Time for Memorable Consumption*

Stefania Minardi† and Andrei Savochkin‡

Preliminary draft: July 9th, 2018

A consumption event is memorable if its memory affects an agent's well-being at times after the material consumption. We develop an axiomatic model of memorable consumption in a dynamic setting. The representation takes the form of exponential discounting, and features additional terms that accumulate utility from the recollection of past consumption. We analyze alternative processes by which the memorable effect accrues over time and show that our model supports well-known phenomena in psychology, such as the peak-end rule, duration neglect, and adaptation trends. We study a prominent special case in which memory evolves according to a Markovian law and develop comparative statics with respect to strength and longevity of memory. As an application, we introduce memorable consumption into the standard linear-quadratic consumption-savings problem and examine its implications for life-cycle patterns.

1 Introduction

In psychology and behavioral science, it has been widely recognized that one's subjective well-being at any point in time is not simply determined by the consumption at that moment — a crucial role is played by the recollection of past experiences. This idea is supported by sizable evidence from different types of experiments.[1] Evoking early ideas of Bentham (1789) and Edgeworth (1881), Kahneman's well-known contributions propose a distinction between 'moment utility' and 'remembered utility:' In his view, a hedonic experience consists of a sequence of moments for which one can instantly measure the degree of pain or pleasure, that is, the moment utility; the ex post judgement of the overall experience gives rise to the remembered utility. When viewed through the lens of modern

*We thank Itzhak Gilboa for many valuable comments and discussions. The authors gratefully acknowledge the financial support of the Investissements d'Avenir program (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047).

†Economics and Decision Sciences Department, HEC Paris, minardi@hec.fr, http://www.hec.fr/minardi/

‡New Economic School, asavochkin@nes.ru, http://pages.nes.ru/asavochkin/

[1] Among many, see Varey and Kahneman (1992), Kahneman, Diener, and Schwarz (1999), and Kahneman (2000a, 2000b). Here we refer to the strand of the literature that is closest to behavioral economics. For earlier references, see, e.g., Elster and Loewenstein (1992) and Diener, Suh, Lucas, and Smith (1999). The idea that past memories may influence well-being goes back to, at least, Smith (1759).

Date: 2018-07-09 14:54:50 Revision: 3abf17a

economics, one problematic aspect of this approach is that its core concepts are not linked to choice behavior. By contrary, moment and remembered "utilities" reflect hedonic states, such as the perceived intensity of pain or pleasure, and their measurements are traditionally based on self-reports of these feelings. While such methodology is common practice in psychological studies, understanding precisely what is being measured and incorporating it into the economic policy analysis remain a serious challenge.[2]

This paper focuses on consumption events that can be referred to as memorable. Indeed, life achievements or, more simply, an exotic vacation, can have enduring effects on a person long after the corresponding events took place. Our goal is to tighten the link between theory and empirical evidence by modeling the notion of memorability within the revealed-preference paradigm. We develop a theory of preferences in which an agent's well-being at a given point in time is affected non only by the current material consumption, but also by the recollection of memorable events experienced in the past. Our agent recognizes that her current choices may generate valuable memories that will affect her future well-being. Through this channel, memorability affects not only the well-being, but also choice behavior. Moreover, the effect of past memories may well depend on various features of the consumption history, such as the intensity or the frequency of the experiences, thereby allowing for a rich dynamics.

Our contribution is threefold. First, we propose a way to separate behaviorally the material effect of consumption for the present moment from its memorable effect that the agent enjoys in subsequent periods. To this end, we lay an axiomatic foundation for a dynamic model of memorable consumption. Second, our theory allows us to determine whether a particular consumption experience is perceived as memorable or ordinary for the agent. If consumption is represented by a bundle of distinct goods (or categories of goods), we can identify which goods are memorable. Thus, the notion of memorable good is endogenous and subjective in our theory. Third, we develop a theory of Markovian memory as a special case of our model and provide an axiomatic foundation for it. Markovian specification is useful for thinking about memory as a dynamic variable, and is suitable for applying standard dynamic programming methods to solve for optimal consumption in applications. The

[2]See Kahneman, Wakker, and Sarin (1997) and Kahneman (1999) for a more formal treatment of the measurement issue and a discussion of the main difficulties.

additional structure also allows us to distinguish between different aspects of the memory effect, such as longevity and strength, and to make interpersonal comparisons along these dimensions.

We study memorable consumption in a dynamic framework of preferences over consumption streams of different finite length.[3] A typical consumption stream of length t is denoted by $f = (f_0, f_1, \dots, f_{t-1})$, where $f_\tau \in \mathcal{C} \subseteq \mathbb{R}^N$ for $N \geq 1$, is the consumption bundle at time $\tau = 0, \dots, t-1$. In its simplest and most general form, our agent evaluates a stream f according to the following criterion:

$$V(f) = \sum_{\tau=0}^{t-1} \beta^\tau \big[u(f_\tau) + M(f_{\tau-1}, \dots, f_0, 0, \dots) \big]. \tag{1}$$

As in the standard theory of exponential discounting, the parameter $\beta \in (0,1)$ is a discount factor and the index $u(f_\tau)$ represents the agent's direct utility of consuming bundle f_τ at time τ. The novel component here is $M(f_{\tau-1}, \dots, f_0, 0, \dots)$ which represents the agent's utility derived from the memory of the consumption history $(f_{\tau-1}, \dots, f_0)$. The expression $u(f_\tau) + M(f_{\tau-1}, \dots, f_0, 0, \dots)$ captures the agent's total subjective well-being that can be attributed to time τ. The value of $M(f_{\tau-1}, \dots, f_0, 0, \dots)$ is positive for pleasant memories; however, M is allowed to take negative values to represent unpleasant memories that the agent would prefer not to carry over into the future, if possible.

The function M in the above representation is identified uniquely, up to multiplying by a positive constant jointly with u. If M equals zero for all streams, then the agent is not subject to the phenomenon of memorability and our representation reduces to the standard exponential discounting. In this case, no consumption or experience is memorable for the agent from the point of view of the effects on choice. Consumption in our model is represented by points in $\mathbb{R}^N$ for arbitrary $N \geq 1$. Hence, the role of memories can be analyzed for the aggregate consumption ($N = 1$) or for single consumption bundles ($N > 1$). If consumption is represented by bundles and the memorable effects of consumption are separable across goods, then our theory allows the analyst to learn from the agent's choice which goods in the bundles are memorable and which are "ordinary."

Note that memorability is not the only reason for the past to affect the current utility. One particularly striking example of such interdependence is Mom's Treat that is discussed

[3] In our formal setup, the domain of preferences consists of lotteries over such streams.

by Machina (1989, p. 1643) and can be traced back to much earlier literature. Suppose that a Mom has a single indivisible treat that she can give either to her daughter or to her son. In general, she is indifferent between giving the treat to each child. However, if her son got a treat on the previous occasion, she will strictly prefer to give the treat to her daughter today. Naturally, such a preference does not rely on whether or not giving a treat to a child is a memorable experience — it is guided by concerns about fairness. There are many other reasons for history dependence, including an intrinsic preference for variety (a person would prefer to accompany his popcorn with a different movie than what was watched last night, even if both movies are not the ones to remember) and habit formation. We would like to emphasize that our model is not a universal theory of history-dependent utility; rather, we are interested precisely in the phenomenon of memorability, its effect on choice, and ways to incorporate it into economic analysis. Our focus manifests distinctly in the proposed axioms and in the derived uniqueness results.

Representation (1) provides a general model to study different processes by which the memorable effect of past consumption accrues over time. The paper discusses three special cases — two examples of well-known time-dependent laws, and, in more detail, the time-invariant Markov law.

Our first example provides a time-dependent specification of the memory function M that accommodates the so called *peak-end rule* and the related phenomenon of *duration neglect* (Fredrickson and Kahneman, 1993). Our second example is based on the idea of adaptation to the exposure of repeated similar experiences; widely studied in psychology, its driving forces are not yet fully understood.[4] We capture a specific process of adaptation in the way future flows of utilities are generated from the current enjoyment of memorable goods. By relating the origins of adaptation to the law of motion of memory, we offer a formal setting to study the determinants of adaptation and its effects on well-being. This has important normative implications. For instance, one can use our model to analyze the efficacy of alternative preventive measures, such as the introduction of breaks into the repeated consumption of pleasant experiences. Furthermore, the concept of adaptation has often been advocated to explain the observed phenomenon of preference for increasing

[4]See, e.g., Diener (1984) and Frederick and Loewenstein (1999) for classic surveys on the topic of subjective well-being and adaptation.

sequences of outcomes over equivalent decreasing ones — a well-known violation of the classic discounted utility model. Our model suggests a behaviorally founded reason for preferring increasing profiles. By choosing upward profiles, agents may inhibit the process of getting used to high levels of consumption that they might be unable to sustain in the future.

In addition to the above two examples, we develop axiomatic foundations of a special case of representation (1) in which memory evolves according to a time-invariant Markov law. A consumption stream $f = (f_0, \ldots, f_{t-1})$ is now evaluated according to

$$V(f) = \sum_{\tau=0}^{t-1} \beta^{\tau} \big[u(f_{\tau}) + m_{\tau-1} \big], \tag{2}$$

where $m_{\tau} = \psi(m_{\tau-1}, f_{\tau})$ for $\tau = 0, \ldots, t-2$ and $m_{-1} = 0$. The key feature is that the utility of memorable consumption can be thought of as a "stock" variable, that is determined at each point in time only by its value in the previous period and the current consumption. Such recursive specification has the advantage of being highly tractable and amenable to be used in macroeconomic applications. With the additional Markovian structure of memory, we conduct a comparative statics analysis and identify two independent channels through which memories can affect the overall utility in representation (2): One is related to the persistency of memory in the agent's mind, and, hence, to the rate at which past memories decay. The other one is related to the sensitivity of the agent's memory to consumption and the ability of consumption to generate new memories. Inter alia, this comparative analysis helps distinguish the relative contribution to the total well-being of the direct utility from consumption and of the past memorable experiences.

Related Literature

We conclude this introduction by placing our work within the theoretical literature. Our theory entails a particular kind of violation of time-separability which places it in stark contrast with other history-dependent phenomena, notably, habit formation. Two forces drive decision-making in our model. On one side, the desire to enjoy current consumption; on the other side, the investment in memories to be enjoyed in the future. These two forces are intertwined since our far-sighted agent envisions that the choice of current consumption will feed the memory function in the future and, vice versa, the joy of memories will affect the overall utility at each point in time. The departure from the standard time-separable approach arises purely from the memory component. Only the latter is influenced by a process of reference dependence: for instance, whether a fine dining experience will generate memories depends on the reference point inherited from past experiences of the sort. Yet, tastes do not change over time: consumption history does not affect the utility of current consumption as it happens in models of habit formation. Our axiomatization captures precisely such reference dependence in *memory*, while ruling out reference dependence in the *direct value* of consumption. Interestingly, these differences are corroborated by opposite predictions: habit formation typically induces a desire for consumption smoothing. On the other side, the investment in memorable goods generates lumpy patterns of consumption, as argued in Hai, Krueger, and Postlewaite (2015).

The notion of memorable good has been developed first by Hai et al. (2015) and Gilboa, Postlewaite, and Samuelson (2016). They extend the standard consumption-saving setting by presenting a model that considers an ordinary good and a memorable good as primitive elements. The key feature of their model is that the consumption of the memorable good generates additional flows of utility only if it exceeds a threshold level determined by previous memorable experiences. They show that optimal consumption profiles of memorable goods exhibit spikes that cannot be justified by the presence of indivisibilities as traditionally argued for durable goods. Hai et al.'s (2015) theoretical findings are complemented by an empirical analysis highlighting that the expenditure volatility on memorable goods is higher than that on nondurable goods and, by contrary, lower than that on durable goods. As important empirical implications, memorable goods may play an important role in reducing the magnitude of the welfare losses due to consumption fluctuations and in rationalizing the evidence on the excess sensitivity of consumption to anticipated income shocks. By taking a more theoretical perspective, Gilboa et al. (2016) provide an axiomatic foundation of the static utility structure, given by $u(x, y) + v(y, z)$, at the basis of their model. The term $u(x, y)$ represents the current utility of consuming the ordinary good x and the memorable good y; the term $v(y, z)$ is the memory utility generated by consuming y currently and z in the past.

Compared to the above papers, we pursue a different line of research. We lay a behavioral foundation of memorable consumption within a temporal framework. We illustrate the empirical relevance by studying alternative laws of motion which support a large body of findings in psychology. By taking a subjective perspective, we let the agent reveal whether a good is memorable through her choice behavior. Furthermore, we analyze the comparative effects of memorable consumption on welfare.

