a general model of intertemporal consumption dependence within Random Utility Maximisation, where a transition function specifies, conditional on a current preference and choice, a distribution over next-period preferences. By contrast, in Section 6 we show that a stochastic version of our model falls outside of the Random Utility Maximisation class.

More directly relevant to the present paper are the contributions of Gilboa and Schmeidler [9] (henceforth GS), Ashkenazy-Golan *et al.* [2] (AKL) and Puig [19] (P) who also study subjectively discouraging effects of repeated consumption or choice. In these models, a decision maker repeatedly chooses from a fixed, finite menu of alternatives, as in our baseline setting. In GS, *cumulative utility* is updated at each period by adding a component—negative in the case of satiation—to the current cumulative utility of the chosen alternative. AKL assume that the intrinsic utility from choosing an alternative x is multiplied by a factor that depends on the frequency with which x was chosen in the past. P introduces the *history-discounted utility* (HDU) model of preferences, whereby the intrinsic utility of x is discounted by a constant satiation parameter each time it is chosen, and increases according to a constant recovery parameter when it is not. P considers structured alternatives of the form (lottery, money).

In terms of behavioural assumptions, our framework is closest in spirit to GS, who posit a "greedy" strategy of period-by-period optimisation and provide an articulated justification for it.[1] However, our approach differs markedly from that of all three papers in that we model the effects of satiation not as changes in the agent's preferences over alternatives, but as arising from a dynamic process of consideration set formation under a stable intrinsic preference. Moreover, unlike AKL, GS, and P, which rely on specific utility representations, our theory is purely *ordinal* in its treatment of preferences. Our analysis focuses on the exact axiomatic characterisation of behaviour and its revealed-preference implications. GS and AKL address different questions concerning the structure of optimal dynamic strategies. For example, AKL examine alternative evaluation methods for infinite streams of payoffs and show that a stationary strategy is optimal under the discounted present-value criterion, while GS relate greedy maximisation to the maximisation of a potential function. P, in turn presents an axiomatisation of HDU binary preferences over finite choice sequences. Earlier works within modified discounted-utility frameworks, with some affinity to P's model, include McAlister [15], Baucells and Sarin [3] and He *et al.* [10].

Another related paper is Ferreira and Gravel [6], which also studies choice behaviour observed as a chronological sequence rather than as timeless menu–choice pairs. Their focus, however, is on preference discovery, where agents learn their preferences through experimentation over time. In contrast, our model assumes stable intrinsic preferences and attributes temporal patterns in choice to repetition aversion operating through dynamically evolving consideration sets, rather than to learning about preferences.

At the conceptual level, our decision maker is reminiscent of the consumer in the advertising model of Eliaz and Spiegler [5], in which consumers on non-retail platforms are assumed to alternate between two mental states: a demand state and a satiation state. In the demand state, a consumer buys a product with positive probability if targeted by advertising; while during satiation, the consumer is uninterested in consumption. A pur-

[1] At the analytical level, GS can be seen as a specialisation of the general case-based choice model of Gilboa and Schmeidler [8]. In this latter setting, however, choice histories are interpreted to affect current preferences through an information channel.

chase triggers entry into satiation. Return to the demand state occurs with a constant per-period probability. While the focus of their paper is very different (namely, the optimal diversification of advertising), the fundamental two-state Markov dynamics for each alternative is a key ingredient in our model, and the average dormancy period in their model is akin to the fixed aversion phase in our baseline model.[2]

To the extent that our model can be interpreted as providing a foundation for stochastic choice—with long-run choice frequencies viewed as probabilities—it also relates to experimental implementations such as Agranov and Ortoleva [1], who study deliberate randomisation in repeated choice from a fixed menu. In our setting, agents do not literally randomise, but also generate a deterministic sequence of choices in which different alternatives may be selected at different times from the same menu. A closer conceptual connection in this respect arises with Lu and Saito [14], who study an agent repeatedly choosing from changing menus, thereby potentially inducing a stochastic choice function in the limit. Beyond this shared point of departure, our analyses diverge sharply. Lu and Saito's main question is when a stochastic choice function can be represented as a limiting frequency, under the assumption that the agent behaves dynamically optimally with respect to a utility function over lotteries evolving according to an ergodic Markov process.

As noted earlier, our model can be viewed through the lens of the theory of choice with consideration sets—a literature too extensive to be surveyed here (see, for early axiomatic treatments of the deterministic and stochastic cases, Masatlioglu *et al.* [17] and Manzini and Mariotti [16], respectively). We return to this connection in Section 6.

2 The Baseline Model

In this section we introduce our baseline model. Here repetition aversion is captured by a fixed aversion phase following the choice of an alternative.

Let X be a non-empty and finite set of alternatives, and $\tilde{X} = X \cup \{a\}$, where $a \notin X$ stands for abstention (the choice of an outside option). We denote $\mathbb{N} = \{1, 2, ...\}$, $\mathbb{Z}_+ =$

[2] We thank Rani Spiegler for bringing this connection to our attention.

$\{0, 1, ...\}$ and $\tilde{\mathbb{Z}}_+ = \mathbb{Z}_+ \cup \{\infty\}$. The observable is a *choice sequence* $c = (c_t)_{t \in \mathbb{N}}$ with $c_t \in \tilde{X}$ for all t. Here, c_t is interpreted as the selection out of $\tilde{X}$ made by the agent at date t. A choice sequence is an idealised dataset, meant to represent an arbitrarily long record of choice observations.

We stick throughout with the conventions that for any real number r, $-\infty < r < \infty$, $r + \infty = \infty$, $r - \infty = -\infty$, and also that $-(-\infty) = \infty$, and $\infty \geq \infty$. For a choice sequence c, time period $t \in \mathbb{N}$, and alternative $x \in \tilde{X}$, denote

$$last^c(x, t) = \max(\{s < t : c_s = x\} \cup \{-\infty\}). \tag{1}$$

Hence, $last^c(x, t)$ is the last period before time t where x is the choice in c. If $c_s = x$ for no $s < t$, then $last^c(x, t) = -\infty$. When the choice sequence being referred to is obvious, we suppress the dependence on c and write $last(x, t)$.

Given a binary relation $\succ$ on $\tilde{X}$ and a set $A \subseteq \tilde{X}$, we denote the $\succ$ $-$maximal set of A by $\max(A, \succ) = \{x \in A : y \succ x \implies y \notin A\}$.

Definition 1. A choice sequence c is *repetition averse* if there exists a pair $(\succ, p)$ where $\succ$ is a linear order[3] on $\tilde{X}$, $x \succ a$ for all $x \in X$, $p : \tilde{X} \to \tilde{\mathbb{Z}}_+$, and $p(a) = 0$, such that

$$\{c_t\} = \max\left(\left\{x \in \tilde{X} : t - last(x, t) - 1 \geq p(x)\right\}, \succ\right) \tag{2}$$

for every t. In this case, we say that $(\succ, p)$ represents c.

In this representation, we say that x *is considered (by c) at t* if $t - last\,(x, t) - 1 \geq p\,(x)$, that is, if the number of periods $(t - last\,(x, t) - 1)$ in which the agent has not chosen x prior to period t weakly exceeds $p(x)$. At any period the agent maximises their preference over the alternatives in $\tilde{X}$ that are being considered at that period. The requirement $p(a) = 0$ ensures that the outside option a is always considered, however, since it is not preferred to any of the other alternatives, it is the choice if and only if no other alternative is considered. When an alternative $x \in X$ is chosen at some period s, it is removed from the set of considered alternatives for $p(x)$ periods, the *aversion phase*. Therefore, the earliest period x can be chosen again is $s + p(x) + 1$. If $last(x, t) = -\infty$, i.e. if x was never

[3] Hence, $\succ$ is complete, transitive, and asymmetric, which guarantees the existence of a unique maximal element in any non-empty subset of $\tilde{X}$ with respect to $\succ$.

the choice before t, then x is considered at t. This is true even if $p(x) = \infty$ since, by our conventions, $t - (-\infty) = \infty \geq \infty$. In particular, all alternatives are considered at $t = 1$ and c_1 maximizes $\succ$ on X.

Since consideration, and consequently the current choice, depend on past choices, the model is recursive. Note also that if $p(x) = \infty$, x can be the choice at most once in the sequence. This captures choices made once in a lifetime, if at all. Otherwise, x may be chosen an infinite number of times.

The restaurants example in the Introduction describes a repetition averse choice sequence. As illustrated by the example, two alternatives may be chosen with the same non-zero frequency along a repetition averse choice sequence even though one dominates the other in both primitives, i.e. ranks higher in the preference ordering and has a shorter aversion phase. It can be shown however that if $x \succ y$ and $p(x) < p(y)$, then x cannot be chosen less frequently than y (see Appendix C, Theorem 24).

We are interested in the observable "signature" of repetition aversion: which features of a choice sequence characterise it as repetition averse? The following example gives an idea of the task at hand.

Example 2. Let $X = \{x, y, z, w\}$. Consider the following two choice sequences:

Sequence I consists of an infinite repetition of the initial subsequence $(c_1, ..., c_{48})$ provided below

$$x\,y\,z\,w\,x\,a\,y\,a\,x\,z\,a\,y\,x\,w\,a\,z\,x\,y\,a\,a\,x\,z\,y\,w\,x\,a\,a\,y\,x\,z\,a\,w\,x\,y\,a\,z\,x\,a\,y\,w\,x\,z\,a\,y\,x\,a\,a\,z$$

Sequence II starts with the subsequence $(c_1, ..., c_{48})$ of Sequence I and then continues with an infinite repetition of the 32-element subsequence $(c_{49}, ..., c_{80})$

$$x\,y\,w\,a\,x\,z\,y\,a\,x\,a\,w\,y\,x\,z\,a\,a\,x\,y\,w\,z\,x\,a\,y\,a\,x\,z\,w\,y\,x\,a\,a\,z$$

As our results will show, Sequence II is repetition averse. In fact, it can be checked that it is represented by the preference $x \succ y \succ z \succ w$ over alternatives and by the aversion phases $p(x) = 3, p(y) = 4, p(z) = 5, p(w) = 7$. We will turn to Sequence I later.

This example highlights the fact that, even with a small menu, the pattern of repetition averse sequences can be quite complex and repetition aversion may be far from obvious to detect.

Towards a characterisation, we need an additional piece of notation. For any c, define the binary relation E^c on $\tilde{X}$ as follows:

$$xE^c y \iff \min(\{t : c_t = x\} \cup \{\infty\}) < \min(\{t : c_t = y\} \cup \{\infty\}). \tag{3}$$

In words, $xE^c y$ if the first appearance of x in the sequence c is earlier than that of y or if x appears in c and y does not.[4] We call E^c the *precedence* relation. We will simply write xEy if the underlying choice sequence is obvious.

We shall require only two properties for our characterisation.

Worse Follows Faster: For all $s, t \in \mathbb{N}$: $c_t E c_s \implies t - last(c_t, t) > s - last(c_t, s)$.

Abstention Last: aEx for no $x \in X$.

Abstention Last says that if the abstention alternative a appears at all in the choice sequence, it cannot do so before all alternatives have been chosen at least once. This property simply captures in terms of choice the nature of a as a default alternative.

Turning to our key axiom, Worse Follows Faster, we first offer an intuitive description, followed by a revealed preference interpretation. Let $c_t = x$ and $c_s = y$. The axiom requires that if x precedes y in the sequence, then *whenever y follows x before x is chosen again, it must do so more rapidly than x ever follows itself*. In fact, the "gap" between a period t at which x is chosen and the preceding period at which it was chosen is required to be larger than the gap between a period s, at which y is chosen, and the most recent earlier period at which x was chosen (i.e., $last(c_t, s)$). This inequality must hold for all such periods t and s.

Note that Worse Follows Faster does not preclude the possibility that some alternative x appears for the first time earlier than another alternative y, and also sometimes follows itself more rapidly than y follows x. This is illustrated by Sequence II in Example 2, where x appears for the first time before z does, but it follows its second appearance faster than z does. Nonetheless, the sequence satisfies Worse Follows Faster: the minimum gap between two choices of x is three periods, whereas whenever z follows x before x appears again, it does so with a delay of at most two periods.