2 Setup

Let $\mathcal{C} \subseteq \mathbb{R}^N$ for some $N \in \mathbb{N}$ be the space of consumption bundles, which we assume to be nondegenerate and connected. Its typical element is denoted by $c = (c^1, \ldots, c^N)$. The set $\mathcal{F}_t = \mathcal{C}^t$ for $t \in \mathbb{N}$ represents the collection of finite consumption streams of length t, with the typical element given by $f = (f_0, f_1, \ldots, f_{t-1})$. Let also $\oslash$ be a stream of length zero and $\mathcal{F}_0 = \{\oslash\}$. We denote by $\mathcal{F} = \bigcup_{t=0}^{\infty} \mathcal{F}_t$ the collection of all finite consumption streams. The sets $\mathcal{F}_t$ for $t \in \mathbb{N}$ are endowed with the sup-norm topology. For an element $f \in \mathcal{F}$, let $\ell(f) := t$ if $f \in \mathcal{F}_t$.

For $t \in \mathbb{N}$, let $\mathcal{L}_t = \Delta(\mathcal{F}_t)$ be the space of lotteries (probability distributions with finite support) over streams of length t, and let $\mathcal{L} = \Delta(\mathcal{F})$ be the space of lotteries over all consumption streams of finite length. The agent's behavior is described by a preference relation (complete preorder) $\succsim$ on $\mathcal{L}$.

As usual, for every $P, Q \in \mathcal{L}$ and $\alpha \in [0,1]$, the lottery $\alpha P + (1-\alpha)Q \in \mathcal{L}$ is defined by $\alpha P(f) + (1-\alpha)Q(f)$ for every $f \in \mathcal{F}$. The spaces $\mathcal{L}_t$ for $t \in \mathbb{N}$ are endowed with the topology of weak convergence: A net $\{P_\alpha\}_\alpha$ converges to some lottery P iff, for any continuous and bounded function $U : \mathcal{F}_t \to \mathbb{R}$, we have $\int U\, dP_\alpha \to \int U\, dP$.

3 Axiomatic Foundations

Notation. For any streams $f = (f_0, f_1, \ldots, f_k)$ and $h = (h_0, h_1, \ldots, h_m)$, let $h|f$ denote a concatenated stream $(h_0, h_1, \ldots, h_m, f_0, f_1, \ldots, f_k)$. Furthermore, for any $P \in \mathcal{L}$ and any $h \in \mathcal{F}$, we use the notation $h|P$ for a lottery Q obtained from P by prepending h to the streams in the support of P: Formally, Q is defined as $Q(f) = P(f')$ if $f = h|f'$ for some $f' \in \mathcal{F}$, and $Q(f) = 0$ otherwise. Also, as usual, we identify a degenerate lottery giving some stream $f \in \mathcal{F}$ with probability one with the stream itself.

3.1 Axioms

Framework assumptions

We start by posing three standard assumptions which are responsible for the discounted expected utility structure of the model.

Axiom A1 (Stationarity)**.** *There exists a consumption bundle that we identify with* 0 *such that, for any* $P, Q \in \mathcal{L}$*, we have*

$$P \succsim Q \quad \Leftrightarrow \quad (0)|P \succsim (0)|Q.$$

Axiom A2 (Impatience)**.** *For any* $c \in \mathcal{C}$ *such that* $(c) \succ (0)$*, we have*

$$(c) \succ (0, c) \succ (0).$$

Axiom A3 (Independence)**.** *For any* $P, Q, R \in \mathcal{L}$ *and* $\alpha \in (0, 1]$*, we have*

$$\alpha P + (1 - \alpha)R \succsim \alpha Q + (1 - \alpha)R \quad \Leftrightarrow \quad P \succsim Q.$$

Axioms pertinent to memory

We now present the two key axioms which delineate the behavioral features of memorable consumption and distinguish it from other forms of history-dependent phenomena. The first axiom asserts that the agent's preferences between streams that differ only in the last-period consumption are independent of the consumption in previous periods.

Axiom A4 (Risk Preference Consistency)**.** *For any* $f, g \in \mathcal{F}$ *and* $p, q \in \mathcal{L}_1$*, we have*

$$f|p \succsim f|q \quad \Rightarrow \quad g|p \succsim g|q.$$

This axiom guarantees that the agent's tastes remain unchanged after varying histories. As an implication, it rules out various types of reference-dependent evaluation of the current consumption, most notably habit formation and preferences for intrinsic variety. Besides that, since the streams f and g can have different lengths, tastes and risk attitudes remain unchanged with the passage of time. In particular, this rules out potential psychological effects that the realization of extreme outcomes may have on future risk-taking behavior.

The second axiom is concerned with tradeoffs between memories and consumption.

Axiom A5 (Memory-Consumption Tradeoff Consistency)**.** *For any $t \in \mathbb{N}$, $f, g \in \mathcal{F}_t$, and $p, q \in \mathcal{L}_1$, we have*

$$f|p \succsim g|(\tfrac{1}{2}p + \tfrac{1}{2}q) \quad \Leftrightarrow \quad f|(\tfrac{1}{2}p + \tfrac{1}{2}q) \succsim g|q.$$

To interpret the axiom, consider the situation of a tradeoff between consumption in a fixed (the last) period, and pleasant memories for the future that are generated by the consumption stream in periods from the first to second before last. Namely, suppose that changes in the initial part of the consumption stream can be counter-balanced by replacing in the last period a consumption lottery p with a lottery that is a midpoint between p and some q. The axiom ensures that, in this case, a similar replacement in the last period of the midpoint between p and q with q — which is a replacement that has the same distance and direction in the space of last-period consumption lotteries — should have the same counter-balancing effect. It thus calibrates the relative effects of memory and consumption in quantitative terms. For further intuition, observe that a simple implication of this axiom is[5] that $f|p \succsim g|p$ if and only if $f|q \succsim g|q$ for all $f, g \in \mathcal{F}_t$, and $p, q \in \mathcal{L}_1$. This implied property highlights the qualitative content of the axiom by making it clear that the desirability of f versus g is independent of last-period consumption. Thus, it rules out additional effects on the subjective well-being that the agent may obtain in early periods from the mere anticipation of her consumption in the last period.

The axioms of this section clearly hold in the standard model of discounted expected utility. What is noteworthy, however, is that they capture consistency properties that hold only with respect to the last-period consumption. The last period becomes significant in our theory (and different from all preceding periods) because effectively the consumption in that period is never memorable — for the reason that there are no subsequent periods in which a memory generated in that period can be enjoyed. Hence, our axioms allow for reference dependence in *memory* (as illustrated by our examples later on), but rule out reference dependence in the *direct value* of consumption. This way, they formally capture our intention to model memorable consumption in isolation from other behavioral phenomena, including habit-formation (ruled out by Risk Preference Consistency) and anticipation (ruled out by Memory-Consumption Tradeoff Consistency).

[5]We refer to Lemma 8 in the Appendix for a proof of this statement.

Technical requirements

We conclude with few technical assumptions.

Axiom A6 (Continuity). *(1) For all* $P \in \mathcal{L}$ *and all* $t \in \mathbb{N}$*, the sets* $\{Q \in \mathcal{L}_t : Q \succsim P\}$ *and* $\{Q \in \mathcal{L}_t : P \succsim Q\}$ *are closed. (2) For all* $P, Q, R \in \mathcal{L}$*, the sets* $\{\alpha \in [0,1] : \alpha P + (1-\alpha)Q \succsim R\}$ *and* $\{\alpha \in [0,1] : R \succsim \alpha P + (1-\alpha)Q\}$ *are closed.*

Axiom A7 (Nondegeneracy). *There exist* $c^*, c_* \in \mathcal{C}$ *such that* $(c^*) \succ (0) \succ (c_*)$ *and* $(c^*, 0) \succ (c_*, 0)$*.*

3.2 Basic representation

Notation. Let $\mathcal{C}_0^\infty$ denote the set of infinite sequences of elements of $\mathcal{C}$ for which only finitely many elements are distinct from 0, where 0 is the element of $\mathcal{C}$ given by the Stationarity axiom.

We endow this set with the following topology: a net $\{f^{(\alpha)}\}_\alpha$ converges to some f in $\mathcal{C}_0^\infty$ iff, for some $T \in \mathbb{N}$ such that $f_t = 0$ for all $t \geq T$, there exist an index α_0 such that $f_t^{(\alpha)} = 0$ for all $\alpha \geq \alpha_0$ and $t \geq T$, and $\sup_{0 \leq t \leq T} |f_t - f_t^{(\alpha)}|$ converges to zero. With a slight abuse of terminology, we will say that a function $\Phi : \mathcal{C}_0^\infty \to \mathbb{R}$ is bounded iff, for any $T \in \mathbb{N}$, there exists $K > 0$ such that, for any $f \in \mathcal{C}_0^\infty$ such that $f_t = 0$ for all $t \geq T$, we have $|\Phi(f)| \leq K$.

Theorem 1. *A complete preorder* $\succsim$ *on* $\mathcal{L}$ *satisfies Axioms (A1)–(A7) if and only if there exist a scalar* $\beta \in (0,1)$*, a continuous and bounded function* $u : \mathcal{C} \to \mathbb{R}$ *such that* $u(0) = 0$ *and* $\operatorname{range} u$ *contains both positive and negative numbers, and a continuous and bounded function* $M : \mathcal{C}_0^\infty \to \mathbb{R}$ *such that* $M(0, 0, \ldots) = 0$ *and*

$$V(P) = \sum_{f \in \operatorname{supp} P} P(f) \sum_{t=0}^{\ell(f)-1} \beta^t \big[u(f_t) + M(f_{t-1}, \ldots, f_0, 0, \ldots)\big] \tag{3}$$

is a utility representation of $\succsim$ *on* $\mathcal{L}$*.*

Theorem 1 delivers a representation of the agent's preferences that enriches the standard exponential discounting formula to allow for the effect of memorability. The usual ingredients of the evaluation formula are the scalar β, that represents the discount factor, and the function u, that measures the utility of a bundle of goods at the time of material consumption. Memorable goods or experiences, besides their direct value at the moment of consumption, generate additional utilities in the future, and their flows are measured by a novel component — the function M. The above representation can also be interpreted as if a stream f is evaluated by an exponential discounting criterion according to which the overall utility at any time t is given by $u(f_t) + M(f_{t-1}, \ldots, f_0, 0, \ldots)$. The agent derives utility from the current consumption, captured by $u(f_t)$, and from the recollection of past memorable experiences, captured by $M(f_{t-1}, \ldots, f_0, 0, \ldots)$. Note that the function M is backward-looking: Its first argument is the most recent past consumption, then the second-to-most-recent one, and so on. The sequence of arguments ends with an infinite sequence of zeroes since all consumption streams are assumed to be finite in this theorem.[6] Note that if the agent does not perceive any good or experience as memorable, we have $M(\cdot) \equiv 0$ and the representation reduces to the standard exponential discounting model.

The parameters β, u, and M that capture the agent's preferences are identified uniquely, as demonstrated by our next result. In comparison to the standard uniqueness results in utility theory, the only minor difference is that the functions u and M are unique only up to a positive multiplicative factor, whereas arbitrary additive constants are not allowed because we impose the convention of assigning the numeric value 0 to the neutral element identified by the Stationarity axiom.

Proposition 2. *Two triples (β, u, M) and $(\hat{\beta}, \hat{u}, \hat{M})$ represent the same binary relation $\succsim$ on $\mathcal{L}$ as in Theorem 1 if and only if $\beta = \hat{\beta}$, $\hat{u} = \lambda u$, and $\hat{M} = \lambda M$ for some $\lambda > 0$.*

3.3 Time- and history-dependent memory

One well-known heuristic about the way people recollect prolonged experiences is called the *peak-end rule*. Originally introduced by Fredrickson and Kahneman (1993), it builds upon the view that any hedonic experience can be thought of as consisting of a sequence of moments which could be identified, for instance, by the unfolding of time. According to the peak-end rule, the evaluation of a retrospective experience, whether positive or negative, is determined by the average of only two salient moments, the most intense value —

[6] A single memory function operating on infinite (but vanishing) streams could be replaced by a collection of functions operating on finite streams — $M_1(f_0)$, $M_2(f_1, f_0)$, and so on. Specifying functions in this way would require imposing additional constraints — it must be that $M_2(c, 0) \equiv M_1(c)$, and so on.

namely, the *peak* — and the value experienced at its end — namely, the *end*.[7] One notable implication is that the duration of an experience has no impact on its retrospective recollection. For instance, a short, but rather exotic, vacation may generate more intense memories in comparison to a longer, but more ordinary, vacation. This pattern, dubbed *duration neglect*, is observed in numerous experimental studies which suggest that prolonging an unpleasant experience by adding some extra moments of diminished discomfort may mitigate the subsequent assessment of the overall experience.[8] Our next example proposes a simple specification of the function M that accommodates the empirical evidence on the peak-end rule and duration neglect.

Example 1. *Let the function M from representation (3) be defined as*

$$M(\underbrace{0,\ldots,0}_{t}, f_k, \ldots, f_{k-l+1}, 0, f_{k-l-1}, \ldots, f_0, 0, \ldots) = \beta^t \, \frac{1}{2}\big(u(f_{j^*}) + u(f_k)\big)\, K, \tag{4}$$

for streams $f \in \mathcal{F}$ such that $u(f_j) > 0$ for all $j \in \mathbb{Z}_+$ with $k - l < j \leq k$, where $t, k \in \mathbb{Z}_+$, $l \in \mathbb{N}$, and $l \leq k+1$. Moreover, $j^* \in \mathbb{Z}_+$ satisfies $k - l < j^* \leq k$ and $u(f(j^*)) \geq u(f(j))$ for all $j \in \mathbb{Z}_+$ with $k - l < j \leq k$. The parameter $K > 0$ stands for the magnitude constant. For streams that do not conform to the above pattern, set M equal to zero.[9] *In this example, we identify the periods of no memorable "experiences" (as understood in the works of Kahnemann) with zero consumption, and the duration of an experience with the number of consecutive positive consumptions: In specification (4), the most recent memorable experience lasted l periods. The factor K is responsible for the magnitude of the memory effect (relative to ordinary utility from consumption), and the factor $\beta^t \in (0,1)$ leads to exponential decay of memory once the experience is over.*

Note also that the rule (4) can be applied not only to the overall consumption (that is, to f_τ representing the entire bundle consumed in period τ), but also to one particular dimension of the consumption, e.g., a generalized "vacation good."

To further illustrate, consider a stream of one-dimensional consumption $(f_0, f_1, \ldots, f_{10}) = (2, 6, 0, 0, 1, 3, 1, 1, 0, 0, 0)$, and suppose that $u(x) = x$ and $K = 1$. Then, the sequence of

[7] For experimental evidence, see, e.g., Ariely and Carmon (2000), Fredrickson (2000), Kahneman (2000a, 2000b), and references therein.

[8] E.g., Varey and Kahneman (1992) and Kahneman, Fredrickson, Schreiber, and Redelmeier (1993).

[9] This example focuses on positive experiences; nevertheless, it can be easily adapted to cover negative experiences, as well.

memory terms

$$M(f_0,0,\dots),M(f_1,f_0,0,\dots),\dots,M(f_{10},f_9,\dots,f_0,0,\dots)$$

is given by $2,6,6\beta,6\beta^2,1,3,2,2,2\beta,2\beta^2,2\beta^3$.

Another important class of behavioral regularities related to past memories is studied in the well-known *adaptation-level theory* in psychology.[10] For economics, the most important prediction of the theory is that repeated exposure to the same good experience will gradually attenuate the initial feeling of pleasure; similarly, a persistent exposure to the same bad experiences will make the feeling of discomfort wane.

Our model is not intended to capture full-fledged adaptation-level theory. Indeed, our axioms imply that the current utility from consumption is not reference-dependent. However, the *memorability* of experiences and their value at *recollection* time may well depend on the past history of similar experiences and exhibit the adaptation features.

Example 2. *Let the function M from representation (3) be defined as*

$$M(f_t,\dots,f_0,0,\dots)=G(f_t,A(f_{t-1},\dots,f_0,0,\dots)), \tag{5}$$

for all $f\in\mathcal{F}$ and $t\in\mathbb{Z}_+$, where $A:\mathcal{C}_0^\infty\to\mathbb{R}$ is defined as $A(f_t,\dots,f_0,0,\dots)=\alpha\sum_{\tau=0}^{\infty}(1-\alpha)^\tau f_{t-\tau}$, $\alpha\in(0,1)$, and $G:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$ is a continuous function that is monotone in the first argument and such that $G(0,0)=0$.

In this specification, $A(f_{t-1},\dots,f_0,0,\dots)$ represents the adaptation level acquired from consumption up to time $t-1$ and sets the reference point for new memories at time t. The coefficient α is the weight attributed to the most recent experience in determining the new adaptation level. In fact, the formula for A can be also written as $A(f_t,\dots,f_0,0,\dots)=\alpha f_t+(1-\alpha)A(f_{t-1},\dots,f_0,0,\dots)$. The function G in (5) measures the utility value of the memory from consuming bundle x after a history of consumption summarized by the reference level r. In Tversky and Griffin's (1991) terminology, $A(f_{t-1},\dots,f_0,0,\dots)$ corresponds to the endowment level accumulated up to time t, whereas G quantifies the contrast effect. The simplest specification for G can be $G(x,r)=\max\{x-r,0\}$, in which a positive flow of

[10] See Helson (1947, 1948) for origins of the theory that started with the perceptual adaptation in vision. For more recent works, see, e.g., Frederick and Loewenstein (1999).

memory utility is generated only if the most recent consumption exceeds the reference level. Naturally, a more general specification for G may accommodate a broader spectrum of adaptation trends in memories' recollection and, in particular, may not necessarily require new experiences to beat the prior record. Indeed, a person may have reached very high standards for fine dining and, at the same time, have a pleasant memory from coffee and pastry from some small bakery.

The broad paradigm of adaptation-level theory gives rise to a number of widely recognized patterns. For instance, the theory suggests that introducing an interval of lower consumption in a lengthy stream of positive consumption may make the agent to appreciate it more.[11] In the formal language, we may be interested in the comparison of mixtures of streams that follow the pattern

$$\frac{1}{3}(c,c,0,c,c,0,\dots)+\frac{1}{3}(0,c,c,0,c,c,\dots)+\frac{1}{3}(c,0,c,c,0,c,\dots)\succ \\ \frac{2}{3}(c,c,c,c,c,c,\dots)+\frac{1}{3}(0,0,0,0,0,0,\dots).$$

According to the standard criterion of expected discounted utility, the agent should be indifferent between them because, at each date, she consumes c with the probability $\frac{2}{3}$ and zero with the probability $\frac{1}{3}$ in both the left-hand and the right-hand sides. However, if memorability is taken into account, a constant stream of high consumption may generate less memory (and less utility from memory) then streams in which high consumption is interrupted. This can easily be accommodated in our model with a suitable choice of the parameters.

Another example is the following well-known challenge to the standard discounted utility criterion: people often exhibit a preference for a sequence of increasingly pleasant outcomes to the same outcomes but in the reverse order.[12] Similarly, when exposed to a sequence of negative outcomes, they often prefer a decreasing order. Such preference statements create a stark contrast with the predictions of the standard theory of discounting, according to

[11]This prediction is supported by evidence from psychology and marketing. See, e.g., Ariely and Zauberman (2000) and Nelson and Meyvis (2008).

[12]For instance, Loewenstein and Sicherman (1991) report evidence in support of a preference for increasing sequences of wages over time. Similarly, Loewenstein and Prelec (1991, 1993) show that, when confronted with the choice between a sequence of increasingly pleasant dining experiences and the same one but in decreasing order, the majority of subjects exhibits a preference for the upward-shaped profile.

which the best experience is the most desirable for an agent if it comes at the earliest moment. These patterns have a natural explanation if the role of adaptation is taken into account, and they can be accommodated in our theory. More importantly, our model suggests a behaviorally founded reason for preferring increasing profiles — if the utility from recollecting past experiences follows an adaptation process, then an increasing profile will generate higher utility flows from memory than an equivalent decreasing profile.

4 Markovian memory

4.1 The Markovian property

In this section we turn to a prominent special case of our general representation that is particularly suitable for applications in macroeconomics and repeated games. Specifically, we provide a behavioral characterization according to which the memory of past consumption adheres to a Markovian law of motion: the utility from memorable consumption at any time t is determined only by its value at time $t-1$ and the consumption at time t, and it does not depend directly on the patterns of consumption at earlier dates. Hence, the utility from memorable consumption can be thought of as a "stock" variable that is driven by the current consumption and evolves according to a time-invariant Markovian process.

Our analysis starts by introducing the notion of a tradeoff between memory and consumption.

Definition 1. We say that the memory after a stream f k-dominates the consumption z, where $k>0$, $f\in\mathcal{F}$, and $z\in\mathcal{C}$, if

$$(f|0)\succsim\frac{1}{k+1}f+\frac{k}{k+1}(f|z). \tag{6}$$

We denote (6) by $f\succsim^{\mathtt{m}-k}z$. A similar strict preference

$$(f|0)\succ\frac{1}{k+1}f+\frac{k}{k+1}(f|z).$$

is denoted by $f\succ^{\mathtt{m}-k}z$.

To understand the gist of this definition, observe that the stream on the left-hand side of (6) in comparison to the lottery on the right-hand side offers the agent a higher chance of enjoying the memory of f in the subsequent time period — on the left-hand side, the probability of enjoying such memory is one, while on the right-hand side it is only $\frac{k}{k+1}$, a difference of $\frac{1}{k+1}$. In exchange for that, the right-hand side offers the agent a potentially higher level of consumption in the last period, z instead of zero.[13] The additional consumption of z is available to the agent with the probability $\frac{k}{k+1}$. Thus, the pattern in (6) describes a preference for enjoying the memory produced by f over the direct benefits of consuming z. Moreover, this preference is quantified: If the agent prefers the left-hand side, then, loosely speaking, the pleasure of the memory produced by f is at least k times higher than the pleasure of consuming z.[14]

The notion of consumption-memory tradeoff allows us to compare consumption streams in terms of their value for the generation of memories at a specific point in time. As shown next, a stream f memory-wise dominates another stream g if, for any consumption bundle z that the agent is willing to give up to enjoy the memory of g, she is willing to give it up to enjoy the memory of f a fortiori.

Definition 2. For $f, g \in \mathcal{F}$, we say that f generates a higher value of memory for the next period in comparison to g if

$$g \succsim^{\mathtt{m}-k} z \;\Rightarrow\; f \succsim^{\mathtt{m}-k} z \qquad \text{for all } k > 0 \text{ and } z \in \mathcal{C}.$$

We denote such a relationship between streams f and g by $f \; \mathcal{R}^{\succsim} \; g$. Extending this definition, we say that f generates a strictly higher value of memory in comparison to g if

$$g \succsim^{\mathtt{m}-k} z \;\Rightarrow\; f \succ^{\mathtt{m}-k} z \qquad \text{for all } k > 0 \text{ and } z \in \mathcal{C},$$

and denote this relationship by $f \; \mathcal{S}^{\succsim} \; y$. Finally, we say that f generates the same value of memory as g if

$$g \succsim^{\mathtt{m}-k} z \;\Leftrightarrow\; f \succsim^{\mathtt{m}-k} z \qquad \text{for all } k > 0 \text{ and } z \in \mathcal{C},$$

and denote this relationship by $f \; \mathcal{I}^{\succsim} \; g$.

[13] The use of the neutral element 0 in the left-hand side of the above definition is convenient but not mandatory. This (and subsequent) definitions can be modified to use a different reference point for measuring tradeoffs.

[14] As usually is the case, it is the usage of lotteries that allows us to give cardinal meaning to relationships between utility levels.

The above definitions achieve two important goals. First, they provide a behavioral notion of what it means for one consumption stream $(f_0, \ldots, f_{t-1})$ to produce a higher-valued memory for the period t relative to another stream $(g_0, \ldots, g_{t-1})$, regardless of the utilities that these streams generate for periods $0, \ldots, t-1$. Second, these definitions enable the comparison of the value of memory for streams of different lengths. This demonstrates that it is behaviorally meaningful to attribute the memory utility $M(f_{t-1}, \ldots, f_0, 0, \ldots)$ to date t in the general representation (3).

We apply the above definition to formulate our key axiom for the Markovian representation.

Axiom A8 (Markovian Property)**.** *For any $f, g \in \mathcal{F}$,*

$$f \, \mathcal{I}^{\succsim} \, g \; \Rightarrow \; (f|c) \, \mathcal{I}^{\succsim} \, (g|c) \qquad \text{for all } c \in \mathcal{C}.$$

The antecedent of this property considers the situation in which the memory effect of f is equivalent to that of g. That is, both streams generate the same value of memory in the period following the consumption of their respective last component. The axiom maintains that if these two streams are extended by one additional period of consumption, then the ranking of their value for memory remains the same. Hence, the impact of g on memory is equivalent to that of f for any given last-period consumption $c \in \mathcal{C}$.

Our next theorem shows that this property, together with our basic axioms (A1)–(A7), delivers a convenient time-invariant Markovian representation.