[4]In the latter case $\min\{t : c_t = x\} < \infty = \min\{t : c_t = y\}$. If both x and y do not appear in c, then $\min(\{t : c_t = x\} \cup \{\infty\}) = \infty \geq \infty = \min(\{t : c_t = y\} \cup \{\infty\})$

By contrast, Sequence I of Example 2 violates Worse Follows Faster. At $t = 51$, z follows itself after a gap of three periods, whereas w sometimes follows z after a gap of four periods. Since w appears for the first time later than z, this contradicts Worse Follows Faster. Similarly, a sometimes follows z after a gap of three periods, and appears for the first time later than z, again contradicting Worse Follows Faster.

Revealed preference interpretation of Worse Follows Faster: There are two distinct ways in which the underlying preference is revealed by a repetition-averse choice sequence. The first exploits the precedence order. It is easy to verify that if c is represented by $(\succ, p)$ and xEy, then $x \succ y$. As previously noted, all alternatives not yet chosen at time t are considered available. Consequently, an alternative can only be chosen for the first time after all strictly preferred alternatives have already been chosen at least once.

We now turn to the second, less direct channel through which $\succ$ is revealed. Denote again $c_t = x$ and $c_s = y$. The interval $t - last(c_t, t)$ provides an upper estimate for $p(x)$. In fact, for x to be chosen at t, it must be under consideration; that is, $t - last^c(x, t) - 1 \geq p(x)$. Now suppose that the conclusion of Worse Follows Faster did *not* hold, that is, $t - last(c_t, t) \leq s - last(c_t, s)$. This would imply that y is chosen at a period s when x is surely under consideration, revealing that $y \succ x$ in contradiction to xEy. Worse Follows Faster thus ensures consistency between the two modes of preference revelation embedded in the model.

Theorem 3. *The following statements are equivalent for a choice sequence c:*

(i) c is repetition averse;

(ii) c satisfies Worse Follows Faster and Abstention Last;

(iii) c satisfies Abstention Last and the binary relation $\succ^c$ on $\tilde{X}$ defined below is asymmetric:

$$x \succ^c y \iff x \neq y, \exists s, t \in \mathbb{N} : c_t = x, c_s = y \text{ and } t - last^c(y, t) \geq s - last^c(y, s) \quad (4)$$

All proofs are provided in Appendix A. Here we sketch the argument for Theorem 3. The revealed preference interpretation of Worse Follow Faster shows the necessity of this axiom, while the necessity of Abstention Last is obvious. To prove sufficiency, we extend the precedence relation to a linear order on $\tilde{X}$ for which a is the worst alternative.

To construct the aversion phases, we define, for every $x \in X$,

$$p(x) = \min \{t - last(x,t) - 1 : x = c_t\} \tag{5}$$

if x is sometimes chosen, and $p(x) = 0$ otherwise. That is, we infer the aversion phase of an alternative x that is sometimes chosen to be the minimum observed gap between two consecutive occurrences of x. As noted earlier, this constitutes only an upper bound on the true aversion phase: in practice the latter may be shorter since an alternative might fail to be chosen at some period not because it is unconsidered, but because a better alternative is considered and chosen. Nevertheless, an inductive argument shows that, when combined with the constructed preference relation, these inferred aversion phases generate precisely the observed sequence.

The equivalent formulation of Worse Follows Faster given in statement (iii) of the theorem disciplines instead the nontrivial revealed preference relation directly, with no reference to the precedence relation E. Rather than requiring consistency between E and the revealed preference, it suffices—equivalently—to demand that the revealed preference is not self-contradictory. The relation $\succ^c$ is precisely the revealed preference for our model of repetition aversion, as discussed above.

Requiring Asymmetry of the revealed preference relation is analogous to the Weak Axiom of Revealed Preference in standard choice models, noting however that in the present dynamic environment the composition of the set over which a preference is maximised at any given period must be inferred from the choice sequence itself.

In the remainder of the paper, we study several extensions and modifications of the setting presented in this section. As we will see, the flavour of the previous results remains unchanged: all the model variations can be characterised in terms of similar axioms with similar interpretations. At the same time, some insights will emerge on the dynamics of behaviour that are hidden in the baseline model.

3 Initial Aversion and Menu Variation

In this section we relax two somewhat idealised assumptions of the baseline model: first, that in the initial period the consumer is not averse to any alternative; and second, that the menu remains fixed over time. Besides characterising the extended model, we argue through examples that repetition aversion can render long-run behaviour highly sensitive to small variations in initial aversion and in the sequencing of menus.

In the preceding analysis we maintained the assumption that the observer has access to the entire choice history from its inception and is aware of this fact. This assumption is appropriate in certain contexts—for instance, when the consumption of a new product is monitored from the moment of its market introduction. In other situations, however, observation may begin only after a choice process is already underway, with the observer lacking access to the full record of past choices or knowledge of when the sequence itself began. In such cases, the consumer may already be averse to some alternatives even at the start of observations. We capture this possibility by adding a vector of unobserved initial conditions among the parameters of the model. We use a slightly different notation that is more convenient to examine this case.

Turning to the second idealised assumption, in practice the set of available alternatives seldom remains perfectly stable: venues may close, products may be temporarily unavailable, transport may be disrupted, budgets may fluctuate, and so forth. Menus may also be strategically manipulated by an actor seeking to influence the choice sequence. We accommodate this by extending the primitive. A *menu-choice sequence* is a sequence $(M, c) = (M_t, c_t)_{t \in \mathbb{N}}$ satisfying $c_t \in M_t \cup \{a\}$ and $M_t \subseteq X$ for every t. Note that $M_t = \emptyset$ is permitted.

Definition 4. A menu-choice sequence (M, c) is *repetition averse* if there is a tuple $(\succ, p, w_1)$, where $\succ$ is a linear order on $X \cup \{a\}$ such that $x \succ a$ for all $x \in X$, $p : \tilde{X} \to \tilde{\mathbb{Z}}_+$ such that $p(a) = 0$, and $w_1 : \tilde{X} \to \tilde{\mathbb{Z}}_+$ such that for every t,

$$\{c_t\} = \max\left(\{x \in M_t \cup \{a\} : w_t(x) \geq p(x)\}, \succ\right) \tag{6}$$

where, for $t > 1$, w_t is defined recursively via

$$w_{t+1}(x) = \begin{cases} w_t + 1 & if\ c_t \neq x \\ 0 & if\ c_t = x \end{cases} \tag{7}$$

In this case, we say that $(\succ, p, w_1)$ represents c.

In this representation, $w_t(x)$ is a "waiting length", namely the number of periods the agent has not been choosing alternative x. We say that x *is considered (by c) at t* if $w_t(x) \geq p(x)$ and $x \in M_t \cup \{a\}$; that is, if the waiting length for x is at least as large as x's aversion phase $p(x)$, and x is available. Note that $w_t(x)$ drops to zero every time that alternative x is chosen in the previous period. The parameters $w_1(x)$, with $x \in X$, are the unobserved initial conditions of the agent's choice sequence; this the waiting length for alternative x at the start of the observed choice. Our baseline model can be represented using this notation by setting $w_t(x) = t - last(x,t) - 1$ for all $x \in X$ and $t \in \mathbb{N}$. Because $last(x,1) = -\infty$, this implies that $w_1(x) = \overline{w}_1(x) = \infty$ for all x in the baseline model.

Any tuple $(\succ, p, w_1)$ that satisfies the conditions of Definition 4 induces, for any sequence of menus M_t, a unique well-defined choice sequence $c_t \in M_t \cup \{a\}$, since abstention is always considered, and $\succ$ is a linear order.

Let us examine how the intrinsic preference can be revealed from the consumer's choice. For x to be preferred to y, it must be the case that x is chosen in some period t when y is available and considered. Since $w_t(x)$ is not observable before the first choice of alternative x, we introduce two additional pieces of notation. Given a choice sequence c, define $\underline{w}_t, \overline{w}_t : \tilde{X} \to \mathbb{N}$ for $t = 1, 2, ...$ as follows: $\underline{w}_1(x) = 0$, $\overline{w}_1(x) = \infty$, and

$$\underline{w}_{t+1}(x) = \begin{cases} \underline{w}_t(x) + 1 & if\ c_t \neq x \\ 0 & if\ c_t = x \end{cases} \qquad \overline{w}_{t+1}(x) = \begin{cases} \overline{w}_t(x) + 1 & if\ c_t \neq x \\ 0 & if\ c_t = x \end{cases} \tag{8}$$

for all $x \in X$. Here, we use our usual convention $\infty + 1 = \infty$.

If (M, c) has a repetition aversion representation $(\succ, p, w_1)$, then for all $x \in X$, we have $\underline{w}_t(x) \leq w_t(x) \leq \overline{w}_t(x)$, clarifying the meaning of our notations. Moreover, if x occurs in the choice before period t, then $\underline{w}_t(x) = w_t(x) = \overline{w}_t(x)$. For any menu-choice sequence (M, c) and for any distinct $x, y \in \tilde{X}$, define the following revealed preference:

$$x \succ^* y \iff \exists t, s \in \mathbb{N}\ such\ that\ c_t = y,\ y \in M_s,\ c_s = x,\ and\ \underline{w}_t(y) \geq \overline{w}_s(y) \tag{9}$$

If (M, c) is a repetition aversion sequence, the condition $\underline{w}_t(y) \geq \overline{w}_s(y)$ implies $w_t(y) \geq w_s(y)$. Thus, alternative y is considered in period t and hence, $x \succ^* y$ implies $x \succ y$. Consequently, $\succ^*$ cannot have cycles, and there is no x such that $a \succ^* x$. Our next result shows that these two properties characterise repetition averse menu-choice sequences.

Theorem 5. *A menu-choice sequence (M, c) is repetition averse if and only if $\succ^*$ is acyclic and $a \succ^* x$ for no $x \in X$.*

To prove the "if" part of the theorem, we construct a repetition averse representation $(\succ, p, w_1)$ of c. For $\succ$, we pick any linear order on $X \cup \{a\}$ that completes the revealed preference $\succ^*$ and satisfies $x \succ a$ for all $x \in X$; the conditions of the theorem allow us to do it. Next, we find, for each alternative x, the lower and upper bounds on its aversion phases consistent with the observed choices and the choice of $\succ$, and show that we can find $p(x)$ that satisfies the corresponding conditions. We present the bounds on $p(x)$ in Section 5, Theorem 14. Finally, for each x, we pick

$$w_1(x) = \max\{0, p(x) - \min\{t : c_t = x\} + 1\} \tag{10}$$

making the model consistent with the data for the periods before the first choice of x.

In the baseline model, the intrinsic preference $\succ$ and the aversion phases p completely determine the choice c and therefore the frequency with which each item is consumed. When the agent may have some initial aversion, this is no longer true:

Example 6. Let $X = \{x, y, z\}$, the agent's preference be $x \succ y \succ z \succ a$, and the aversion phases be $p(x) = 3$, $p(y) = 3$, and $p(z) = 1$ (with $p(a) = 0$ as usual). Let the menu sequence be constant: $M_t = X$ for all t. Suppose the agent is not initially averse to any item, that is, $w_1(x) = w_1(y) = w_1(z) = \infty$; this case corresponds to the baseline model. Then the choice is an infinitely repeating cycle $c = (x, y, z, a, ...)$. Hence, each item and the outside option are chosen with equal frequencies $1/4$. Suppose instead that $w_1(y) = 1$, and $w_1(x) > 3$. In this case, the choice is an infinitely repeating cycle $c = (x, z, y, z, ...)$. Thus, items x and y are still chosen with frequencies $1/4$, but item z is chosen with a double frequency $1/2$, and the outside option is never chosen.

Consider a hypothetical seller seeking to maximise long-run sales of products x, y, z to the consumer described in Example 6. If the consumer were initially averse just enough to product y and not averse to product x at the start of the relationship, the seller would not need to take any action to induce the choice sequence $c = (x, z, y, z, ...)$, thereby maximising sales. However, the consumer might initially not be averse to any product. It is implausible that the seller can manipulate the consumer's initial aversion levels, and if the seller takes no action, the resulting choice sequence is $c = (x, y, z, a, ...)$, which yields substantially lower sales. Yet, the seller can manipulate the menu in a single period—namely, period two—to induce the sequence $c = (x, z, y, z, ...)$:

Example 7. Let $X = \{x, y, z\}$, the agent's preference be $x \succ y \succ z \succ a$, and the aversion parameters be $p(x) = 3$, $p(y) = 3$, and $p(z) = 1$. Suppose the agent is not initially averse to any item, that is, $w_1(x) = w_1(y) = w_1(z) = \infty$. If $M_t = X$ for all t, then the choice is an infinitely repeating cycle $c = (x, y, z, a, ...)$, and each item and the outside option are chosen with equal frequencies $1/4$. If the sequence of menus is instead $X, \{x, z\}, X, X, X, ...$, that is, if $M_t = X$ for $t \neq 2$, and $M_2 = \{x, z\} = X \backslash \{y\}$, then the choice is an infinitely repeating cycle $c = (x, z, y, z, ...)$, items x and y are chosen with frequencies $1/4$, item z is chosen with frequency $1/2$, and the outside option is never chosen. The same choice occurs if the sequence of menus is $\{x, z\}, \{x, z\}, X, X, X, ...$

Examples 6 and 7 show that the long-run frequencies of choice may depend on the details of the environment at the beginning of observations: how averse the consumer is at the start, and what particular menus appear in the early periods.