Theorem 3. *Let $\succsim$ be a complete preorder on $\mathcal{L}$. The following statements are equivalent:*

(i) $\succsim$ satisfies Axioms (A1)–(A7) and (A8);

(ii) there exist a scalar $\beta \in (0,1)$, a continuous and bounded function $u : \mathcal{C} \to \mathbb{R}$ with $u(0) = 0$, a bounded interval I of $\mathbb{R}$ that contains 0, and a continuous function $\psi : I \times \mathcal{C} \to I$ with $\psi(0,0) = 0$ and $\operatorname{range} \psi = I$, such that a utility representation of $\succsim$ on $\mathcal{L}$ is $V(P) = \sum_{f \in \operatorname{supp} P} P(f) V(f)$ for all $P \in \mathcal{L}$, where $V(f)$ for all $f \in \mathcal{F}$ is computed as

$$V(f) = \sum_{t=0}^{\ell(f)-1} \beta^t \big[u(f_t) + m_{t-1} \big], \tag{7}$$

$$\text{where} \quad m_\tau = \psi(m_{\tau-1}, f_\tau) \text{ for } \tau = 0, \ldots, \ell(f) - 2,$$

$$m_{-1} = 0.$$

According to representation (7), the evaluation of a stream f at any time t is given by $u(f_t) + m_{t-1}$, where $u(f_t)$ is the material utility of f_t and m_{t-1} is the stock of memory accumulated up to time t. The function ψ describes the process of incorporating the memory effect of consuming f_t into m_{t-1}, giving rise to the next-period value, m_t. Similarly to all specifications of the memory utility discussed earlier, memory may have long-lasting effects here, as well. However, the dependence of m_t on consumption in periods $t-1, \dots, 1, 0$ is encapsulated in the previous stock of memory, m_{t-1}. This is the nature of our Markovian evolution of memory. Note that such recursive process of computing the values of m_t is particularly tractable because the function ψ is independent of time.

Theorem 3 represents preferences in terms of quadruples of the form (β, u, I, ψ). These quadruples are essentially unique, as shown next.

Proposition 4. *Two quadruples* (β, u, I, ψ) *and* $\left(\hat{\beta}, \hat{u}, \hat{I}, \hat{\psi}\right)$ *represent the same binary relation* $\succsim$ *on* $\mathcal{L}$ *as in Theorem 3 if and only if* $\hat{\beta} = \beta$ *and there exists* $\lambda > 0$ *such that* $\hat{u} = \lambda u$, $\hat{I} = \lambda I$, *and* $\hat{\psi}(m, c) = \lambda \psi(m/\lambda, c)$ *for all* $m \in \hat{I}$ *and* $c \in \mathcal{C}$.

4.2 Properties of the memory evolution function

In studying possible specifications for the law of motion of memory, monotonicity of the function ψ in its first or second argument stands out as a desirable feature. Here we show that these monotonicity properties have natural behavioral counterparts.

Axiom (Monotonicity in the Past Memory)**.** *For any* $f, g \in \mathcal{F}$, $f \; \mathcal{R}^{\succsim} \; g$ *implies* $(f|c) \; \mathcal{R}^{\succsim} \; (g|c)$ *for any* $c \in \mathcal{C}$.

This axiom prescribes that the relationship between any two non-degenerate streams f and g in terms of value of memory is preserved if they are both extended by the same one period of extra consumption c. In passing, note that Monotonicity in the Past Memory implies the Markovian property.

Axiom (Monotonicity in the Current Consumption)**.** *For any* $x, y \in \mathcal{C}$, $(x) \succsim (y)$ *implies* $(f|x) \; \mathcal{R}^{\succsim} \; (f|y)$ *for any* $f \in \mathcal{F}$.

This axiom simply ensures that a better bundle in terms of direct consumption should generate a higher value of memory if adjoined to any consumption stream.

The following proposition confirms that each of these properties is equivalent to the monotonicity of the memory evolution function in the respective argument.

Proposition 5. *Suppose that $\succsim$ is a complete preorder on $\mathcal{L}$ that satisfies Axioms (A1)–(A7) and (A8), and let (β, u, I, ψ) be its representation as in Theorem 3. Then:*

(i) $\succsim$ satisfies Monotonicity in the Past Memory if and only if $\psi(m_1, c) \geq \psi(m_2, c)$ for all $m_1, m_2 \in I$ such that $m_1 \geq m_2$ and all $c \in \mathcal{C}$;

(ii) $\succsim$ satisfies Monotonicity in the Current Consumption if and only if $\psi(m, c_1) \geq \psi(m, c_2)$ for all $m \in I$ and all $c_1, c_2 \in \mathcal{C}$ such that $u(c_1) \geq u(c_2)$.

4.3 Comparative statics analysis

This section presents a rigorous comparative statics analysis of the effects of memorable consumption on well-being. We propose alternative ways to compare two agents and determine if one of them is more sensitive to memorable experiences than the other. We thus lay out a foundation for studying well-being across individuals by recognizing the role of past experiences in affecting decision-making. Moreover, this inquiry proves useful for the purpose of developing parametric examples of the memory evolution function.

As standard in the literature on comparative statics, we start by disentangling the effects of memorable consumption from other unrelated determinants of decision-making. To this end, we consider agents that differ in their attitudes toward memorable consumption, while being identical for the other components of their tastes. That is, we require that they assess atemporal risk (including deterministic consumption bundles) and the passage of time in the same way. The following routine definition formalizes these assumptions.

Definition 3. Let $\succsim_1$ and $\succsim_2$ be relations on $\mathcal{L}$.

(i) We say that $\succsim_1$ and $\succsim_2$ have the same risk attitudes if

$$(p) \succsim_1 (q) \quad \Leftrightarrow \quad (p) \succsim_2 (q) \qquad \text{for all } p, q \in \mathcal{L}_1; \tag{8}$$

(ii) Provided that condition (8) holds, we say that $\succsim_1$ and $\succsim_2$ have the same time attitudes if

$$(0|p) \succsim_1 (q) \quad \Leftrightarrow \quad (0|p) \succsim_2 (q) \qquad \text{for all } p, q \in \mathcal{L}_1.$$

If $(\beta_1, u_1, I_1, \psi_1)$ and $(\beta_2, u_2, I_2, \psi_2)$ are representations of $\succsim_1$ and $\succsim_2$, respectively, as in Theorem 3, then they have the same risk attitude if and only if there exists $\lambda > 0$ such that $u_2 = \lambda u_1$; and, in addition, they have the same time attitude if and only if $\beta_2 = \beta_1$.

The following definition introduces the language that will be used to establish our comparative statics results. It therefore constitutes a key step for the subsequent development. It extends Definition 2 about the value of memory for a single agent to comparisons across different agents.

Definition 4. Let $\succsim_1$ and $\succsim_2$ be relations on $\mathcal{L}$ that have the same risk and time attitudes. For $f, g \in \mathcal{F}$, we say that f generates a higher value of memory for the next period for $\succsim_1$ in comparison to g for $\succsim_2$ if

$$g \succsim_2^{\mathtt{m}-k} z \Rightarrow f \succsim_1^{\mathtt{m}-k} z \qquad \text{for all } k > 0 \text{ and } z \in \mathcal{C}.$$

We denote the above relationship by $f \ {}^{\succsim_1}\mathcal{R}^{\succsim_2}\ g$. We also say that f generates a strictly higher value of memory for $\succsim_1$ in comparison to g for $\succsim_2$ if

$$g \succsim_2^{\mathtt{m}-k} z \Rightarrow f \succ_1^{\mathtt{m}-k} z \qquad \text{for all } k > 0 \text{ and } z \in \mathcal{C}.$$

We denote this relationship by $f \ {}^{\succsim_1}\mathcal{S}^{\succsim_2}\ g$. Finally, we say that f generates the same value of memory for $\succsim_1$ as g for $\succsim_2$ if

$$g \succsim_2^{\mathtt{m}-k} z \Leftrightarrow f \succsim_1^{\mathtt{m}-k} z \qquad \text{for all } k > 0 \text{ and } z \in \mathcal{C}.$$

We denote this relationship by $f \ {}^{\succsim_1}\mathcal{I}^{\succsim_2}\ g$.

We are now ready to study the comparative attitudes towards memorable consumption. Throughout, we focus on two agents respectively endowed with a consumption stream that has the same impact on their memory. That is, agents 1 and 2 face, respectively, non-degenerate streams f and g which generate the same value of memory for both of them.[15] Thus, our natural prerequisite to make meaningful comparisons is that $f \ {}^{\succsim_1}\mathcal{I}^{\succsim_2}\ g$.

Our first behavioral notion establishes when we can say that such an equivalent stock of memories has a more persistent effect for one agent compared to another.

[15] Note that it would not be sufficient to consider one common stream for both agents. From a subjective viewpoint, the same stream may give rise to different values of memory for different agents. By considering different streams and let individuals express their preferences, we fully take this aspect into account.

Definition 5. Let $\succsim_1$ and $\succsim_2$ be relations on $\mathcal{L}$ that have the same risk and time attitudes. For any $f, g \in \mathcal{F}$ such that $f \ {}^{\succsim_1}\mathcal{I}^{\succsim_2}\ g$, we have that:

(a) *Positive* memorable consumption has *longer effects* for $\succsim_2$ in comparison to $\succsim_1$ if

$$f\ \mathcal{R}^{\succsim_1}\ (0) \quad \Rightarrow \quad (g|0)\ {}^{\succsim_2}\mathcal{R}^{\succsim_1}\ (f|0)\ \mathcal{R}^{\succsim_1}\ (0).$$

(b) *Negative* memorable consumption has *longer effects* for $\succsim_2$ in comparison to $\succsim_1$ if

$$(0)\ \mathcal{R}^{\succsim_1}\ f \quad \Rightarrow \quad (0)\ \mathcal{R}^{\succsim_1}\ (f|0)\ {}^{\succsim_1}\mathcal{R}^{\succsim_2}\ (g|0).$$

Parts (a) and (b) provide symmetric definitions that distinguish between positive and negative experiences. Focus on part (a) first. The antecedent defines the sign of the stock of memory: if the stream f generates a higher value of memory than the neutral element (0) for agent 1, then it must be that f brings pleasant memories to her. Note that this assumption, together with $f \ {}^{\succsim_1}\mathcal{I}^{\succsim_2}\ g$, implies that g has a higher value of memory than (0) for agent 2, as well (i.e., $g\ \mathcal{R}^{\succsim_2}\ (0)$).[16] Thus, item (a) restricts attention to comparisons of agents with an equivalent baggage of positive experiences. Now, extend streams f and g by adding one last period of zero consumption. That is, consider streams $f|0$ and $g|0$: such profiles offer both agents to enjoy the memory of f and g for one extra period. This is the only difference in comparison to f and g. Then, we say that a positive memory has longer effects for agent 2 compared to agent 1 if $g|0$ generates a higher value of memory for agent 2 than $f|0$ does for agent 1. We interpret this pattern as evidence that g persists longer in agent 2's mind than does f for agent 1.

Part (b) presents a symmetric notion for unpleasant experiences. The implication now moves in the opposite direction to ensure that the sign of memory is negative. For a memory to qualify as negative, it must be the case that zero consumption (and hence zero memory) makes the agent better off. Moreover, the indices of 1 and 2 are reversed in the consequent because, when studying the effects of negative experiences, we seek to capture the greater absolute value of the effect.

Next proposition provides a characterization in terms of the Markovian representation and shows that comparative attitudes concerning memory persistency are determined by the Markovian function ψ.

[16]Conversely, note that positive memories necessarily satisfy the antecedent condition: that is, a stock of memory is (weakly) positive if no consumption bundle with positive utility can be dominated by memory.

Proposition 6. *Suppose that $\succsim_1$ and $\succsim_2$ are complete preorders on $\mathcal{L}$ that satisfy Axioms (A1)–(A7) and (A8) and have the same risk and time attitudes. Let (β, u, I_1, ψ_1) and (β, u, I_2, ψ_2) be their representations, as in Theorem 3. Then,*

(i) $\succsim_2$ exhibits longer effects of positive memory if and only if $\psi_2(m,0) \geq \psi_1(m,0)$ for all $m \in I_1 \cap I_2$ such that $m \geq 0$;

(ii) $\succsim_2$ exhibits longer effects of negative memory if and only if $\psi_2(m,0) \leq \psi_1(m,0)$ for all $m \in I_1 \cap I_2$ such that $m \leq 0$.

The above proposition characterizes when one can say that the same stock of memory has more lingering effects for one agent than another. It therefore provides a comparative measure of how long a memory persists in one agent's mind.

The comparative statics analysis can be performed along another dimension that reflects how easy it is for an agent to generate valuable memory. More precisely, we propose a ranking criterion that is based on the minimal level of consumption that is sufficient to maintain the stock of memory at a particular level. Since the stock of memory in the Markovian specification is measured in terms of its utility, this criterion can be interpreted as a way to compare the *strength* of the effects of memorable consumption — the smaller is the consumption that maintains a particular level of utility from memory, the stronger is the effect of memorable consumption on the agent's behavior.