The structural reason for this fragility of long-run consumer behaviour lies in the dynamics of the state variable in the underlying dynamic process. The agent's choice is governed by a vector w of wait lengths, whose components are truncated at the relevant aversion thresholds and evolve deterministically over time. This evolution defines a deterministic Markov chain on the state space of wait vectors w. Its absorbing sets coincide with the directed cycles of the associated graph.

If there is exactly one cycle, long-run choice frequencies are independent of the initial conditions. Otherwise, the initial conditions determine which absorbing cycle of the Markov chain is reached and hence the long-run choice behaviour. In Examples 6 and 7, the associated Markov chain has 3 cycles. These correspond to the choice cycles (x, y, z, a), (x, z, y, z) and (y, x, z, a) discussed above. The first and third cycles generate the same long-run choice frequencies, whereas the second yields different frequencies.

In Example 7 the seller might incur some cost of menu adjustment, but even so large effects may be obtained at low cost: note the menu needs changing only once or twice to achieve the desired frequency of choices.

4 History Dependent Aversion

In this section we consider two generalisations of the baseline model that allow for richer forms of history dependence. First, we allow the aversion phase to depend arbitrarily on past choices, subject only to the requirement that it never decreases. This specification may generate choice sequences that do not eventually cycle, unlike the baseline model.

The second extension captures the idea that recent and frequent consumption of an alternative increases aversion to it, by assuming that consideration depends on a finite window of the alternative's own past consumption. The resulting choice sequence eventually cycles, as in the baseline model, but can accommodate behavioural patterns that neither the baseline model nor the first history-dependent specification can capture.

4.1 Monotonic Repetition Aversion

It is possible that the gradual accumulation of past experiences of an option makes repetition aversion more acute. In the baseline model, an alternative is either chosen infinitely often, or it is chosen exactly once (in the case of an infinite repetition aversion), or it is never chosen. The model cannot handle more nuanced situations of increasing fatigue, whereby, for example, choices repeat at increasing intervals, possibly until they stop altogether. Thus, we generalise the baseline model and axioms by allowing the aversion phase associated with an alternative to depend on the entire history of past choices involving that alternative, while imposing the core principle that more past consumption of an alternative cannot reduce aversion towards it.[5]

Given a choice sequence c, denote

$$n_t^c(x) = |\{s \in \{1,...,t-1\} : c_s = x\}| \tag{11}$$

Thus, $n_t^c(x)$ counts the number of times the agent selected alternative x before period t.

[5]We work in the fixed-menu setting and assuming no initial aversion in order to economise on notation. An extension to the setting with initial aversion and menu variation is provided in the Appendix B.

We allow the aversion phase p to depend on $n_t^c(x)$. Aversion is now described by a map $p : \tilde{X} \times \mathbb{Z}_+ \to \tilde{\mathbb{Z}}_+$, where we interpret $p(x,k)$ to be the aversion phase of the alternative x after the agent has chosen it k times in the past. Naturally, we are interested in the case when $p(x, n_t^c(x))$ (weakly) increases if x has been chosen more times in the past.

Definition 8. A choice sequence c is *monotonic repetition averse* if there exists a pair $(\succ, p)$, where $p : \tilde{X} \times \mathbb{Z}_+ \to \tilde{\mathbb{Z}}_+$ is non-decreasing with respect to the second argument for all x, $p(x,0) = 0$ for all $x \in X$, $p(a,k) = 0$ for all k, and $\succ$ is a linear order on $\tilde{X}$ such that $x \succ a$ for all $x \in X$ and, for every t,

$$\{c_t\} = \max\left(\left\{x \in \tilde{X} : t - last^c\left(x,t\right) - 1 \geq p(x, n_t^c(x))\right\}, \succ\right). \tag{12}$$

In this case we say that $(\succ, p)$ is a monotonic repetition aversion representation of c.

A seemingly more general model allows the aversion phase of the alternative to weakly increase with time according to the following definition:

Definition 9. A choice sequence c is *generally monotonic repetition averse* if there exists a pair $(\succ, P)$, where $P : \tilde{X} \times \mathbb{N} \to \tilde{\mathbb{Z}}_+$ is non-decreasing with respect to the second argument for all x, $P(x,1) = 0$ for all $x \in X$, $P(a,t) = 0$ for all t, and $\succ$ is a linear order on $\tilde{X}$ such that $x \succ a$ for all $x \in X$ and, for every t,

$$\{c_t\} = \max\left(\left\{x \in \tilde{X} : t - last^c(x,t) - 1 \geq P(x,t)\right\}, \succ\right). \tag{13}$$

We say that $(\succ, P)$ is a general monotonic repetition aversion representation of c.

One can see that if the choice c is monotonic repetition averse, then c is generally monotonic repetition averse, since time-dependent monotonic aversion phase can always mimic monotone aversion phase that depends only on the number of times the considered alternative has been chosen prior to a given period. Our results show that the two models have the same empirical content. We introduce a directional version of the Worse Follows Faster property:

Worse Follows Faster Earlier: For all $s, t \in \mathbb{N}$ such that $t > s$: $c_t E^c c_s \implies t - last\left(c_t, t\right) > s - last\left(c_t, s\right)$.

Parallel to Worse Follows Faster, axiom Worse Follows Faster Earlier states that if the earliest appearance of x in the choice sequence comes before that of y, then x always must follow itself more slowly than y has followed x before x followed itself *at any earlier date*.

A revealed preference interpretation of Worse Follows Faster Earlier can be given along the same lines as for Worse Follows Faster. Suppose that x, the choice at t, is better than y, the choice at an earlier date s, as revealed by the precedence relation E^c. Then whenever y is chosen, x cannot be considered, implying that x is not yet considered at s after x was chosen at $last\,(x,s)$; that is, the aversion phase for x must be larger than $s - last\,(x,s)$. If the aversion phases are monotonic, either with respect to time or with respect to the number of times x has been chosen before, then x must wait at least as long before being chosen again at any later date.

Theorem 10. *The following statements are equivalent for a choice sequence c:*

(i) c is monotonic repetition averse;

(ii) c is generally monotonic repetition averse;

(iii) c satisfies Worse Follows Faster Earlier and Abstention Last.

As in the baseline model, the analyst is able to identify the agent's preference $\succ$ from the precedence order E^c of the choice sequence c. To illustrate how the analyst can construct the aversion phases p in the monotonic repetition aversion model (given by Definition 8), consider the case in which each of the alternatives is chosen infinitely many times. For alternative x, let $T(x,k)$ be the time period in which alternative x is chosen for the k-th time. Then for all $k \geq 1$, set

$$p(x,k) \;=\; \max\left(\{0\} \cup \{s - last^c(x,s) : x \succ c_s, 1 \leq s < T(x,k+1)\}\right) \tag{14}$$

In words, the aversion phase for x after the k-th time it is chosen is set to be the length of the maximal interval beginning with a period at which x is chosen and ending with a period s, before x is chosen for the $(k+1)$-th time, such that an alternative worse than x is chosen at s; and if x is the only alternative chosen during several first periods, we set $p(x,k) = 0$. Reflecting on formula (14) shows that $p(x,\cdot)$ is monotone and $(\succ, p)$ represents c, provided that c satisfies Worse Follows Faster Earlier and Abstention Last.

4.2 Repetition Aversion with Memory Horizon

A choice pattern that seems natural in the presence of repetition aversion is *intermittent consumption*: an alternative is repeatedly chosen in consecutive periods with regular interruptions, as for example in the sequence $c = (x, x, a, x, x, a, ...)$. Such patterns are not compatible either with the baseline model or with the model of the previous section. A parsimonious way to accommodate them in our framework is to assume that the consideration of alternative x depends on the history of its choice in the last two periods: if the last two choices were (a, a), (x, a), or (a, x), then x is considered, and if the last two choices were (x, x), x is not considered. We now generalise this idea.

Formally, given a choice sequence c and number $M \in \mathbb{N}$, define, for all $t \in \mathbb{N}$, the M-history $h_t^M(c) \in \tilde{X}^M$ as

$$h_t^M(c) = \begin{cases} (c_{t-1}, c_{t-2}, ..., c_{t-M}) & if\ t > M \\ (c_{t-1}, ..., c_1, a, ..., a) & if\ t \leq M \end{cases} \tag{15}$$

We interpret M as a memory horizon. For time periods exceeding this horizon, the M-history contains the history of choices in the last M periods preceding period t (written in reverse order); for the time periods $t \leq M$, we use an implicit assumption that before the start of observations, the agent consumed an outside option.

To ensure that repetition aversion is alternative-specific, we define for each x a binary encoding of its recent choice history. Define the function $\delta : \tilde{X} \times \tilde{X}^M \to \{0, 1\}$ given, for $(x, h) \in \tilde{X} \times \tilde{X}^M$, by

$$(\delta_x(h))_i = \mathbf{1}\{h_i = x\} \tag{16}$$

For example, suppose $c = (x, x, y, z, a, x, y, x, ...)$ and $M = 4$. Then $h_3^4(c) = (x, x, a, a)$, $h_9^4(c) = (x, y, x, a)$. Thus, $\delta_x(h_3^4(c)) = (1, 1, 0, 0)$, $\delta_y(h_3^4(c)) = \delta_z(h_3^4(c)) = (0, 0, 0, 0)$, $\delta_x(h_9^4(c)) = (1, 0, 1, 0)$, $\delta_y(h_9^4(c)) = (0, 1, 0, 0)$, and $\delta_z(h_9^4(c)) = (0, 0, 0, 0)$.

A natural dominance relation on M-histories (mapped into the vector in $\{0, 1\}^M$ by δ_x) is defined by one history dominating another if, in the former, x is chosen at least as recently and at least as often as in the latter. Formally, for $l, k \in \{0, 1\}^M$, we say that l *dominates* k, or $l \trianglerighteq k$, if for all $m = 1, ..., M$, $\sum_{i=1}^{m} l_i \geq \sum_{i=1}^{m} k_i$. Note that $\trianglerighteq$ is a transitive and antisymmetric relation.

For example, for $M = 4$, the vector $(1,0,1,0)$ dominates itself, and also vectors $(1,0,0,1)$, $(0,1,1,0)$, $(0,1,0,1)$, $(0,0,1,1)$, and all vectors that contain one or zero ones; it does not, however, dominate the vector $(0,1,1,1)$, and, vice versa, $(0,1,1,1)$ does not dominate $(1,0,1,0)$. Neither of the vectors $(1,0,0,1)$ and $(0,1,1,0)$ dominates the other.

We say that a numerical function f with domain $\{0,1\}^M$ is weakly decreasing with respect to $\trianglerighteq$ if $l \trianglerighteq k$ implies $f(l) \leq f(k)$. We are now ready to present the model.

Definition 11. A choice sequence c is *M-repetition averse* if there exists a pair $(\succ, R)$, where $\succ$ is a linear order on $\tilde{X}$, $x \succ a$ for all $x \in X$, and $R : \tilde{X} \times \{0,1\}^M \to \{0,1\}$ such that $R(a,h) = 1$ for all h, $R(x,(0,...0)) = 1$ for all x, and for all x, $R(x,\cdot)$ is weakly decreasing with respect to $\trianglerighteq$ such that

$$\{c_t\} = \max\left(\left\{x \in \tilde{X} : R(x, \delta_x(h_t^M(c))) = 1\right\}, \succ\right) \tag{17}$$

for every t. In this case, we say that $(\succ, R)$ represents c.

As usual, at each period t the agent chooses the most preferred considered alternative, where x is considered if $R(x, \delta_x(h_t^M(c))) = 1$ and x is not considered if $R(x, \delta_x(h_t^M(c))) = 0$. Since $R(a,h) = 1$ for all h, the outside option is always considered, and since $R(x,(0,...0)) = 1$ for all x, all alternatives are considered at least before the first time they are chosen. Monotonicity of R with respect to $\trianglerighteq$ formalises repetition aversion: more frequent and more recent past consumption weakly reduces consideration.

Suppose that $x \succ y$. Let x be chosen in period t, and y in period s. Then x is considered at t (after M-history $\delta_x(h_t^M(c))$) and not considered at s (after $\delta_x(h_s^M(c))$). Since consideration is monotonic in histories, it cannot be that x has been consumed at t as often and as recently as it was in the period preceding the choice of y, that is, it cannot be $\delta_x(h_t^M(c)) \trianglerighteq \delta_x(h_s^M(c))$. We show that this property and Abstention Last are also sufficient for a choice sequence c to be M-repetition averse.

Theorem 12. *A choice sequence c is M-repetition averse if and only if it satisfies Abstention Last, and for all s,t,*

$$xE^c y,\ c_t = x,\ c_s = y \implies \delta_x(h_t^M(c)) \ntrianglerighteq \delta_x(h_s^M(c)) \tag{18}$$

In the sufficiency proof we use the following construction: we first define a preference $\succ$ consistent with E^c, and then derive intuitive lower and upper bounds on feasible consideration consistent with observed choices and monotonicity. Property (18) ensures these bounds define a non empty interval that includes a monotone R. We then prove that $(\succ, R)$ thus constructed represents c.

Any baseline repetition aversion model $(\succ, p)$ for which $p(x) < \infty$ for all x is a special case of an M-repetition aversion model with $M \geq \max_{x \in X}(p(x))$. To see this, set

$$R(x, h) = \mathbf{1}\{h_i = 0 \; for \; all \; i \leq p(x)\} \tag{19}$$

One can see that $R(x, \cdot)$ is decreasing with respect to $\trianglerighteq$, and that $(\succ, R)$ represents the same choice as $(\succ, p)$. Moreover, one can derive the Worse Follows Faster property from eq. (18), since x being chosen in period t means that $R(x, \delta_x(h_t^M(c))) = 1$ and hence, $t - last(x, t) > p(x)$, while y being chosen in period s means that $R(x, \delta_x(h_s^M(c))) = 0$ and hence, $s - last(x, s) \leq p(x) < t - last(x, t)$.

5 Identification of Aversion Phases

In this section we discuss the extent to which the parameters of the model, and in particular aversion phases, can be inferred from the data. Denote by $X_0 = \{x \in \tilde{X} : x = c_t \text{ for some } t\}$ the set of alternatives chosen at least once. If there is an alternative that is never chosen, it can have an arbitrary aversion phase. Moreover, we cannot infer any information about the intrinsic order of such alternatives except that they should be worse than any alternative that is chosen in some period.[6] To avoid excessive notation, we will assume that each alternative, with the possible exception of the outside option, is chosen in some period; that is, $X \subseteq X_0$. All identification results extend straightforwardly to the case in which $X \backslash X_0 \neq \varnothing$ as discussed above.

We start with the baseline model from Section 2. Since the agent is not averse to any alternative in the initial period, the intrinsic preference $\succ$ is fully identified from the data. In Section 2 eq. (5), we construct an upper bound on the true aversion phase $p(x)$ for any alternative x; our next result provides also a lower bound on the aversion phase.

[6] In the model with initial aversion, there is one more possibility: the counter $w_1(x)$ of times that the alternative x has been chosen before the start of observations can be non-zero, and the corresponding aversion phase can be $+\infty$.

Theorem 13. *If a choice sequence c is repetition averse (according to Definition 1), then it is represented by $(\succ, p)$ if and only if $\succ = E^c$ and $\underline{p}(x) \leq p(x) \leq \overline{p}(x)$ for all $x \in X$, where*

$$\begin{aligned} \underline{p}(x) &= \max(\{t - last(x,t) : xE^c c_t\} \cup \{0\}) \\ \overline{p}(x) &= \min\{t - last(x,t) - 1 : x = c_t\} \end{aligned} \tag{20}$$

Note that if x is the best alternative according to $\succ$, then $\underline{p}(x) = \overline{p}(x)$. For the other alternatives, $\underline{p}(x) \leq \overline{p}(x)$ and in general, it can be that $\underline{p}(x) < \overline{p}(x)$.

Next, consider a repetition averse model from Section 3 with unknown initial aversion and menu variation, and let $\succ^*$ be the revealed preference relation according to eq. (9).

Theorem 14. *If a menu-choice sequence (M, c) is repetition averse (according to Definition 4), then it is represented by $(\succ, p, w_1)$ with some w_1 if and only if $\succ$ is a linear order such that $\succ^* \subseteq \succ$, and for all $x \in X$, $\underline{p}(x, \succ) \leq p(x) \leq \overline{p}(x)$, where*

$$\begin{aligned} \underline{p}(x,\succ) &= \max(\{0\} \cup \{\underline{w}_t(x) + 1 : x \in M_t \text{ and } x \succ c_t\}) \\ \overline{p}(x) &= \min\{\overline{w}_t(x) : x = c_t\} \end{aligned} \tag{21}$$

In Theorem 14, we remain silent about the multiplicity of w_1. It is not difficult to find the range of possible w_1 which, in general, depends on $\succ$ and p. However, we are primarily interested in identifying aversion phases. We therefore sacrifice some precision for the sake of simplicity.

In the model with initial aversion and menu variation, as in the baseline model, the upper bound on the aversion phase of an alternative does not depend on the choice of the linear order $\succ$. However, the lower bound on the aversion phase of an alternative, in general, depends on the choice of the linear order $\succ$. If for two preferences $\succ$ and $\succ'$, for any y, $x \succ y$ implies $x \succ' y$, then $\underline{p}(x, \succ) \geq \underline{p}(x, \succ')$. Thus, the range of possible aversion phases narrows when the alternative ascends in the preference order. The following simple example shows that the choice of $\succ$ does matter for identification of p.

Example 15. Let $X = \{x, y\}$, the sequence of menus be $M = (X, X, X, X, X, X, ...)$, and the sequence of choices be $c = (x, y, x, y, x, y, ...)$. We have $\underline{w}(x) = (0, 0, 1, 0, 1, 0, 1, 0, ...)$, $\overline{w}(x) = (\infty, 0, 1, 0, 1, 0, 1, 0, ...)$, $\underline{w}(y) = (0, 1, 0, 1, 0, 1, 0, 1, ...)$, $\overline{w}(y) = (\infty, \infty, 0, 1, 0, 1, 0, 1, ...)$. One can see that the revealed preference $\succ^*$ is empty, hence all linear orders are possible.

Next, $\overline{p}(x) = \overline{p}(y) = 1$. Consider $\succ$ such that $x \succ y$. Then $\underline{p}(x, \succ) = 1$ and $\underline{p}(y, \succ) = 0$. Thus, we must have $p(x) = 1$, $p(y) \in \{0, 1\}$, and $w_1(x) \geq 1$ in the representation in this case. Consider now $\succ'$ such that $y \succ' x$. Then $\underline{p}(y, \succ) = 1$, and $\underline{p}(x, \succ) = 0$. Thus, we must have $p(x) \in \{0, 1\}$, $p(y) = 1$, and $w_1(y) = 0$ in the representation in this case.

Consider the model of history dependent aversion analysed in Section 4.1. Analogous to identification results for previous models, we have the following theorem:

Theorem 16. *If a choice sequence c is monotonic repetition averse (according to Definition 8), then it is represented by $(\succ, p)$ if and only if $\succ = E^c$, for all $x \in X$, $p(x, \cdot)$ is weakly increasing, and for all $k \in \mathbb{Z}_+$, $\underline{p}(x, k) \leq p(x, k) \leq \overline{p}(x, k)$, where*

$$\underline{p}(x,k) \quad = \quad \max\left(\{0\} \cup \{t - last(x,t) : xE^c c_t,\ 1 \leq t < T(x, k+1)\}\right), \tag{22}$$

$$\overline{p}(x,k) = \begin{cases} T(x,k+1) - last(x, T(k+1, x)) - 1 & if\ x \text{ is chosen at least } k+1 \text{ times} \\ \infty & otherwise \end{cases}$$

and $T(x, k+1)$ denotes the period at which the alternative x is chosen in $(k+1)$-st time.

Finally, in the M-repetition aversion model, for each alternative that is chosen, the consideration function $R(x, \cdot)$ is partially identified via $\underline{R}(x, h) \leq R(x, h) \leq \overline{R}(x, h)$, where the lower and upper bounds are given by

$$\begin{aligned} \underline{R}(x,h) &= \mathbf{1}\left\{\exists t \in \mathbb{N} : c_t = x \text{ and } h_t^M(c) \trianglerighteq h\right\} \\ \overline{R}(x,h) &= 1 - \mathbf{1}\left\{\exists t \in \mathbb{N} : xE^c c_t \text{ and } h \trianglerighteq h_t^M(c)\right\} \end{aligned} \tag{23}$$

(these are the same bounds used in the proof of Theorem 12). In this case, even for the $\succ$-best alternative x, the consideration function $R(x, \cdot)$ might not be point identified, since in general, $\underline{R}(x, h) \neq \overline{R}(x, h)$, because the relation $\trianglerighteq$ is a partial, but not a linear order, and the choice sequence might miss some M-histories relevant for the identification of $R(x, \cdot)$.

To summarise, in all models with well-defined aversion phases—that is, in all models except the M-repetition aversion model, which involves a more complex consideration function—each alternative admits a range of aversion phases consistent with the data, independently of the aversion phases of other alternatives. We have characterised these ranges exactly for each model. The upper bounds are independent of intrinsic preference, whereas the lower bounds are not: higher-ranked alternatives in the intrinsic preference order correspond to larger lower bounds and therefore narrower ranges of admissible aversion phases. In particular, the lower and upper bounds coincide for the top-ranked alternative.

If the menu varies sufficiently across periods, the agent is exposed to a rich set of choice environments. In this case, the inferred lower bounds on aversion phases will approach, and may even coincide with, the corresponding upper bounds.

For example, assume that each period the menu is drawn i.i.d. according to a full-support distribution from all subsets of X. Then, with probability one, each alternative x appears alone for arbitrarily long stretches. The analyst can therefore identify its exact aversion phase as follows:

$$p(x) = \min\{t - last(x,t) - 1 : x = c_t\} = \min\{\#\text{periods between two choices of } x\}. \tag{24}$$

Moreover, for any pair of alternatives x and y, with probability one there is a period t such that neither alternative is subject to aversion and $M_t = \{x, y\}$. The choice c_t identifies the agent's intrinsic preference between alternatives x and y. The same reasoning applies when the menu-generating process exhibits serial correlation, provided that the conditional probability of drawing any menu—given any history—is bounded away from zero.

If the aversion phases are allowed to vary (an in particular to increase) over time, identification generally fails—even if the sequence of menus[7] is a realisation of the stochastic process described above. Indeed, the data on agent's choices after the $(k+1)$-st choice of alternative x can only help to establish an upper bound on the aversion phase $p(x,k)$ using the monotonicity of $p(x,\cdot)$, but not the exact value of $p(x,k)$.

[7]We consider a straightforward generalisation of the monotonic repetition aversion model to the case of varying menus and initial aversion; see Appendix B for the details.

6 Concluding Remarks

While most of the formal literature in economics on intertemporal choice dependence has concentrated on the "opposite" phenomena of habit formation or addiction, repetition aversion is psychologically intuitive, well documented and arises in a wide range of contexts, from food consumption to gaming.[8] Satiation also plays a central role in explanations of observed advertising strategies.[9] Sellers are concerned not only with a consumer's intrinsic valuation of a product but also with how satiation constrains frequent purchases. From a normative standpoint, it is equally important to identify environments—such as publicly funded projects—in which alternatives chosen relatively infrequently may nonetheless be intrinsically preferred to those selected more often.

The model of repetition aversion developed here admits a transparent characterisation in terms of observed choices and yields a clean separation between intrinsic preference and the repetition-aversion component. It also provides a novel foundation for limited consideration. To our knowledge, it is the first deterministic framework to incorporate intertemporal dependence in consideration formation. Although random consideration sets have been extensively studied in static environments, the deterministic dynamics of our model naturally induce a long-run distribution of consideration sets that admits a probabilistic interpretation.