Definition 6. Let $\succsim_1$ and $\succsim_2$ have the same risk and time attitudes. For any $f_1, f_2 \in \mathcal{F}$ such that $f_1 \,{}^{\succsim_1}\mathcal{I}^{\succsim_2}\, f_2$, we have that:

(a) *Positive* memorable consumption has *stronger effects* for $\succsim_2$ in comparison to $\succsim_1$ if $f_1 \,\mathcal{R}^{\succsim_1}\, (0)$, and

   (i) if $(f_1|x) \,\mathcal{R}^{\succsim_1}\, f_1$, $(f_2|y) \,\mathcal{R}^{\succsim_2}\, f_2$, and $(y) \succsim_2 (x) \succsim_2 (0)$ for some $x, y \in \mathcal{C}$, then $(f_1|y) \,\mathcal{R}^{\succsim_1}\, f_1$ and $(f_2|x) \,\mathcal{R}^{\succsim_2}\, f_2$;

   (ii) if $(f_1|x) \,\mathcal{S}^{\succsim_1}\, f_1$, $(g|y) \,\mathcal{S}^{\succsim_2}\, f_2$, and $(y) \succsim_2 (x) \succ_2 (0)$ for some $x, y \in \mathcal{C}$, then $(f_1|y) \,\mathcal{S}^{\succsim_1}\, f_1$ and $(f_2|x) \,\mathcal{S}^{\succsim_2}\, f_2$.

(b) *Negative* memorable consumption has *stronger effects* for $\succsim_2$ in comparison to $\succsim_1$ if $(0) \,\mathcal{R}^{\succsim_1}\, f_1$, and

   (i) if $f_1 \,\mathcal{R}^{\succsim_1}\, (f_1|x)$, $f_2 \,\mathcal{R}^{\succsim_2}\, (f_2|y)$, and $(0) \succsim_2 (x) \succsim_2 (y)$ for some $x, y \in \mathcal{C}$, then

$f_1 \, \mathcal{R}^{\succsim_1} \, (f_1|y)$ and $f_2 \, \mathcal{R}^{\succsim_2} \, (f_2|x)$;

(ii) if $f_1 \, \mathcal{S}^{\succsim_1} \, (f_1|x)$, $f_2 \, \mathcal{S}^{\succsim_2} \, (f_2|y)$, and $(0) \succ_2 (x) \succsim_2 (y)$ for some $x, y \in \mathcal{C}$, then $f_1 \, \mathcal{S}^{\succsim_1} \, (f_1|y)$ and $f_2 \, \mathcal{S}^{\succsim_2} \, (f_2|x)$.

Analogously to the previous notion of comparative longevity, also the above criterion focuses on the comparison of two streams that generate the same stock of memory for the agents, as revealed by the comparative relation ${}^{\succsim_1}\mathcal{I}^{\succsim_2}$. As before, part (a) considers the case in which such stock of memory represents positive experiences for agent 2 (and, hence, for agent 1, as well). Suppose that x and y are consumption bundles that make the extended streams $f_1|x$ and $f_2|y$ more valuable memory-wise than the original f_1 and f_2 for agents 1 and 2, respectively. Importantly, suppose that y is weakly preferred to x (by both agents, clearly). Then, the axiom prescribes that x and y are interchangeable as both of them should lead to an increase in the stock of memory for both individuals. Note that the definition applies to situations in which x and y are ranked in a non-expected way; this formalization rules out the case in which y is preferred to x, x and y increase the stock of memory of agents 1 and 2 respectively, and yet, x outperforms y in terms of generation of valuable memory (i.e., x fails to increase the value for the more sensitive agent 2). Part (a)(ii) repeats the same requirement for strict increases of the stock of memory, and Part (b) states a symmetric definition for negative stocks of memory and consumption levels (in which case the rankings of streams are reversed because what matters is the absolute magnitude of the effect).

The next proposition translates the behavioral notion of comparative strength into a parametric characterization.

Proposition 7. *Suppose that $\succsim_1$ and $\succsim_2$ are complete preorders on $\mathcal{L}$ that satisfy Axioms (A1)–(A7) and (A8) and have the same time and risk preferences. Let (β, u, I_1, ψ_1) and (β, u, I_2, ψ_2) be their representations as in Theorem 3.*

(a) For $i = 1, 2$, let $c_i^+ : I_i \rightrightarrows \mathbb{R}$ and $\mathring{c}_i^+ : I_i \rightrightarrows \mathbb{R}$ be correspondences (possibly, empty valued) defined as $c_i^+(m) = \{r \geq 0 : u(c) = r \text{ for some } c \in \mathcal{C} \text{ and } \psi_i(m, c) \geq m\}$ and $\mathring{c}_i^+(m) = \{r > 0 : u(c) = r \text{ for some } c \in \mathcal{C} \text{ and } \psi_i(m, c) > m\}$. Then, positive memory has stronger effects for $\succsim_2$ in comparison to $\succsim_1$ if and only if $c_2^+(m)$ and $\mathring{c}_2^+(m)$ are dominated by $c_1^+(m)$ and $\mathring{c}_1^+(m)$, respectively, in the strong set order for all $m \in I_1 \cap I_2 \cap \mathbb{R}_+$;

(b) For $i = 1, 2$, let $c_i^- : I_i \rightrightarrows \mathbb{R}$ and $\mathring{c}_i^- : I_i \rightrightarrows \mathbb{R}$ be correspondences (possibly, empty valued) defined as $c_i^-(m) = \{r \le 0 : u(c) = r$ for some $c \in \mathcal{C}$ and $\psi_i(m, c) \le m\}$ and $\mathring{c}_i^-(m) = \{r < 0 : u(c) = r$ for some $c \in \mathcal{C}$ and $\psi_i(m, c) < m\}$. Then, negative memory has stronger effects for $\succsim_2$ in comparison to $\succsim_1$ if and only if $c_2^-(m)$ and $\mathring{c}_2^-(m)$ dominate $c_1^-(m)$ and $\mathring{c}_1^-(m)$, respectively, in the strong set order for all $m \in I_1 \cap I_2 \cap \mathbb{R}_-$.

To illustrate the comparative statics notions developed in this section, we present two parametric rules for the evolution of memory and analyze the traits of longevity and strength in terms of the parameters of these rules.

Example 3. *Suppose that agents 1 and 2 are characterized by a Markovian memory representation with the same β and u, and their evolution functions ψ_i have the following linear autoregressive form:*

$$\psi_i(m, c) = \alpha_i m + (1 - \alpha_i) K_i u(c), \tag{9}$$

where $\alpha_i \in (0, 1)$ and $K_i > 0$ for $i = 1, 2$.

As follows from Proposition 6, the agent with longer positive memory has greater values of $\psi_i(m, 0) \equiv \alpha_i m$ for all non-negative values of m. Therefore, agent 2's positive memory has longer effect if and only if $\alpha_2 \ge \alpha_1$ (and if and only if agent 2's negative memory has longer effect). Note that the longevity of memory is controlled only by the parameter α and is unaffected by changes in K.

Next, we apply Proposition 7. The strength of the effect of positive memory is determined by the ordering of the set of solutions of inequalities $\alpha_i m + (1 - \alpha_i) K_i u(c) \ge m$ for $i = 1, 2$ — that is, $u(c) \ge \frac{1}{K_i} m$ — for all non-negative values of m, as well as the set of solutions of similar strict inequalities. Such a set for agent 2 is dominated if and only if $\frac{1}{K_2} \le \frac{1}{K_1}$. Hence, agent 2's positive memory has stronger effect if and only if $K_2 \ge K_1$. Similar to longevity, the strength of memory effect is controlled by one parameter, namely K, and this is independent of any changes in the other parameter, α.

Example 4. *Suppose that agents 1 and 2 are characterized by a Markovian memory representation with the same β and u, and their evolution functions ψ_i have the following form:*

$$\psi_i(m, c) = \alpha_i \max\{m, K_i u(c)\} + (1 - \alpha_i) K_i u(c), \tag{10}$$

where $\alpha_i \in [0,1]$ and $K_i > 0$ for $i = 1,2$.

As in the previous example, agent 2's positive memory has longer effect if and only if $\alpha_2 \geq \alpha_1$. Note that the case $\alpha_i = 1$ corresponds to maximum longevity. In this case, memory of a single positive experience never decays, and continues to contribute to the agent's per-period utility forever (or until a stronger positive experience occurs). To compare the strength of effects across agents, we need, again, to order the sets of solutions of inequalities $\alpha_i \max\{m, K_i u(c)\} + (1-\alpha_i)K_i u(c)\} \geq m$, as well as $\alpha_i \max\{m, K_i u(c)\} + (1-\alpha_i)K_i u(c)\} > m$, for $i = 1,2$. The solutions of the weak inequalities become trivial for $\alpha = 1$. However, weak and strict inequalities together unambiguously determine the way the agents are compared: agent 2's positive experience has stronger memory effect if and only if $K_2 \geq K_1$.

Specification (9) for the evolution function can be slightly generalized. Let $v : \mathbb{R} \to \mathbb{R}$ be a strictly increasing function with $v(0) = 0$ and $w : \mathcal{C} \to \mathbb{R}$. Then, we can consider the following law for the evolution of memory:

$$\psi(m,c) = v\big(\alpha v^{-1}(m) + (1-\alpha)w(c)\big),$$

where $\alpha \in (0,1)$.[17] In this case, we can change variables by defining $\tilde{m}_t = v^{-1}(m_t)$, and re-write representation (7) as

$$V(f) = \sum_{t=0}^{\ell(f)-1} \beta^t \big[u(f_t) + v(\tilde{m}_{t-1})\big], \tag{11}$$
$$\text{where} \quad \tilde{m}_\tau = \alpha \tilde{m}_{\tau-1} + (1-\alpha)w(f_\tau) \quad \text{for } \tau = 0, \ldots, \ell(f) - 2,$$
$$\tilde{m}_{-1} = 0.$$

Here, the stock of memory $\tilde{m}_t$ is measured in different "units," which leads to a specification convenient for applications: $AR(1)$-type law for memory and (potentially) non-linear function in the agent's objective.

5 A consumption-savings problem

This section presents an application of our Markovian memory specification by introducing memorable consumption into the standard linear-quadratic consumption-savings problem.

[17] Example 3 corresponds to $v(m) = m$ and $w(c) = K_i u(c)$.

Suppose that, in periods $t = 0, 1, 2 \ldots$, a consumer receives income y_t that is stochastic and iid across time. There are no borrowing constraints, so that she can reallocate income between periods by borrowing or saving at the gross interest rate $R > 0$. Time horizon is infinite and the future is discounted using discount factor $\beta \in (0, 1)$. For simplicity, we assume there is only one good ($N = 1$).[18] Utility from physical consumption is given by $u(c) = c - \frac{1}{2}c^2$; utility from consuming memories conforms to specification (11) with $v(\tilde{m}) = b\tilde{m} - \frac{1}{2}a\tilde{m}^2$, $a, b > 0$, and $w(c) = c$.[19] Thus, the consumer faces the following maximization problem:

$$\begin{aligned}
&\underset{\{c_t\}_{t=0}^{\infty}, \{s_t\}_{t=0}^{\infty}, \{m_t\}_{t=0}^{\infty}}{\text{maximize}} \ \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \left(c_t - \frac{1}{2}c_t^2 + b\tilde{m}_{t-1} - \frac{1}{2}a\tilde{m}_{t-1}^2 \right) \right] \qquad (12) \\
&\text{s.t. } c_t + s_t = y_t + Rs_{t-1} \qquad \text{for } t = 0, 1, \ldots, \\
&\qquad \tilde{m}_t = \alpha\tilde{m}_{t-1} + (1 - \alpha)c_t \quad \text{for } t = 0, 1, \ldots, \\
&\qquad s_{-1} \text{ is given,} \\
&\qquad \tilde{m}_{-1} = 0.
\end{aligned}$$

Finally, assume that $R = \frac{1}{\beta}$, $s_{-1} \geq 0$, and $\mathbb{E}[y] > 0$.

We are interested in extending the standard Permanent-Income-type solution of the problem to account for the role of memories.[20] Our goal is to identify some life-cycle implications of memorable consumption.

The Lagrangian of the problem is

$$\begin{aligned}
\mathcal{L} = \mathbb{E}_0 \Bigg[\sum_{t=0}^{\infty} \beta^t \Bigg(& c_t - \frac{1}{2}c_t^2 + b\tilde{m}_{t-1} - \frac{1}{2}a\tilde{m}_{t-1}^2 \\
& - \lambda_t \left(c_t + s_t - y_t - \frac{1}{\beta}s_{t-1} \right) - \mu_t(\tilde{m}_t - \alpha\tilde{m}_{t-1} - (1 - \alpha)c_t) \Bigg) \Bigg].
\end{aligned}$$

[18] There is no loss of generality in focusing on this special case. Allowing for multiple, ordinary or memorable, goods ($N > 1$) would not affect our analysis.

[19] Note that this example deviates from our specification (11) by considering infinite consumption streams and a non-monotone function v.

[20] Doing that, we will ignore the usual issues related to non-monotonicity of the utility from consumption and the form of conditions at infinity.

The First-Order Conditions are

$$
\begin{aligned}
&1 - c_t - \lambda_t + (1-\alpha)\mu_t = 0, \\
&-\lambda_t + \mathbb{E}_t[\lambda_{t+1}] = 0, \\
&\beta(b - a\tilde{m}_t) - \mu_t + \beta\alpha\mathbb{E}_t[\mu_{t+1}] = 0.
\end{aligned}
$$

By combining this system with the constraints, we eventually obtain the following solution:

$$
c_t = \left((1-\beta)y_t + \beta\mathbb{E}[y] + \frac{1-\beta}{\beta}s_{t-1}\right)(1+\kappa) - \kappa\tilde{m}_{t-1}, \tag{13}
$$

where $\kappa \geq 0$ is a constant given by

$$
\kappa = \frac{\sqrt{1 - 2\beta(\alpha^2 - a(1-\alpha)^2) + \beta^2(\alpha^2 + a(1-\alpha)^2)^2} - (1 - \beta(\alpha^2 - a(1-\alpha)^2))}{2\alpha(1-\beta\alpha)}.
$$
[21]

Expression (13) has an intuitive interpretation. If the effect of memorability is absent ($a = 0$ or $\alpha = 1$), then $\kappa = 0$ and we recover Hall's (1978) classic result that "consumption follows a random walk." In this case, the agent consumes the sum of the fraction $(1-\beta)$ of the income shock $y_t - \mathbb{E}[y]$, the average income $\mathbb{E}[y]$, and the interest from savings $\frac{1-\beta}{\beta}s_{t-1}$; the fraction β of her income shock and the body of the savings are kept as savings. In the presence of memorability ($a > 0$, $0 < \alpha < 1$), we observe that the agent exhibits a stronger reaction to income shocks and consumes more out of them ($\kappa > 0$). Albeit framed within a simplified setting, our finding suggests that memorable consumption may help explaining the well-known empirical evidence on excess sensitivity of consumption to income changes.[22] Moreover, the optimal level of consumption is negatively correlated with the stock of memory — it increases as the stock of memory decreases, and vice versa.

To further illustrate, assume for a moment that there is no income uncertainty and $y_t = \mathbb{E}[y]$ for all t. Then, the consumption rule can be rewritten as $c_t = \bar{c}_t + \kappa(\bar{c}_t - \tilde{m}_{t-1})$,

[21] This consumption rule is supported by

$$
\mu_t = \left((1-\beta)y_t + \beta\mathbb{E}[y] + \frac{1-\beta}{\beta}s_{t-1} - \tilde{m}_{t-1}\right)\kappa' + \frac{\beta}{1-\beta\alpha}(b - a\tilde{m}_{t-1}),
$$

where

$$
\kappa' = \frac{\sqrt{1 - 2\beta(\alpha^2 - a(1-\alpha)^2) + \beta^2(\alpha^2 + a(1-\alpha)^2)^2} - (1 - \beta(\alpha^2 - a(1-\alpha^2)))}{2\alpha(1-\beta\alpha)^2}
$$

[22] See, e.g., the surveys of Attanasio (1999) and Jappelli and Pistaferri (2010).

where $\bar{c}_t = \mathbb{E}[y] + \frac{1-\beta}{\beta} s_{t-1}$. In the standard linear-quadratic consumption-savings model, the expression for $\bar{c}_t$ corresponds to the permanent income. In our model, it becomes a reference that determines the level of consumption taking into account the accumulation of memory. If $\tilde{m}_{t-1} = \bar{c}_t$ then the agent is in a steady state and both her consumption and the stock of memory will stay constant; if $\tilde{m}_{t-1}$ exceeds $\bar{c}_t$ then she will consume less than $\bar{c}_t$ and opt for depleting part of her stock of memory; and, if $\tilde{m}_{t-1}$ has not reached $\bar{c}_t$ then she will consume more than $\bar{c}_t$ in order to build up her stock of memory. From the viewpoint of life-cycle profiles, this dynamics implies that agents tend to under-save and over-consume when they are young (as they start with $\tilde{m}_{-1} = 0 < \bar{c}_0$). As the stock of memory accumulates in subsequent periods, the gap will reduces and over-consumption will attenuate. If we compare consumption paths across agents, then those with higher κ over-consume more at young age, save less, and approach the steady state with lower savings. This behavior is rational, and can be interpreted as hidden savings in the form of investment in pleasant memories that substitutes for investment in financial assets. Furthermore, this dynamics may represent one key source of support for the empirical evidence according to which individuals consume too little at the retirement age compared to the predictions of the canonical model.[23]

The magnitude of the agent's excessive reaction to income shocks (relative to predictions of the standard model), as well as features of the life-cycle consumption pattern such as over-consumption when young, depend on the parameters of preferences through the value of κ. Holding everything else fixed, κ is an increasing function of the parameter a that, jointly with b, captures the strength of memorable effects of consumption. Hence, stronger memorable effect leads to greater over-consumption at young age and also stronger reaction to income shocks. As a function of the longevity of memory that is captured by the parameter α, κ has an inverse U-shape. Indeed, as α approaches one, memory becomes very persistent and is hardly affected by consumption. In the limit, the law of motion for memory takes the form of $\tilde{m}_t = \tilde{m}_{t-1}$, and additional investments in future memory are fruitless. On the other side of the range, as α approaches zero, memory loses its lasting effect, and the decision problem transforms into the standard question of consuming today

[23]It is worth mentioning that some works point to the role of demographics and changes in family size in explaining life-cycle consumption patterns (see, e.g., Fernández-Villaverde and Krueger (2007)). This seems consistent with our findings as a substantial portion of memorable expenditures might be household-related.

versus tomorrow. The effect of memorability is the greatest at intermediate levels of α.

Appendix

A Proof of the basic representation

Lemma 8. *Suppose that $\succsim$ is a preference relation on $\mathcal{F}$ that satisfies the Memory-Consumption Tradeoff Consistency and Continuity axioms. Then, for any $t \in \mathbb{N}$, $f, g \in \mathcal{F}_t$, and $p, q \in \mathcal{L}_1$, we have*

$$f|p \succsim g|p \quad \Leftrightarrow \quad f|q \succsim g|q.$$

Proof. First, we claim that, for any $n \in \mathbb{N}$, and for any $t \in \mathbb{N}$, $f, g \in \mathcal{F}_t$, and $p, q \in \mathcal{L}_1$,

$$f|p \succsim g|(\tfrac{n-1}{n}p + \tfrac{1}{n}q) \quad \Leftrightarrow \quad f|(\tfrac{1}{n}p + \tfrac{n-1}{n}q) \succsim g|q. \tag{14}$$

Indeed, for $n = 1$, this statement is a triviality. Suppose that it holds for some $n \in \mathbb{N}$, and that $f|p \succsim g|(\frac{n}{n+1}p + \frac{1}{n+1}q)$ for some $t \in \mathbb{N}$, $f, g \in \mathcal{F}_t$, and $p, q \in \mathcal{L}_1$. Let $q' := \frac{1}{n+1}p + \frac{n}{n+1}q$. Note that $\frac{n-1}{n}p + \frac{1}{n}q' = \frac{n}{n+1}p + \frac{1}{n+1}q$ and, hence, it follows from assumptions that $f|(\frac{1}{n}p + \frac{n-1}{n}q') \succsim g|q'$. Now, observe that q' is the midpoint between $\frac{1}{n}p + \frac{n-1}{n}q'$ and q. Therefore, by Memory-Consumption Tradeoff Consistency, $f|q' \succsim f|q$, which completes the inductive step.

Now, the claim of the lemma follows from (14) by taking the limit $n \to \infty$. Indeed, fix arbitrary $t \in \mathbb{N}$, $f, g \in \mathcal{F}_t$, and $p, q \in \mathcal{L}_1$. If $f|p \succ g|p$ then, by continuity, for all sufficiently large n we have $f|p \succ g|(\frac{n-1}{n}p + \frac{1}{n}q)$, which gives $f|(\frac{1}{n}p + \frac{n-1}{n}q) \succ g|q$ by the previous step and, in the limit as $n \to \infty$, $f|q \succsim g|q$. If $g|p \succ f|p$, then the claim similarly holds. By the symmetry of the claim with respect to renaming p and q, the only remaining case is $f|p \sim g|p$ and $f|q \sim g|q$, in which the claimed equivalence holds, as well. □

Lemma 9. *Let X be a connected separable topological space, Y a convex subset of a separable topological vector space, and $\succcurlyeq$ a continuous complete preorder on $X \times Y$ that has the following properties:*

(i) There exist $x, x', x_0 \in X$ and $y, y', y_0 \in Y$ such that $(x, y_0) \succ (x', y_0)$ and $(x_0, y) \succ (x_0, y')$.

(ii) For all $x, x' \in X$ *and* $y, y' \in Y$*,* $(x,y) \succcurlyeq (x',y) \Rightarrow (x,y') \succcurlyeq (x',y')$*.*

(iii) For all $x, x' \in X$ *and* $y, y' \in Y$*,* $(x,y) \succcurlyeq (x,y') \Rightarrow (x',y) \succcurlyeq (x',y')$*.*

(iv) For all $x, x' \in X$ *and* $y, y' \in Y$*,* $(x,y) \succcurlyeq (x', \frac{1}{2}y + \frac{1}{2}y') \Leftrightarrow (x, \frac{1}{2}y + \frac{1}{2}y') \succcurlyeq (x',y')$*.*

Then, there exist a continuous function $U_x : X \to \mathbb{R}$ *and a continuous affine function* $U_y : Y \to \mathbb{R}$ *such that*

$$(x,y) \succcurlyeq (x',y') \quad \Leftrightarrow \quad U_x(x) + U_y(y) \geq U_x(x') + U_y(y').$$

Proof. To verify the conditions of Wakker (1989, Th. III.4.1), observe that the assumptions of the lemma immediately guarantee the existence of two essential coordinates and that the coordinate independence property is satisfied. It remains to show that the hexagon condition holds. Indeed, suppose that $x_1, y_1, v_1 \in X$ and $a_2, b_2, c_2 \in Y$ are such that $(y_1, a_2) \asymp (x_1, b_2)$[24] and $(v_1, a_2) \asymp (y_1, b_2) \asymp (x_1, c_2)$. Let $d = \frac{1}{2}a_2 + \frac{1}{2}c_2$. We claim that $(x_1, d) \asymp (x_1, b_2)$. If $(y_1, a_2) \succ (x_1, d)$ then, by (iv), $(y_1, d) \succ (x_1, c_2)$. By transitivity, $(x_1, b_2) \succ (x_1, d)$ and $(y_1, d) \succ (y_1, b_2)$, a contradiction with (iii). The situation $(x_1, d) \succ (y_1, a_2)$ similarly leads to a contradiction. We conclude that $(x_1, d) \asymp (y_1, a_2) \asymp (x_1, b_2)$. By the assumption and (iii), we have $(v_1, a_2) \asymp (y_1, b_2) \asymp (y_1, d)$. Then, $(v_1, d) \asymp (y_1, c_2)$ by (iv) and $(v_1, d) \asymp (v_1, b_2)$ by (iii), which gives by transitivity $(v_1, b_2) \asymp (y_1, c_2)$, and the hexagon condition is proven.

Now, we can apply Wakker (1989, Th. III.4.1) to obtain that there exist nonconstant continuous functions $U_x : X \to \mathbb{R}$ and $U_y : Y \to \mathbb{R}$ such that

$$(x,y) \succcurlyeq (x',y') \quad \Leftrightarrow \quad U_x(x) + U_y(y) \geq U_x(x') + U_y(y'). \tag{15}$$

It remains to show that U_y must be affine. Indeed, (15) and property (iv) imply that, for any $y, y' \in Y$,

$$U_y(y) - U_y(\tfrac{1}{2}y + \tfrac{1}{2}y') \geq U_x(x') - U_x(x) \Leftrightarrow U_y(\tfrac{1}{2}y + \tfrac{1}{2}y') - U_y(y') \geq U_x(x') - U_x(x) \quad \text{for all } x, x' \in X. \tag{16}$$

Fix an arbitrary $[a,b] \subseteq range\, U_x$, where $a < b$, and let $\varepsilon \in (0, b-a)$. Then, the arbitrariness of x and x' in (16) gives that, for any $y, y' \in Y$ such that $|U_y(y) - U_y(y')| \leq \varepsilon$,

$$U_y(y) - U_y(\tfrac{1}{2}y + \tfrac{1}{2}y') = U_y(\tfrac{1}{2}y + \tfrac{1}{2}y') - U_y(y').$$

[24] Here, we use $\asymp$ to denote the symmetric part of $\succcurlyeq$.