We sketch how the model can serve as a foundation for stochastic choice and show that, in this extension, it lies outside the class of Random Utility Maximisation (RUM). Consider a setting with multiple menus, i.e. non-empty subsets of the grand set X on which $\succ$ is defined. Let $c_t(A)$ denote the choice from menu A at t. Assume that choice sequences are observed for each menu, and define the *choice frequency* $\varphi(x, A)$ of an alternative x as the long-run fraction of times that x is chosen. That is,

$$\varphi(x, A) = \lim_{n \to \infty} \frac{1}{n} \sum_{t=1}^{n} \mathbf{1}(c_t(A) = x) \tag{25}$$

This limit is well defined in the baseline model: a Markov process is induced on a finite state space — the set of vectors of outstanding aversion phases — and therefore

[8] E.g., Garbinsky *et al.*[7], Lei *et al.* [13]

[9] See Eliaz and Spiegler [5]

eventually enters a cycle. Long-run choice frequencies correspond to frequencies within that cycle.

We recall two benchmark properties of stochastic choice, namely Regularity and Luce Independence. Regularity requires that the introduction of new alternatives cannot increase the choice probabilities of existing ones and it is satisfied by Random Utility Models. Luce Independence states that the odds of choosing between any two alternatives are independent of the other alternatives in the menu and it is satisfied by Logit. Both properties are violated by the frequencies induced by our model (this result and the formal definition of the properties are presented in Appendix C). The resulting stochastic representation therefore lies strictly outside the RUM family, offering a novel perspective on violations of these axioms driven by intertemporal dependence. For example, introducing a new product may benefit some of the existing products, and may alter their relative market shares, by inducing repetition averse consumers to adopt a new purchasing routine.

Exploring the testable implications of this class of models appears to be a promising direction for future research. This is particularly true in richer extensions of the deterministic framework in which the repetition aversion associated with an alternative depends not only on its own choice history but also on that of other alternatives. Repetition aversion may, for instance, extend across similar products. Speculatively, in the corresponding stochastic version of the model, such interdependence could generate correlation structures akin to those captured in a static setting by nested stochastic choice specifications.[10]

The analysis of "optimal menus" in section 3 and the unusual properties of choice frequencies noted above provide another promising avenue for future research. We have highlighted the structural fragility of choice behaviour with repetition aversion and the possibility of inducing substantial behavioural changes by manipulating just a few initial menus. Our examples are deliberately simplified. In practice, many additional factors may matter, including consumer heterogeneity and information about individual characteristics such as preferences, aversion phases, and initial aversion. This opens the door to a more systematic investigation of menu design.

[10]See Kovach and Tserenjigmid [12] for an axiomatic treatment of these models, and the literature therein.

References

[1] Agranov, Marina and Pietro Ortoleva (2017) "Stochastic Choice and Preferences for Randomization", *Journal of Political Economy* 125(1): 40-68.

[2] Ashkenazi-Golan, Galit, Dominik Karos and Ehud Lehrer (2024) "A Taste for Variety", *arXiv*:2405.00561 [econ.TH].

[3] Baucells, M. and R. K. Sarin (2010) "Predicting utility under satiation and habit formation" *Management Science* 56: 286–30.

[4] Becker, Gary S. and Kevin M. Murphy (1988) "A Theory of Rational Addiction". *Journal of Political Economy* 96(4): 675-700.

[5] Eliaz, K. and R. Spiegler (2020) "Incentive-compatible advertising on nonretail platforms", *The Rand Journal of Economics* 51(2): 323-345

[6] Ferreira, Joao V. and Nicolas Gravel (2025) "Revealing preference discovery: a chronological choice framework", *Theory and Decision* 98: 1-32

[7] Garbinsky, Emily N., Carey K. Morewedge, and Baba Shiv (2014) "Does liking or wanting determine repeat consumption delay?", *Appetite* 72: 59-65.

[8] Gilboa, Itzhak and David Schmeidler (1995) "Case-based Decision Theory". *The Quarterly Journal of Economics* 110 (3): 605–639.

[9] Gilboa, Itzhak and David Schmeidler (1997) "Cumulative Utility Consumer Theory", *International Economic Review* 38(4): 737-761.

[10] He, Y., J. S. Dyer, and J. C. Butler (2013) "On the Axiomatization of the Satiation and Habit Formation Utility Models". *Operations Research* 61: 1399–1410.

[11] Kaiser, Karen and Rainer Schwabe (2012) "Preference for Variety", *The B.E. Journal of Theoretical Economics* 12(1).

[12] Kovach, Matthew and Gerelt Tserenjigmid (2022) "Behavioral Foundations of Nested Stochastic Choice and Nested Logit", *Journal of Political Economy* 130(9): 2411-2461.

[13] Lei, Xiao, Beichen Wan, and Shixin Wang (2023) "Content rotation in the presence of satiation effects". SSRN 4593945, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4593945

[14] Lu, Jay and Kota Saito (2020) "Repeated choice: a theory of stochastic intertemporal preferences", Working paper, Social Science Working Paper, 1449. California Institute of Technology.

[15] McAlister, Leigh (1982) "A Dynamic Attribute Satiation Model of Variety-Seeking Behavior ", *Journal of Consumer Research* 9(2): 141–150.

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

[17] Masatlioglu, Yusufcan, Daisuke Nakajima and Erkut Y. Ozbay (2012) "Revealed Attention", *American Economic Review* 102(5): 2183-2205.

[18] Pollak, Robert A. (1970): "Habit formation and dynamic demand functions", *Journal of Political Economy* (78): 745–763.

[19] Puig, David (2024) "Taste for Variety: An Intertemporal Choice Model", Unpublished manuscript.

[20] Rozen, Kareen (2010) "Foundations of Intrinsic Habit Formation", *Econometrica* 78(4): 1341-1373.

[21] Turansick, Christopher (2024) "Consumption dependent random utility", *arXiv*:2412.05344, 2024

Appendix A: Proofs

Proof of Theorem 3

(i) $\Longrightarrow$ (ii) Let c be a repetition averse choice sequence represented by $(\succ, p)$. To see that Worse Follows Faster holds, suppose that $x = c_t$, $y = c_s$ and xEy. Then $x \succ y$ or $y = a$. In either case, x cannot be considered at s and is considered at t. Hence $s - last(x,s) \leq p(x) < t - last(x,t)$, proving Worse Follows Faster. To see that Abstention Last holds, note that a can only be chosen if all alternatives are not considered, which means that they have been chosen before.

(ii) $\Longrightarrow$ (i) Suppose that c satisfies Worse Follows Faster and Abstention Last. Denote $X_0 = \{x \in \tilde{X} : x = c_t \text{ for some } t\}$. Since E is a linear order, and xEa for all $x \in X_0 \backslash \{a\}$, E can be extended to a linear order $\succ$ on $\tilde{X}$ that satisfies $x \succ a$ for all $x \in X$. Next, for each $x \in X$, let $p(x) = 0$ for all $x \in X \backslash X_0$, $p(a) = 0$, and for all $x \in X_0 \backslash \{a\}$, pick $p(x)$ that satisfies $\underline{p}(x) \leq p(x) \leq \overline{p}(x)$, where

$$\begin{aligned}\underline{p}(x) &= \max(\{t - last(x,t) : xE^c c_t\} \cup \{0\}) \\ \overline{p}(x) &= \min\{t - last(x,t) - 1 : x = c_t\}\end{aligned}$$

Let us show that for any $x \in X_0 \backslash \{a\}$, such p exists; that is, that $\underline{p}(x) \leq \overline{p}(x)$. If $\{t - last(x,t) : xE^c c_t\} = \varnothing$, then $\underline{p}(x) = 0 \leq \overline{p}(x)$. Otherwise, towards a contradiction, assume that $\underline{p}(x) > \overline{p}(x)$, then there are s, t such that $c_t = x$, $xE^c c_s$ and $t - last(x,t) - 1 = \overline{p}(x) < \underline{p}(x) = s - last(x,s)$, and hence, $t - last(x,t) \leq s - last(x,s)$, in contradiction to Worse Follows Faster.

The model $(\succ, p)$ induces a choice d. We are going to show that $d_t = c_t$ for all t by induction in $t = 1, 2, ...$ It is without loss to consider $X \backslash X_0 = \varnothing$. Indeed, if we show that $c = d$ in this case, then in case if $X \backslash X_0 = \varnothing$, any $x \in X \backslash X_0$ does not appear neither in c (by construction) nor in d (since in any period, there is always $y \in X_0$ considered by d, and $y \succ x$ for any $x \in X \backslash X_0$).

Since $\succ = E^c$, then $d_t = c_t$ for $t = 1$, proving the induction base. For $t > 1$, if $x = c_t$ was not chosen before, then it is considered by d. Otherwise,

$$p(x) \leq \overline{p}(x) = \min\{s - last^c(x,s) - 1 : x = c_s\} \leq \tag{26}$$

$$\leq t - last^c(x,t) - 1 = t - last^d(x,t) - 1$$

where we used $last^d(x,t) = last^c(x,t)$ since both quantities depend only on the history of choices prior to period t, which are identical by the induction assumption. Thus, x is considered by d in period t.

Consider any alternative y such that $y \succ x = c_t$, if such alternative exists. We have

$$p(y) \geq \underline{p}(y) = \max(\{s - last^c(x,s) : yE^c c_s\} \cup \{0\}) \geq t - last^c(t,y) = t - last^d(t,y) \quad (27)$$

Thus, y is not considered by d in period t, and $d_t = x = c_t$, proving the induction step.

(ii) $\iff$ (iii). We show that a choice sequence satisfies Worse Follows Faster if and only if it satisfies Asymmetry. Let c be a choice sequence which satisfies Worse Follows Faster but fails Asymmetry: $x \succ^c y$ and $y \succ^c x$ for some x and y. By definition of $\succ^c$ both x and y appear in c. It follows that either $xE^c y$ or $yE^c x$, which means, by Worse Follows Faster and the definition of $\succ^c$, that either $y \not\succ^c x$ or $x \not\succ^c y$, a contradiction. In the other direction, suppose that $\succ^c$ is asymmetric, and $xE^c y$. Let t be the period in which x is chosen for the first time, and $s > t$ be the period in which y is chosen for the first time. Then $t - last(y,t) = t - (-\infty) = \infty$, and $s - last(y,s) = s - (-\infty) = \infty$. Using the convention $\infty \geq \infty$, we get $t - last(y,t) \geq s - last(y,s)$ and hence, $x \succ^c y$. Thus, $y \not\succ^c x$ since $\succ^c$ is asymmetric, and we conclude that c satisfies Worse Follows Faster. $\square$

Proof of Theorem 5

Let (M,c) be a repetition aversion menu-choice sequence represented by $(\succ, p, w_1)$, and let $\succ^*$ be its associated revealed preference. Suppose that $x \succ^* y$. Hence, there are periods t,s such that $c_s = x$, $y \in M_s$, $c_t = y$, and $\underline{w}_s(y) \geq \overline{w}_t(y)$. It follows that $w_s(y) \geq w_t(y) \geq p(y)$. Therefore, y is considered at s and hence, $x \succ y$. Since $\succ$ is acyclic, and $x \succ a$ for all $x \in X$, we conclude that $\succ^*$ is acyclic, and for all $x \in X$, not $a \succ^* x$.

Conversely, let $\succ^*$ be the revealed preference associated with the menu-choice sequence (M,c), let $\succ^*$ be acyclic and for all $x \in X$, not $a \succ^* x$. Denote $X_0 = \{x \in X : \exists t : c_t = x\}$. Consider a tuple $(\succ, p, w_1)$, where $\succ$ is some completion of $\succ^*$ to a linear order on $X \cup \{a\}$ such that $x \succ a$ for all $x \in X$, $p(a) = 0$, $w_1(a) = 0$, $\underline{p}(x) \leq p(x) \leq \overline{p}(x)$, where

$$\begin{aligned} \underline{p}(x, \succ) &= \max(\{0\} \cup \{\underline{w}_t(x) + 1 : x \in M_t \textit{ and } x \succ c_t\}) \\ \overline{p}(x) &= \min\{\overline{w}_t(x) : x = c_t\}, \end{aligned} \tag{28}$$

and

$$w_1(x) = \max\{0, p(x) - \min\{t : c_t = x\} + 1\} \tag{29}$$

for any $x \in X_0$; finally, $p(x) = \infty$, $w_1(x) = 0$ for any $x \in X \backslash X_0$. Note that $\succ$ exists given the properties of $\succ^*$.