Applying it repeatedly, this equation can be extended to all $y, y' \in Y$. Moreover, it can be re-written as $U_y(\frac{1}{2}y + \frac{1}{2}y') = \frac{1}{2}U_y(y) + \frac{1}{2}U_y(y')$. By continuity, it implies that U_y is affine. □

Proof of Theorem 1. *Only if* part. Suppose that $\succsim$ is a complete preorder on $\mathcal{L}$ that satisfies Axioms (A1)–(A7). Let z_t for $t \in \mathbb{N}$ denote an element of $\mathcal{F}_t$ such that $z_t = (0, 0, \dots, 0)$.

Step 1. On the subset $\mathcal{L}_1 \subset \mathcal{L}$, $\succsim$ admits an expected utility representation: there exists a continuous and bounded function $u : \mathcal{C} \to \mathbb{R}$ such that $p \succsim q \Leftrightarrow \mathbb{E}_p[u] \geq \mathbb{E}_q[u]$ for all $p, q \in \mathcal{L}_1$. Let u be normalized such that $u(0) = 0$. Moreover, Nondegeneracy directly implies that range u admits both positive and negative values.

Step 2. By Independence, Continuity (item (2)), and Nondegeneracy, we have that the conditions of the mixture space theorem (?) are satisfied and, therefore, there exists an affine function $V : \mathcal{L} \to \mathbb{R}$ that represents $\succsim$ on $\mathcal{L}$: $P \succsim Q \Leftrightarrow V(P) \geq V(Q)$ for all $P, Q \in \mathcal{L}$. By the uniqueness of the expected utility representation on $\mathcal{L}_1$, it must be that the restriction of V to $\mathcal{L}_1$ is a positive affine transformation of the mapping $p \mapsto \mathbb{E}_p[u]$ for $p \in \mathcal{L}_1$. Normalizing if necessary, assume that $V(p) = \mathbb{E}_p[u]$ for all $p \in \mathcal{L}_1$. Note that, by the continuity axiom, V must be continuous when restricted to convex sets $\mathcal{L}_t$ for all $t \in \mathbb{N}$.

Step 3. Risk Preference Consistency and Stationarity imply that, for all $f, g \in \mathcal{F}$ and $p, q \in \mathcal{L}_1$,

$$f|p \succsim f|q \quad \Leftrightarrow \quad g|p \succsim g|q \quad \Leftrightarrow \quad p \succsim q.$$

Hence, by the uniqueness of the expected utility representation, it must be that for all $t \in \mathbb{N}$ there exist $\alpha_t : \mathcal{F}_t \to \mathbb{R}$ and $\beta_t : \mathcal{F}_t \to \mathbb{R}_{++}$ such that

$$V(f|p) = \alpha_t(f) + \beta_t(f)\mathbb{E}_p[u] \qquad \text{for all } f \in \mathcal{F}_t \text{ and } p \in \mathcal{L}_1.$$

Step 4. This step establishes an alternative representation for $\succsim$ restricted to $\mathcal{F}_t \times \mathcal{L}_1$ for all $t \in \mathbb{N}$: we claim that there exist continuous functions $W_t : \mathcal{F}_t \to \mathbb{R}$ such that

$$f|p \succsim g|q \quad \Leftrightarrow \quad W_t(f) + \mathbb{E}_p[u] \geq W_t(g) + \mathbb{E}_q[u]$$

for all $f, g \in \mathcal{F}_t$ and $p, q \in \mathcal{L}_1$. Indeed, Nondegeneracy and Stationarity, Risk Preference Consistency, Lemma 8, and Memory-Consumption Tradeoff Consistency make the conditions of Lemma 9 hold. Therefore, there exist continuous $W_t : \mathcal{F}_t \to \mathbb{R}$ and continuous affine $W'_t : \mathcal{L}_1 \to \mathbb{R}$ such that $f|p \succsim g|q \Leftrightarrow W_t(f) + W'_t(p) \geq W_t(g) + W'_t(q)$ for all $f, g \in \mathcal{F}_t$ and

$p, q \in \mathcal{L}_1$. Then, as follows from Risk Preference Consistency, $p \succsim q \Leftrightarrow W'_t(p) \geq W'_t(q)$ for all $p, q \in \mathcal{L}_1$. Hence, by the uniqueness of the expected utility representation, it must be that, for all $t \in \mathbb{N}$, W'_t are positive affine transformations of our representation of $\succsim$ restricted to $\mathcal{L}_1$ obtained in Step 1. Then, normalizing if necessary, we can assume that, for all $t \in \mathbb{N}$, $W'_t(p) = \mathbb{E}_p[u]$ for all $p \in \mathcal{L}_1$.

Step 5. For all $t \in \mathbb{N}$, the range of the mapping $(f, p) \mapsto W_t(f) + \mathbb{E}_p[u]$ is convex, and, therefore, there must exist continuous and strictly increasing functions $\zeta_t : \mathbb{R} \to \mathbb{R}$ such that $V(f|p) = \alpha_t(f) + \beta_t(f)\mathbb{E}_p[u] = \zeta_t(W_t(f) + \mathbb{E}_p[u])$ for all $f \in \mathcal{F}_t$ and $p \in \mathcal{L}_1$. Observe that, for any $t \in \mathbb{N}$ and any fixed $f \in \mathcal{F}_t$, the left-hand side of this equality is an affine function of $p \in \mathcal{L}_1$. Hence, ζ_t must be positive affine functions for all $t \in \mathbb{N}$: $\zeta_t(t) = A_t + B_t t$ for some $A_t \in \mathbb{R}$ and $B_t \in \mathbb{R}_{++}$. If we let $\tilde{W}_t(f) := A_t + B_t W_t(f)$ for all $t \in \mathbb{N}$ and $f \in \mathcal{F}_t$, we obtain:

$$V(f|p) = \tilde{W}_t(f) + B_t \mathbb{E}_p[u] \qquad \text{for all } t \in \mathbb{N},\ f \in \mathcal{F}_t, \text{ and } p \in \mathcal{L}_1. \tag{17}$$

Step 6. The Impatience axiom asserts that $(c) \succ (0, c) \succ (0)$ for all $c \in \mathcal{C}$ such that $(c) \succ (0)$. By taking the limit $c \to 0$ in the above and using continuity, we obtain $(0, 0) \sim (0)$. Using Stationarity and mathematical induction, it can be seen that $z_t \sim (0)$ and $V(z_t) = 0$; in turn, by (17), we also have $\tilde{W}_t(z_t) = 0$.

Let $\beta := B_1$ and note that Impatience implies that $\beta < 1$. For any $c \in \mathcal{C}$, let $p \in \mathcal{L}_1$ be defined as $p := \beta\delta_c + (1 - \beta)\delta_0$ and observe that $V(0, c) = \beta u(c) = \mathbb{E}_p[u] = V(p)$ by (17). For any $t \in \mathbb{N}$, Stationarity gives that $z_t|0|c \sim z_t|p$ and, hence, by (17), $B_{t+1}u(c) = B_t\mathbb{E}_p[u] = B_t\beta u(c)$. Since c was arbitrarily chosen, we have that $B_t = \beta^t$ for all $t \in \mathbb{N}$ and

$$V(f|p) = \tilde{W}_t(f) + \beta^t \mathbb{E}_p[u] \qquad \text{for all } t \in \mathbb{N},\ f \in \mathcal{F}_t, \text{ and } p \in \mathcal{L}_1. \tag{18}$$

This equation holds also for $t = 0$ by letting $\tilde{W}_0 := 0$.

Step 7. Let $M_0 : \mathcal{F}_0 \to \mathbb{R}$ be zero and $M_t : \mathcal{F}_t \to \mathbb{R}$ for $t \in \mathbb{N}$ be defined as

$$M_t(f_{t-1}, \ldots, f_0) := \beta^{-t}(\tilde{W}_t(f_0, \ldots, f_{t-1}) - V(f_0, \ldots, f_{t-1})).$$

Using this definition in (18), we obtain that, for all $t \in \mathbb{N}$ and $f \in \mathcal{F}_{t+1}$,

$$V(f_0, \ldots, f_{t-1}, f_t) = V(f_0, \ldots, f_{t-1}) + \beta^t M_t(f_{t-1}, \ldots, f_0) + \beta^t u(f_t).$$

Then, for all $t \in \mathbb{N} \cup \{0\}$,

$$V(f_0, \ldots, f_t) = \sum_{\tau=0}^{t} \beta^\tau \big[u(f_\tau) + M_\tau(f_{\tau-1}, \ldots, f_0)\big] \qquad \text{for all } f \in \mathcal{F}_{t+1}.$$

Step 8. We clam that, for all $t \in \mathbb{N}$ and $P \in \mathcal{L}$, $V(z_t|P) = \beta^t V(P)$. Indeed, fix $t \in \mathbb{N}$, and observe that, by Stationarity, both $P \mapsto V(P)$ and $P \mapsto V((0)|P)$ are representations of the restriction of $\succsim$ to $\mathcal{L}_t$. Hence, by uniqueness of affine representations, there exists $b > 0$ such that $V((0)|P) = bV(P)$ for all $P \in \mathcal{L}_t$. As follows from (18), it must be that $b = \beta$. The claim now follows by induction. Note that, by (18), we also have $\tilde{W}(z_t|f) = \beta^t \tilde{W}(f)$ for all $t \in \mathbb{N}$ and $f \in \mathcal{F}$.

Step 9. Now, we can define $M : \mathcal{C}_0^\infty \to \mathbb{R}$ for all $h \in C_0^\infty$ by letting $M(h) = M_l(h)$ for an arbitrary $l \in \mathbb{N}$ such that $h_\tau = 0$ for all $\tau \geq l$. (Note that, by the result of the previous step, this definition does not depend on the choice of l.) Then, for all $t \in \mathbb{N} \cup \{0\}$,

$$V(f_0, \ldots, f_t) = \sum_{\tau=0}^{t} \beta^\tau \big[u(f_\tau) + M(f_{\tau-1}, \ldots, f_0, 0, 0, \ldots) \big] \qquad \text{for all } f \in \mathcal{F}_{t+1}.$$

Representation (3) is now proven. Note that M is continuous since all M_l for $l \in \mathbb{N}$ are continuous due to the continuity of V and $\tilde{W}_l$.

If part. Suppose that $\succsim$ admits a utility representation via a function V as specified in (3). We next show that the axioms hold.

Stationarity. Let $0 \in \mathcal{C}$ denote an element that is mapped by u into $0 \in \mathbb{R}$. Then, equation (3) gives $V((0)|P) = \beta V(P)$ for all $P \in \mathcal{L}$, which implies that $(0)|P \succsim (0)|Q \Leftrightarrow P \succsim Q$ for all $P, Q \in \mathcal{L}$.

Impatience. If $(c) \succ (0)$ for some $c \in \mathcal{C}$, then, by (3), $u(c) > 0$. Then, $V(c) = u(c) > \beta u(c) = V(0, c) > 0$.

Independence. Follows directly from the representation.

Risk Preference Consistency. For any $f, g \in \mathcal{F}$ and $p, q \in \mathcal{L}_1$, we have by (3) that $f|p \succsim f|q \Leftrightarrow \mathbb{E}_p[u] \geq \mathbb{E}_q[u] \Leftrightarrow g|p \succsim g|q$.

Memory-Consumption Tradeoff Consistency. Define $S : \mathcal{F}_t \to \mathbb{R}$ as

$$S(f) := \sum_{\tau=0}^{t-1} \beta^\tau \big[u(f_\tau) + M(f_{\tau-1}, \ldots, f_0, 0, 0, \ldots) \big] + \beta^\tau M(f_{\tau-1}, \ldots, f_0, 0, 0, \ldots).$$

Then, for any $t \in \mathbb{N}$, $f, g \in \mathcal{F}_t$, and $p, q \in \mathcal{L}_1$, we have by (3) that

$$\begin{aligned} f|p \succsim g|(\tfrac{1}{2}p + \tfrac{1}{2}q) \quad &\Leftrightarrow \quad S(f) + \beta^t \mathbb{E}_p[u] \geq S(g) + \beta^t \left(\tfrac{1}{2}\mathbb{E}_p[u] + \tfrac{1}{2}\mathbb{E}_q[u] \right) \quad \Leftrightarrow \\ &S(f) - S(g) \geq \beta^t \left(\tfrac{1}{2}\mathbb{E}_q[u] - \tfrac{1}{2}\mathbb{E}_p[u] \right) \quad \Leftrightarrow \\ S(f) + \beta^t \left(\tfrac{1}{2}\mathbb{E}_p[u] + \tfrac{1}{2}\mathbb{E}_q[u] \right) &\geq S(g) + \beta^t \mathbb{E}_q[u] \quad \Leftrightarrow \quad f|(\tfrac{1}{2}p + \tfrac{1}{2}q) \succsim g|q. \end{aligned}$$

Continuity. For each $t \in \mathbb{N}$, the mapping $\mathcal{F}_t \to \mathbb{R}$ defined as $f \mapsto \sum_{\tau=0}^{t-1} \beta^\tau \big[u(f_\tau) + M(f_{\tau-1}, \dots, f_0, 0, \dots)\big]$ is continuous by the continuity of u and M; hence, V defined by (3) is continuous on $\mathcal{L}_t$, which establishes the first part of the axiom. The second part holds immediately.