Let us verify that there exists p that satisfies the condition given above. It suffices to show that $\underline{p}(x, \succ) \leq \overline{p}(x)$ for all $x \in X_0$. Suppose not, then $\underline{p}(x, \succ) > 0$, and there are periods t, s such that $x \in M_s$, $x \succ c_s$, $x = c_t$, and

$$\underline{w}_s(x) + 1 = \underline{p}(x, \succ) > \overline{p}(x) = \overline{w}_t(x) \quad \Longrightarrow \quad \underline{w}_s(x) \geq \overline{w}_t(x) \tag{30}$$

It follows that $c_s \succ^* x$, in contradiction to $x \succ c_s$.

Let d be the sequence of choices represented by $(\succ, p, w_1)$. We are going to show that $d_t = c_t$ for all t by induction in $t \in \mathbb{N}$. It is without loss to consider $X \backslash X_0 = \varnothing$. Indeed, if $c = d$ in this case, then in case if $X \backslash X_0 = \varnothing$, any $x \in X \backslash X_0$ does not appear neither in c (by construction) nor in d (since in any period, there is always $y \in X_0$ that is available and considered by d, and $y \succ x$ for any $x \in X \backslash X_0$).

For the induction base $t = 1$, suppose $c_1 = x \in X$. Then $\min\{t : c_t = x\} = 1$ and $w_1(x) = \max\{0, p(x)\} \geq p(x)$. Thus, x is considered by d at $t = 1$. If $c_1 = a$, then a is considered by d at $t = 1$ as well. For any other alternative $y \neq c_1, a$, if $y \succ x$, then

$$p(y) \geq \underline{p}(y, \succ) = \max(\{0\} \cup \{\underline{w}_t(x) + 1 : x \in M_t \textit{ and } x \succ c_t\}) \geq \underline{w}_1(y) + 1 \geq 1 \tag{31}$$

We also have $\min\{t : c_t = y\} > 1$ and, hence, $w_1(y) < p(y)$, meaning that y is not considered by d at $t = 1$. Thus, $d_1 = c_1$, proving the induction base.

For the induction step, assume that $d_r = c_r$ for all $r < t$. Let $x = c_t$. If $x = a$, then x is considered by d at t. Consider now $x \in X$. Suppose x appears in c in some period $s < t$. Without loss, $c_r = d_r \neq x$ for all $s < r < t$. Hence, $w_t(x) = \overline{w}_t(x) = \overline{p}(x) \geq p(x)$. In the remaining case, t is the first period when x appears in c. Then $w_1(x) \geq p(x) - t + 1$ and, hence, $w_t(x) \geq p(x) - t + 1 + (t - 1) = p(x)$. Thus, x is considered by d in period t.

Consider $y \in M_t$ such that $y \succ x$. Suppose first that there is $s < t$ such that $c_s = y$. Then $\underline{w}_t(y) = w_t(y) = \overline{w}_t(y)$ and

$$p(y) \geq \underline{p}(y) \geq \underline{w}_t(y) + 1 = w_t(y) + 1 \tag{32}$$

Thus, $w_t(y) \leq p(y) - 1 < p(y)$ and hence, y is not considered by d in period t.

Suppose now that y does not appear in c, and, therefore, in d before period t. In this case, $p(y) \geq \underline{p}(y) \geq \underline{w}_t(y) + 1 = (t-1) + 1 = t$. If $w_1(y) = 0$, then $w_t(y) = t - 1 < p(y)$ and hence, y is not considered in period t. Otherwise, $w_1(y) = p(y) - \min\{s : c_s = y\} + 1$. Since $\min\{s : c_s = y\} > t$, then $w_1(y) < p(y) - t + 1$ and thus, $w_t(y) = w_1(y) + (t-1) < p(y)$. In this case, y is not considered by d in period t as well. This shows that $d_t = x = c_t$, proving the induction step. □

Proof of Theorem 10

We will show implications (i) $\Longrightarrow$ (ii) $\Longrightarrow$ (iii) $\Longrightarrow$ (i).

To prove (i) $\Longrightarrow$ (ii), let $(\succ, p)$ be a monotonic repetition averse representation of c. Construct a general monotonic averse model $(\succ, P)$ as follows: $P(x,t) = p(x, n_t^c(x))$. Note that since $n_t^c(x)$ is weakly increasing with respect to t, and $p(x,n)$ is weakly increasing with respect to n, then $P(x,t)$ is weakly increasing with respect to t. Finally,

$$\max\left(\{x \in X : t - last^c(x,t) - 1 \geq P(x,t)\}, \succ\right) = \tag{33}$$

$$= \max\left(\{x \in X : t - last^c\,(x,t) - 1 \geq p(x, n_t^c(x))\}, \succ\right) = \{c_t\}$$

proving that $(\succ, P)$ is a general monotonic repetition averse representation of c.

To prove (ii) $\Longrightarrow$ (iii), let c have a general monotonic repetition aversion representation $(\succ, P)$. Since $P(x,0) = 0$ for all x, $xEy \iff x \succ y$ for all $x, y \in A$, and also aEx for no $x \in X$, hence Abstention Last holds. Suppose now $t > s$, $c_t E c_s$. Then $c_t \succ c_s$, and $P(x,t) \geq P(x,s)$. Since c_t is not chosen at period s and it is chosen in period t, $t - last^c(c_t,t) - 1 \geq P(x,t) \geq P(x,s) \geq s - last^c(c_t,s) - 1$, proving that Worse Follows Faster Earlier holds.

To prove (iii) $\Longrightarrow$ (i), suppose that c satisfies Worse Follows Faster Earlier and Abstention Last. Construct a representation $(\succ, p)$ as follows. Let $X_0 = \{x \in X : x =$

c_t for some $t\}$. Define $x \succ y \iff xEy$ for all $x, y \in X_0$, $x \succ y$ for any $x \in X_0$, $y \in X \backslash X_0$, and $x \succ a$ for all $x \in X$. Define $p(x, 0) = 0$ for all x. Pick $p(x, \cdot)$ as an arbitrary weakly increasing function for all $x \in X \backslash X_0$. Next, for $x \in X_0$ and $k > 0$, pick $p(x, k)$ that is weakly increasing in the second argument and satisfies $\underline{p}(x, k) \leq p(x, k) \leq \overline{p}(x, k)$, where

$$\underline{p}(x,k) \quad = \quad \max\left(\{0\} \cup \{t - last(x,t) : xE^c c_t,\ 1 \leq t < T(x, k+1)\}\right) \tag{34}$$

$$\overline{p}(x,k) = \begin{cases} T(x,k+1) - last(x, T(x,k+1)) - 1 & \textit{if } x \textit{ is chosen at least } k+1 \textit{ times} \\ \infty & \textit{otherwise} \end{cases}$$

and $T(x, k+1)$ is the time period when x is chosen for the $(k+1)$-th time with the convention that $T(x, k+1) = \infty$ if x is not chosen $k+1$ times.

Let us verify that there exist p that satisfies the conditions above. In particular, consider $p(x, k) = \underline{p}(k, x)$. Since $m < k$ implies $\{t - last(x,t) : xE^c c_t,\ 1 \leq t < T(x, m+1)\} \subseteq \{t - last(x,t) : xE^c c_t,\ 1 \leq t < T(x, k+1)\}$, the function $\underline{p}(x, \cdot)$ is weakly increasing. It remains to show that $\underline{p}(x,k) \leq \overline{p}(x,k)$ for all $x \in X_0$. We only need to check the case when x is chosen at least $k+1$ times, and the set $\{t - last(x,t) : xE^c c_t,\ 1 \leq t < T(x, k+1)\}$ is not empty. Towards a contradiction, assume $\underline{p}(x,k) > \overline{p}(x,k)$, then there is a period t such that $xE^c c_t$, $1 \leq t < T(x, k+1)$, and $t - last(x,t) > T(x,k+1) - last(x, T(x, k+1)) - 1$. Then $t - last(x,t) \geq T(x,k+1) - last(x, T(x,k+1))$, $c_{T(x,k+1)} = x$, $xE^c c_t$, and $t < T(x, k+1)$, in contradiction to Worse Follows Faster Earlier axiom.

We show that $(\succ, p)$ represents c. First, let us verify that for any t, $x \succ c_t$ implies $t - last^c(x,t) - 1 \geq p(x, n_t^c(x))$. It suffices to consider $x \in X_0$, since if $X_0 \neq X$, then by Abstention Last there is no t such that $c_t = a$, and hence $c_t \in X_0$ for all t; $x \succ c_t$ and $x \notin X_0$ is impossible in this case. If x is not chosen before period t, $x \succ c_t$ is impossible since $c_t Ex \implies c_t \succ x$. It remains to consider $n_t^c(x) \geq 1$ and $last^c(t, x) \geq 1$. In this case, either $p(x, n_t^c(x)) = \infty$ and $t - last^c(x,t) \leq p(x, n_t^c(x))$, or there exists $T(x, n_t^c(x)) > t$, and

$$p(x, n_t^c(x)) \geq \underline{p}(x, n_t^c(x)) \geq t - last^c(x, n_t^c(x)) \tag{35}$$

as desired, where we used the fact that $x \succ c_t$ implies $c_t \neq x$ and hence, $\{s - last^c(x,s) : x \succ c_s, 1 \leq s < T(x, n_t^c(x) + 1)\} \neq \varnothing$.

Finally, consider $c_t = x \in X$; denote k to be the number of times x is chosen before periods t. If $k = 0$, then $p(x,k) = p(x,0) = 0 \leq t - last(x,t) - 1$. Otherwise, $p(x,k) \leq \overline{p}(x,k) = t - last(x,t) - 1$. In both cases, x is considered by c in period t; thus, $d_t = c_t$ $\square$

Proof of Theorem 12

Suppose first that c has an M-repetition aversion representation $(\succ, R)$. If t is the first period in which alternative x is chosen, and $s \leq t$, then $\delta_x(h_s^M(c)) = (0,...,0)$ and hence, $R(x, \delta_x(h_s^M(c))) = R(x,(0,...,0)) = 1$. Thus, every alternative is considered at least before its first choice; Abstention Last must, therefore, hold. Moreover, $xE^c y$ implies $x \succ y$. Suppose now that $c_t = x$, $c_s = y$, and $xE^c y$. It follows $R(x, \delta_x(h_s^M(c))) = 0$ and $R(x, \delta_x(h_t^M(c))) = 1$. Since $R(x, \cdot)$ is weakly decreasing with respect to $\trianglerighteq$, then $\delta_x(h_t^M(c)) \not\trianglerighteq \delta_x(h_s^M(c))$.

Suppose now that c satisfies Abstention Last, and xEy, $c_t = x$, $c_s = y$ implies $\delta_x(h_t^M(c)) \not\trianglerighteq \delta_x(h_s^M(c))$. Pick $\succ$ that extends E to a linear order on $\tilde{X}$ with $x \succ a$ for all $x \in X$. Next, for $x \in X \backslash X_0$, set $R(x,h) = 1$ for all h, and also $R(a,h) = 1$ for all h. Finally, for $x \in X_0 \backslash \{a\}$, set $\underline{R}(x,h) \leq R(x,h) \leq \overline{R}(x,h)$ for all h, where

$$\underline{R}(x,h) = \mathbf{1}\left\{\exists t \in \mathbb{N} : c_t = x \textit{ and } h_t^M(c) \trianglerighteq h\right\}$$

$$\overline{R}(x,h) = 1 - \mathbf{1}\left\{\exists t \in \mathbb{N} : xE^c c_t \textit{ and } h \trianglerighteq h_t^M(c)\right\}$$

Let us first show that such $R(x,\cdot)$ exists. It suffices to show that for all $x \in X_0 \backslash a$, $\underline{R}(x,h) \leq \overline{R}(x,h)$ for all h. Assume this does not hold for some x and h, then there are periods t, s such that $c_t = x$, xEc_s, and $h_t^M(c) \trianglerighteq h$, $h \trianglerighteq h_s^M(c)$. Since $\trianglerighteq$ is transitive, $h_t^M(c) \trianglerighteq h_s^M(c)$, in contradiction.

Let d be the choice induced by the M-repetition aversion model $(\succ, R)$. We are going to show that $d_t = c_t$ by induction in t. As with the proofs of the other representation theorems, it suffices to consider the case when $X \subseteq X_0$. Since $R(x,(0,...,0)) = 1$ for all x, c_1 is considered at $t = 1$ by d and chosen (since $\succ = E$). Let us prove the induction step. By the induction assumption, $h_t^M(d) = h_t^M(c)$. Since c_t is chosen by c at t, then

$$R(x, \delta_x(h_t^M(d))) = R(x, \delta_x(h_t^M(c))) \geq \underline{R}(x, \delta_x(h_t^M(c))) = 1$$

Thus, c_t is considered by d in period t. Consider any $y \succ c_t$, then $yE^c c_t$, and

$$R(y, \delta_y(h_t^M(d))) = R(y, \delta_y(h_t^M(c))) \leq \overline{R}(y, \delta_y(h_t^M(c))) = 0$$

Thus, y is not considered by d in period t. Therefore, c_t is chosen by d in period t. $\square$

Proof of Theorem 13

Let c be repetition averse and $(\succ, p)$ be its representation. Clearly, it must be that $\succ = E^c$, since before the first choice of an alternative x, $last(x,t) = -\infty$. Suppose that $p(x) < \underline{p}(x) = \max(\{t - last(x,t) : xE^c c_t\} \cup \{0\})$ for some x. Then $\underline{p}(x) > 0$ and hence, there is a period t such that $p(x) < \underline{p}(x) = t - last(x,t)$, and, hence, $p(x) \leq t - last(x,t) - 1$, and also $x \succ c_t$, in contradiction to $\{c_t\} = \max\left(\left\{x \in \tilde{X} : t - last(x,t) - 1 \geq p(x)\right\}, \succ\right)$. Suppose that $p(x) > \overline{p}(x) = \min\{t - last(x,t) - 1 : x = c_t\}$. Then, there is a time period t such that $p(x) > \overline{p}(x) = t - last(x,t) - 1$ and $c_t = x$, in contradiction to $c_t = x$. This proves the "only if" part of the theorem.

The "if" part of the theorem is already proven in the proof of Theorem 3. $\square$

Proof of Theorem 14

Let (M, c) be a menu-choice sequence represented by $(\succ, p, w_1)$. Towards a contradiction, assume $\succ^* \not\subseteq \succ$, then there are alternatives $x, y \in \tilde{X}$ such that $x \succ^* y$ and $y \succ x$. Thus, there are periods s, t such that $c_t = y$, $y \in M_s$, $c_s = x$, and $w_s(y) \geq \underline{w}_s(y) \geq \overline{w}_t(y) \geq w_t(y) \geq p(y)$. This is impossible, since y is considered by c in period s, but a $\succ$-worse alternative x is chosen. Now, suppose that $p(x) > \overline{p}(x) = \min\{\overline{w}_t(x) : x = c_t\}$ for some x. Then, there is a period t such that $p(x) > \min\{\overline{w}_s(x) : x = c_s\} = \overline{w}_t(x) \geq w_t(x)$, and $c_t = x$, in contradiction. Suppose now that $p(x) < \underline{p}(x) = \max(\{0\} \cup \{\underline{w}_t(x) + 1 : x \in M_t \text{ and } x \succ c_t\})$ for some x. Then $\underline{p}(x) > 0$, and there is a period t such that $p(x) < \underline{w}_t(x) + 1 \leq w_t(x) + 1$, $x \in M_t$, and $x \succ c_t$. Thus, x is considered by c in period t, but $c_t \succ x$, in contradiction. This proves the "only if" part of the theorem.

The "if" part of the theorem is already proven in the proof of Theorem 5. $\square$

Proof of Theorem 16

Let c be a choice sequence that has a monotonic repetition averse representation $(\succ, p)$. Clearly, it must be that $\succ = E^c$, since $p(x, 0) = 0$ for all x, and before the first choice of an alternative x, $last(x, t) = -\infty$. Towards a contradiction, assume that $p(x, k) < \underline{p}(x, k)$, then $\underline{p}(x, k) > 0$ and hence, there is a period t such that $1 \leq t < T(x, k+1)\}$, $p(x, k) < \underline{p}(x, k) = t - last(x, t)$, and hence, $p(x, k) \leq t - last(x, t) - 1$, contradicting $x \succ c_t$. Towards a contradiction, assume that $p(x, k) > \overline{p}(x, k)$, then x is chosen at least $k+1$ times, and $p(x, k) > T(x, k+1) - last(x, T(x, k+1)) - 1$, in contradiction to $c_{T(x,k+1)} = x$.

The "if" part of the theorem is already proven in the proof of Theorem 10. □

Appendix B: Extensions

Variants of a model with fixed aversion phases

The model of repetition aversion considered in Section 3 extends the baseline model considered in Section 2 in two ways: allowing the agent to be averse to some alternatives at the start of observations, and allowing for menu to change from period to period. In this section, we complement the analysis done in Sections 2 and 3 by considering a baseline model with each of these extensions in turn.

If the agent may have the initial aversion, but the menu is fixed, a special case of the model from Section 3 applies. The characterisation and identification theorems repeat Theorems 5 and 14 with the only difference that the menu is fixed and hence, we can drop the requirement $y \in M_s$ from the definition of revealed preference given by eq. (9):

$$x \succ^* y \iff \exists t, s \in \mathbb{N} \text{ such that } c_t = y,\ c_s = x, \text{ and } \underline{w}_s(y) \geq \overline{w}_t(y) \tag{36}$$

Corollary 17. *If menu $M_t = X$ is fixed, but the agent may have initial aversion, a choice sequence c has a repetition averse representation $(\succ, p, w_1)$ if and only if the revealed preference relation $\succ^*$ given by eq. (36) is acyclic and $a \succ^* x$ for no $x \in X$. If, additionally, all alternatives are chosen, then $(\succ, p, w_1)$ with some w_1 represents c if and only if $\succ$ is a linear order such that $\succ^* \subseteq \succ$, and* *for all* $x \in X$, $\underline{p}(x, \succ) \leq p(x) \leq \overline{p}(x)$, *where*

$$\begin{aligned} \underline{p}(x, \succ) &= \max(\{0\} \cup \{\underline{w}_t(x) + 1 : x \succ c_t\}) \\ \overline{p}(x) &= \min\{\overline{w}_t(x) : x = c_t\} \end{aligned} \tag{37}$$

Similarly, if the menu can vary, but there is no initial aversion, Theorems 5 and 14 work as well, but this time, $w_1(x) = \infty$ for all x and hence, we need to use $\overline{w}_t(x)$ everywhere. Alternatively, we can use $\overline{w}_t(x) = t - last(x, t) - 1$ and re-write the definition of the revealed preference as follows:

$$x \succ^* y \iff \exists t, s \in \mathbb{N} : \; c_t = y, \; y \in M_s, \; c_s = x, \; and \; s - last(y, s) \geq t - last(y, t) \tag{38}$$

We can now re-write Theorems 5 and 14 in notations "$last(x, t)$" instead of "$w_t(x)$:"

Corollary 18. *If the agent does not have initial aversion, but the menu may vary, a menu-choice sequence* (M, c) *has a repetition averse representation* $(\succ, p)$ *if and only if the revealed preference relation* $\succ^*$ *given by eq. (38) is acyclic and* $a \succ^* x$ *for no* $x \in X$*. If, additionally, all alternatives are chosen, then* $(\succ, p)$ *represents* c *if and only if* $\succ$ *is a linear order such that* $\succ^* \subseteq \succ$*, and for all* $x \in X$, $\underline{p}(x, \succ) \leq p(x) \leq \overline{p}(x)$, *where*

$$\begin{aligned} \underline{p}(x, \succ) &= \max(\{0\} \cup \{t - last(x, t) : x \in M_t \text{ and } x \succ c_t\}) \\ \overline{p}(x) &= \min\{t - last(x, t) - 1 : x = c_t\} \end{aligned} \tag{39}$$

Even without the initial aversion, the bounds on the aversion phases may depend on the choice of the linear order $\succ$. We illustrate it with a slight modification of Example 15 to make use of the variation of the menu over time:

Example 19. Let $X = \{x, y\}$, the sequence of menus be $M = (\{x\}, X, X, X, X, X, ...)$, and the sequence of choices be $c = (x, y, x, y, x, y, ...)$. In this case, the revealed preference $\succ^*$ given by eq. (38) is empty. Therefore, all linear orders on X are consistent with the data. Consider first the order $x \succ y$. Then $\underline{p}(x, \succ) = \overline{p}(x) = 1$. Accordingly, we must have $p(x) = 1$. On the other hand, since $y \not\succ x$, we have $\underline{p}(y, \succ) = 0$, and $\overline{p}(y) = 1$. Hence, both $p(y) = 0$ and $p(y) = 1$ are possible. Consider now the order $y \succ' x$. In this case, only $p(y) = 1$ is consistent with the data; however, $\underline{p}(x) = 0$ and one can see that we can have both $p(x) = 0$ and $p(x) = 1$ in the representation.

History dependent aversion: initial aversion and menu variation

Let us now consider the most general model that combine monotonic history dependent aversion, menu variation, and unknown initial conditions. In this case, there is no need to introduce an additional vector of initial conditions w_1; instead, we allow an aversion phase before the first encounter of an alternative to be non-zero. A seemingly more general class of monotonic history dependent aversion models with arbitrary w_1 has, in fact, the same empirical content.[11]

Definition 20. A menu-choice sequence (M, c) is *monotonic repetition averse* if there is a pair $(\succ, p)$, where $\succ$ is a linear order on $X \cup \{a\}$ such that $x \succ a$ for all $x \in X$, $p : \tilde{X} \times Z_+ \to \tilde{\mathbb{Z}}_+$ is non-decreasing with respect to the second argument for all x, and $p(a,k) = 0$ for all k, such that for every t,

$$\{c_t\} = \max\left(\{x \in M_t : w_t(x) \geq p(x, n_t^c(x))\}, \succ\right) \tag{40}$$

where

$$n_t^c(x) = \left|\{s \in \{1, ..., t-1\} : c_s = x\}\right|, \tag{41}$$

$w_1 = 0$, and w_t for $t > 1$ is defined recursively via

$$w_{t+1}(x) = \begin{cases} w_t + 1 & if\ c_t \neq x \\ 0 & if\ c_t = x \end{cases} \tag{42}$$

In this case, we say that $(\succ, p)$ represents c.

Note that w_t is now observable from the data, in contrast to w_t in the model from Section 3, when the aversion phase has been fixed, but initial values of w have been allowed to take arbitrary values. Define the following revealed preference relation:

$$x \succ^{**} y \iff \exists t, s \in \mathbb{N}\ such\ that\ t > s,\ c_t = y,\ y \in M_s,\ c_s = x,\ and\ w_s(y) \geq w_t(y) \tag{43}$$

[11] To see this, let $(\succ, p, w_1)$ be a natural generalization of the model considered in Definition 20 that allow for the possibility to have $w_1(x) > 0$. Consider a monotonic repetition aversion model $(\hat{\succ}, \hat{p})$ according to Definition 20, where $\hat{\succ} = \succ$, and for all x, $\hat{p}(x,0) = p(x,0) - w_1(x)$, $\hat{p}(x,k) = p(x,k)$ for all $k > 0$. Then $\hat{p}(x,0) \leq p(x,0) \leq p(x,1) = \hat{p}(x,1)$, so $\hat{p}$ is weakly increasing with respect to the second argument for every x. Also, $(\hat{\succ}, \hat{p})$ represents the same menu-choice as model $(\succ, p, w_1)$.

The revealed preference $\succ^{**}$ given by eq. (43) differs from the revealed preference $\succ^{*}$ given by eq. (9) in two ways: first, by the additional requirement that $t > s$, and second, by replacing $\underline{w}_s(y)$ with $w_s(y)$ and $\overline{w}_t(y)$ with $w_t(y)$. The latter is of no consequence[12]. To understand the role of the requirement $t > s$, note that in the monotonic repetition averse model, $w_s(y) \geq w_t(y)$ implies that y is considered in period s only if $p(y, n_s(y)) \leq p(y, n_t(y))$, which is guaranteed only if $t > s$. The characterisation and indentification theorem for this model has a familiar form:

Theorem 21. *A menu-choice sequence (M, c) is monotonic repetition averse if and only if $\succ^{**}$ is acyclic and $a \succ^{**} x$ for no $x \in X$. If, additionally, all alternatives are chosen, then $(\succ, p)$ represents (M, c) if and only if $\succ$ is a linear order such that $\succ^{**} \subseteq \succ$, for all $x \in X$, $p(x, \cdot)$ is weakly increasing, and for all $k \in \mathbb{Z}_+$, $\underline{p}(x,k) \leq p(x,k) \leq \overline{p}(x,k)$, where*

$$\underline{p}(x,k) \;=\; \max\left(\{0\} \cup \{w_t(x) + 1 \;:\; x \in M_t,\; x \succ c_t,\; 1 \leq t < T(x, k+1)\}\right) \tag{44}$$

$$\overline{p}(x,k) = \begin{cases} w_{T(x,k+1)} & if\ x \text{ is chosen at least } k+1 \text{ times} \\ \infty & \text{otherwise} \end{cases} \tag{45}$$

Proof. Let (M, c) be a monotonic repetition averse menu-choice sequence represented by $(\succ, p)$, and let $\succ^{**}$ be its associated revealed preference given by eq. (43). Suppose that $x \succ^{**} y$. Hence, there are periods t, s such that $t > s$, $c_s = x$, $y \in M_s$, $c_t = y$, and $w_s(y) \geq w_t(y)$. Then $w_s(y) \geq w_t(y) \geq p(y, n_t(y)) \geq p(y, n_s(y))$. Therefore, y is considered in period s and hence, $x \succ y$. Since $\succ$ is acyclic, and $x \succ a$ for all $x \in X$, we conclude that $\succ^{**}$ is acyclic, and for all $x \in X$, not $a \succ^{**} x$.

The argument above shows that $\succ^{**} \subseteq \succ$. Suppose, additionally, that all alternatives are chosen. If $p(x,k) < \underline{p}(x,k)$ for some k, x, then there is a period t such that $1 \leq t < T(x, k+1)$, $p(x,k) < w_t(x) + 1$, $x \in M_t$, and $x \succ c_t$, in contradiction to $(\succ, p)$ representing c. Similarly, $p(x,k) > \overline{p}(x,k) = w_{T(x,k+1)}$ implies $x \neq c_{T(x,k+1)}$, in contradiction.

Conversely, let $\succ^{**}$ be the revealed preference associated with the menu-choice sequence (M, c) according to eq. (43), let $\succ^{**}$ be acyclic and for all $x \in X$, not $a \succ^{**} x$. Let

[12] In this setup, $w_s(y)$ coincides with $\underline{w}_s(y)$; next, $w_s(y) \geq w_t(y)$ implies that y is chosen before t and hence, $w_t(y) = \overline{w}_t(y)$; conversely, if $w_s(y) \geq \overline{w}_t(y)$, then $\overline{w}_t(y) < \infty$ and hence, $w_t(y) = \overline{w}_t(y)$.

$X_0 = \{x \in X : \exists t : c_t = x\}$. Consider a pair $(\succ, p)$, where $\succ$ is some completion of $\succ^{**}$ to a linear order on $X \cup \{a\}$ such that $x \succ a$ for all $x \in X$ (this linear order exists since $\succ^{**}$ is acyclic, and $a \succ^{**} x$ for no x), and $p : \tilde{X} \times Z_+ \to \tilde{Z}_+$ satisfies the following conditions. First, $p(a,k) = 0$ for all k. Second, $p(x,k) = \infty$ for all $x \in X \backslash X_0$ for all k. Third, for all $x \in X_0$, $p(x, \cdot)$ is weakly increasing and satisfies $\underline{p}(x,k) \leq p(x,k) \leq \overline{p}(x,k)$ for all k, where $\underline{p}(x,k)$ and $\overline{p}(x,k)$ are given by eq. (44), (45).

Let us verify that there exists p that satisfies the condition given above. Pick $p(x,k) = \underline{p}(x,k)$; note that $\underline{p}(x,\cdot)$ is weakly increasing by construction. It remains to show that $\underline{p}(x,k) \leq \overline{p}(x,k)$. Suppose not, then $\underline{p}(x,k) > 0$, $\overline{p}(x,k) \neq \infty$, and there are periods t, s, $t = T(x,k+1) > s$, $x \in M_s$, $x \succ c_s$, $x = c_t$, and $w_s(x) + 1 = \underline{p}(x,k) > \overline{p}(x,k) = w_t(x)$. Then $w_s(x) \geq w_t(x)$, and we conclude that $c_s \succ^{**} x$, in contradiction to $x \succ c_s$.

Let d be the sequence of choices represented by the monotonic repetition averse model $(\succ, p)$. We are going to show that $d_t = c_t$ for all t. It is without loss to consider $X \backslash X_0 = \varnothing$. If we prove that in this case, $c = d$, then in the case $X \backslash X_0 \neq \varnothing$, any $x \in X \backslash X_0$ does not appear neither in c nor in d.

Let $x = c_t$. If $x = a$, then x is considered by d in period t. If $x \in X$, then $t = T(x, n_t^c(x) + 1)$ and hence, $p(x, n_t^c(x)) \leq \overline{p}(x,k) = w_{T(x,n_t^c(x)+1)} = w_t(x)$. Thus, x is considered by d in period t.

Towards a contradiction, assume there is $y \in X$ such that $y \succ x$, and y is considered by d in period t; that is, $w_t(y) \geq p(y, n_t^c(y))$, and $y \in M_t$. Then $t \in \{s \in \mathbb{N} : y \in M_s,\ y \succ c_s,\ 1 \leq s < T(y, n_t^c(y) + 1)\} \neq \varnothing$. Hence,

$$p(y, n_t^c(y)) \geq \underline{p}(y, n_t^c(y)) = \tag{46}$$

$$= \max\left(\{w_s(y) + 1 \ :\ y \in M_s,\ y \succ c_s,\ 1 \leq s < T(y, n_t^c(y) + 1)\}\right) \geq w_t(y) + 1 > w_t(y)$$

in contradiction. Thus, $c = d$; that is, $(\succ, p)$ represents c. $\square$

The exponential discounting aversion model

We discuss a class of parametric models that allow the consideration of an alternative to depend on the whole history of its own choice, although the effect of consumption on consideration gradually vanishes for consumption that happened further in the past.

For every $x \in X$, fix parameters $\delta_x \in (0,1]$ and $v_x \in \mathbb{R}_+$. Given a choice sequence c, for any partial history of choices $h = (c_1, ..., c_t)$, define

$$K(h) = \left\{ x \in X : \sum_{s=0}^{t-1} \mathbf{1}(c_{t-s} = x)\, \delta_x^{s+1} < v_x \right\} \cup \{a\}. \tag{47}$$

$K(h)$ is interpreted as the set of alternatives that are considered after the partial history h. The unit weight of each instance of past consumption of an alternative x fades at exponential rate δ_x, and for x to be considered the accumulated discounted weight must fall below the critical threshold v_x.

Let $\succ$ be a linear order on $X \cup \{a\}$ with $x \succ a$ for all $x \in X$. The choice sequence c has an *exponential discounting representation* if it satisfies, for every t:

$$\{c_t\} = \max\left(\{x : x \in K(c_1, ..., c_{t-1})\}, \succ\right). \tag{48}$$

The baseline model with constant aversion phases $p(x)$ can be obtained as a special case of an exponential discounting representation. Let $\delta_x = \frac{1}{2}$ and $v_x = \left(\frac{1}{2}\right)^m$ for some $m \in \mathbb{Z}_+$. If at any period t alternative x has been chosen in at least one of the previous m periods, then

$$\sum_{s=0}^{t-2} \mathbf{1}(c_{t-1-s} = x)\, \delta_x^{s+1} \geq \left(\frac{1}{2}\right)^m = v_x \tag{49}$$

Conversely, if x has not been chosen at *any* of the previous m periods, then

$$\sum_{s=0}^{t-2} \mathbf{1}(c_{t-s} = x)\, \delta_x^{s+1} \leq \sum_{s=m+1}^{t-2} \delta_x^{s+1} < \frac{1}{1-\delta_x} \delta_x^{m+1} = \left(\frac{1}{2}\right)^m = v_x \tag{50}$$

Thus, x is considered if and only if it has not been chosen in any of the previous m periods, which corresponds to setting $p(x) = m < \infty$ in the baseline model. The case $p = \infty$ is dealt with by setting $\delta_x = 1$ and $v_x = 1$.

The next two examples illustrate that the exponential discounting model and monotonic repetition aversion models are not nested.

Example 22. Let $X = \{x\}$. Define the aversion map p by setting $p(x,k) = k$ for all k. The resulting choice sequence is $c = (x, a, x, a, a, x, a, a, a, x, a, a, a, a, x, a, a, a, a, a, x, ...)$. Suppose, towards a contradiction, that c has an exponential discounting representation with parameters (v, δ). Denote by $t_k = T(x, k+1) - 1$ the period preceding the $(k+1)$-st choice of x. Since $c_{t_k} = a$, and $c_{t_k - 1} = ... = c_{t_k - k + 1} = a$, it must be that $v \leq \delta^k + \delta^{k+1} + \delta^{k+2} + ... = \dfrac{\delta^k}{1-\delta}$. On the other hand, since $c_3 = x$, we have $v \geq \delta^2$. These two inequalities contradict each other for any $k > \log_\delta(1-\delta) + 2$.

Example 23. Let $X = \{x\}$. Consider an exponential discounting model with parameters $\delta_x = 0.8$ and $v_x = 1$. The resulting choice is $c_t = (x, x, a, a, x, a, a, x, a, a, a, x, a, a, x, ...)$ which is inconsistent with the monotonic repetition aversion model since the gap in choice of alternative x between periods 8 and 12 is 3 periods, between periods 12 and 15 is 2 periods.

Appendix C: Long-run frequencies in the baseline model

As noted in Section 6, the choice sequence in the baseline model eventually cycles and thus, the long-run choice frequencies $\varphi(\cdot, X)$ given by eq. 25 are well-defined.

Theorem 24. *Consider the baseline model. If $x \succ y$ and $p(x) \leq p(y)$, then the long-run frequency of choice of x is at least as large as that of y; that is, $\phi(x, X) \geq \phi(y, X)$.*

Proof. The only non-trivial case is when $p(x) \leq p(y) < \infty$, and both x and y are chosen. Let $T(x, k)$ be the period in which x is chosen for the k-th time, and define similarly $T(y, k)$. If y is chosen finitely many times, then $\phi(y, X) = 0$, and the statement follows. It remains to consider the case when y is chosen infinitely many times. Let us show by induction in $k = 1, 2, ...$ that $T(x, k) \leq T(y, k)$. The base of the induction $k = 1$ follows from the fact that all alternatives are considered before their first choice, and $x \succ y$. Let us prove the induction step. Towards a contradiction, assume there is k such that x is not chosen for k-th time at or before period $T(y, k)$. By the induction assumption, $T(x, k-1) \leq T(y, k-1)$. Hence, $T(y, k) - T(x, k-1) - 1 \leq T(y, k) - T(y, k-1) - 1 \leq p(y) \leq p(x)$, which implies that x is considered at period $T(y, k)$, but y is chosen, in contradiction. □

Next, we show by example that the model fails both properties referred to in Section 6 of the text:

Regularity: For all $x \in A$ and $y \in X$: $\varphi(x, A) \geq \varphi(x, A \cup \{y\})$

Luce Independence: For all $x, y \in A \cap B$: $\frac{\varphi(x,A)}{\varphi(y,A)} = \frac{\varphi(x,B)}{\varphi(y,B)}$

Example 25. (Violation of Regularity and Luce Independence) Let $A = \{y, z\}$ and $B = \{x, y, z\}$ with $x \succ y \succ z$ and $p(x) = 1$, $p(y) = 2$ and $p(z) = 3$. The cycle in B is $zxyx$, so that $\varphi(z, A) = \frac{1}{4}$. The cycle in A is $yzayazyaa$, so that $\varphi(z, A) = \frac{2}{9} < \frac{1}{4}$, violating Regularity. Moreover, $\frac{\varphi(y,A)}{\varphi(z,A)} = \frac{2}{3} \neq 1 = \frac{\varphi(y,B)}{\varphi(z,B)}$, violating Luce Independence.

Example 26. (Violation of Luce Independence) Let $A = \{x, z\}$ and $B = \{x, y, z\}$, with $x \succ y \succ z$ and $p(x) = 3, p(y) = 4, p(z) = 5$. The cycle in A is

$$xzaaxaazxaaa$$

while in B the first four choices are $xyza$ and they are followed by the cycle

$$xayaxzayxaazxyaaxzyaxaayxzaaxyaz.$$

We have $\varphi(x, A) = \frac{1}{4} = \varphi(x, B)$, and $\varphi(z, A) = \frac{1}{6} > \varphi(z, B) = \frac{5}{32}$. Therefore, the addition of y to the menu has a strictly negative impact on z and no impact on y, violating Luce Independence without violating Regularity.