Nondegeneracy. The first part holds directly since the range of u take both positive and negative values. □

Proof of Proposition 2. Let (β, u, M) and $(\hat{\beta}, \hat{u}, \hat{M})$ represent the same binary relation $\succsim$ on $\mathcal{L}$ as in Theorem 1. By Wakker (1989, Obs. III.6.6′), there exist $\lambda > 0$ and $d, d' \in \mathbb{R}$ such that $\hat{u} = \lambda u + d$, and $\hat{M} = \lambda M + d'$. As required by Theorem 1, it must be that $u(0) = 0 = \hat{u}(0)$. Thus, $d = 0 = d'$, implying that $\hat{u} = \lambda u$ and $\hat{M} = \lambda M$. Moreover, it clearly must be that $\beta = \hat{\beta}$ for the two triples to represent the same binary relation. The sufficiency of the conditions can be directly verified. □

B Proofs of Theorem 3 and Related Results

We start with a preliminary lemma that will be useful to prove Theorem 3.

Lemma 10. *Suppose that a complete preorder $\succsim$ on $\mathcal{L}$ satisfies Axioms (A1)–(A7), and let (β, u, M) be its representation as in Theorem 1. Then, for all $z \in \mathcal{C}$ and $k > 0$,*

$$(f|0) \succsim \frac{1}{k+1} f + \frac{k}{k+1}(f|z) \quad \Leftrightarrow \quad M(f_{\ell(f)-1}, \dots, f_0, 0, \dots) \geq k u(z).$$

Proof. Let $t := \ell(f)$. Using representation (3), we obtain

$$\begin{aligned}
&(f|0) \succsim \frac{1}{k+1} f + \frac{k}{k+1}(f|z) \quad \Leftrightarrow \\
&\quad V(f) + \beta^t M(f_{t-1}, \dots, f_0, 0, \dots) \geq \\
&\qquad \frac{1}{k+1} V(f) + \frac{k}{k+1}\big[V(f) + \beta^t u(z) + \beta^t M(f_{t-1}, \dots, f_0, 0, \dots)\big] \quad \Leftrightarrow \\
&\qquad\qquad M(f_{t-1}, \dots, f_0, 0, \dots) \geq k u(z).
\end{aligned}$$

□

Proof of Theorem 3. *Only if* part. Suppose that $\succsim$ is a complete preorder on $\mathcal{L}$ that satisfies the specified axioms.

Date: 2018-07-09 14:54:50 Revision: 3abf17a

Step 1. Let $V : \mathcal{L} \to \mathbb{R}$ be a utility representation of $\succsim$ as in (3) with β, u, and M as specified in Theorem 1. Let $I := \{M(f_{t-1}, \ldots, f_0, 0, \ldots) \mid f \in \mathcal{F}_t, t \in \mathbb{N}\}$ and note that I is a bounded interval of $\mathbb{R}$ that contains 0 bacause M is bounded and $M(0, 0, \ldots) = 0$. Define $\psi : I \times \mathcal{C} \to I$ as follows: For $r \in \mathbb{R}$ and $c \in \mathcal{C}$, $\psi(r, c) := M(c, f_{t-1}, \ldots, f_0, 0, \ldots)$, where $t = \ell(f)$ and $f \in \mathcal{F}$ is an arbitrary act such that $M(f_{t-1}, \ldots, f_0, 0, \ldots) = r$.

Step 2. We claim that, in the above definition of ψ, the value of $\psi(r, c)$ is independent of the choice of f. Indeed, fix an arbitrary $c \in \mathcal{C}$, and let $f \in \mathcal{F}$ and $f' \in \mathcal{F}$ be such that $M(f_{t-1}, \ldots, f_0, 0, \ldots) = M(f'_{t'-1}, \ldots, f'_0, 0, \ldots)$, where $t = \ell(f)$ and $t' = \ell(f')$. Let $\hat{f} \in \mathcal{F}_{t-1}$ and $\hat{f}' \in \mathcal{F}_{t'-1}$ be the truncated streams: $f = \hat{f}|f_{t-1}$ and $f' = \hat{f}'|f'_{t-1}$. We have

$$M(f_{t-1}, \ldots, f_0, 0, \ldots) \geq ku(z) \Leftrightarrow M(f'_{t'-1}, \ldots, f'_0, 0, \ldots) \geq ku(z) \quad \text{for all } z \in \mathcal{C} \text{ and } k > 0,$$

and, therefore, by Lemma 10,

$$\hat{f}|f_{t-1} \succsim^{\mathtt{m}-k} z \quad \Leftrightarrow \quad \hat{f}'|f'_{t-1} \succsim^{\mathtt{m}-k} z \qquad \text{for all } z \in \mathcal{C} \text{ and } k > 0.$$

By Axiom A8, we have

$$f|c \succsim^{\mathtt{m}-k} z \quad \Leftrightarrow \quad f'|c \succsim^{\mathtt{m}-k} z \qquad \text{for all } z \in \mathcal{C} \text{ and } k > 0.$$

Applying Lemma 10 again, and since z and k are arbitrary, we conclude that $M(c, f_{t-1}, \ldots, f_0, 0, \ldots) = M(c, f'_{t'-1}, \ldots, f'_0, 0, \ldots)$.

Step 3. Now, we show that $\operatorname{range}\psi = I$. Let $r \in I$ be chosen arbitrarily. By definition, $r = M(f_{t-1}, \ldots, f_0, 0, \ldots)$ for some $t \in \mathbb{N}$ and $f \in \mathcal{F}_t$. Let $\tilde{r} = M(f_{t-2}, \ldots, f_0, 0, \ldots)$, and observe that $\psi(\tilde{r}, f_{t-1}) = r$ by the result of the previous step. Hence, $r \in \operatorname{range}\psi$.

Step 4. Finally, we prove that ψ is continuous. Suppose, by contradiction, that ψ is not continuous: There exist sequences $\{r_n\}_{n=1}^{\infty}$ in I and $\{c_n\}_{n=1}^{\infty}$ in $\mathcal{C}$ such that $r_n \to r \in I$, $c_n \to c \in \mathcal{C}$, $\psi(r_n, c_n) \to K \in \mathbb{R} \cup \{-\infty, +\infty\}$ as $n \to \infty$, but $K \neq \psi(r, c)$. Passing to a subsequence, we can assume that the sequence $\{r_n\}_{n=1}^{\infty}$ is either increasing or decreasing.

Note that $I = \bigcup_{t=1}^{\infty} I_t$, where $I_t = \{M(f_{t-1}, \ldots, f_0, 0, \ldots) \mid f \in \mathcal{F}_t\}$. Recall that M is continuous; for each $t \in \mathbb{N}$, $\mathcal{F}_t$ is connected and, hence, I_t is an interval; moreover, $0 \in I_t$. Therefore, we can find some $t \in \mathbb{N}$ such that $r \in I_t$ and $r_n \in I_t$ for all $n \in \mathbb{N}$. Let $f^{(1)}$ and f in $\mathcal{F}_t$ be such that $M\left(f^{(1)}_{t-1}, \ldots, f^{(1)}_0, 0, \ldots\right) = r_1$ and $M\left(f_{t-1}, \ldots, f_0, 0, \ldots\right) = r$. For $n \in \mathbb{N}$, $n \geq 2$, let $f^{(n)} := (1 - \gamma_n)f^{(1)} + \gamma_n f$, where, for each $n \in \mathbb{N}$, $n \geq 2$, γ_n is chosen such that

$M\left(f_{t-1}^{(n)},\dots,f_0^{(n)},0,\dots\right)=r_n$, which is possible by continuity. Passing to a subsequence, $\{\gamma_n\}_{n=1}^{\infty}$ converges, and, hence, $\{f^{(n)}\}_{n=1}^{\infty}$ converges to some $f^{(\infty)}\in\mathcal{F}_t$. Observe that $r=\lim_{n\to\infty}r_n=\lim_{n\to\infty}M\left(f_{t-1}^{(n)},\dots,f_0^{(n)},0,\dots\right)=M\left(f_{t-1}^{(\infty)},\dots,f_0^{(\infty)},0,\dots\right)$ by the continuity of M. By the result of Step 2, we have $\psi(r_n,c_n)=M\left(c_n,f_{t-1}^{(n)},\dots,f_0^{(n)},0,\dots\right)$ for all $n\in\mathbb{N}$ and $\psi(r,c)=M\left(c,f_{t-1}^{(\infty)},\dots,f_0^{(\infty)},0,\dots\right)$; by the continuity of M, we have $\lim_{n\to\infty}M\left(c_n,f_{t-1}^{(n)},\dots,f_0^{(n)},0,\dots\right)=M\left(c,f_{t-1}^{(\infty)},\dots,f_0^{(\infty)},0,\dots\right)$; and obtain $\lim_{n\to\infty}\psi(r_n,c_n)=\psi(r,c)$, a contradiction to our assumption.

If part. Assume that there exist a scalar $\beta\in(0,1)$, a function $u:\mathcal{C}\to\mathbb{R}$, and a function $\psi:I\times\mathcal{C}\to I$ for some interval $I\subseteq\mathbb{R}$ as described in the theorem such that $V(P)=\sum_{f\in\operatorname{supp}P}P(f)V(f)$ for all $P\in\mathcal{L}$, where $V(f)$ is computed as in (7).

Let $M:\mathcal{C}_0^{\infty}\to\mathbb{R}$ be defined as follows. For $h=(h_0,h_1,\dots,h_{t-1},0,\dots)\in\mathcal{C}_0^{\infty}$, let $m_{-1}=0$; for $\tau=0,\dots,t-1$, $m_\tau=\psi(m_{\tau-1},h_\tau)$; and, finally, $M(h)=m_{t-1}$. Note that in this construction the value of $M(h)$ does not depend on the choice of t as long as $h_\tau=0$ for all $\tau\geq t$.

Now, we establish the continuity of M. Suppose that a net $\{h^{(\alpha)}\}_\alpha$ converges to some h in $\mathcal{C}_0^{\infty}$. Hence, for some $T\in\mathbb{N}$ such that $h_t=0$ for all $t\geq T$, there exist an index α_0 such that $h_t^{(\alpha)}=0$ for all $\alpha\geq\alpha_0$ and $t\geq T$, and $\sup_{0\leq t\leq T}|h_t-h_t^{(\alpha)}|$ converges to zero. Then, $M\left(h^{(\alpha)}\right)=\psi\left(\psi\left(\dots\psi\left(0,h_{T-1}^{(\alpha)}\right),\dots,h_1^{(\alpha)}\right),h_0^{(\alpha)}\right)\to M(h)=\psi(\psi(\dots\psi\left(0,h_{T-1}\right),\dots,h_1),h_0)$ because of the continuity of ψ.

Clearly, M satisfies the normalization condition and it is bounded because ψ is bounded. Thus, Axioms (A1)–(A7) hold by Theorem 1. It remains to show that Axiom (A8) holds, as well.

Suppose that $f,g\in\mathcal{F}$ and $x,y\in\mathcal{C}$ are such that

$$f|x\succsim^{\mathsf{m}-k}z\quad\Leftrightarrow\quad g|y\succsim^{\mathsf{m}-k}z\qquad\text{for all } z\in\mathcal{C}\text{ and } k>0.$$

By Lemma 10, this gives

$$M(x,f_{\ell(f)-1},\dots,f_0,0,\dots)\geq ku(z)\Leftrightarrow M(y,g_{\ell(g)-1},\dots,g_0,0,\dots)\geq ku(z)\quad\forall z\in\mathcal{C},k>0.$$

Due to the arbitrariness of z and k and the fact that $\operatorname{range}u$ takes both positive and negative values, it must be that $M(x,f_{\ell(f)-1},\dots,f_0,0,\dots)=M(y,g_{\ell(g)-1},\dots,g_0,0,\dots)$. Fix an arbitrary $c\in\mathcal{C}$. Then, $M(c,x,f_{\ell(f)-1},\dots,f_0,0,\dots)=\psi(M(x,f_{\ell(f)-1},\dots,f_0,0,\dots),c)=$

$\psi(M(y, g_{\ell(g)-1}, \ldots, g_0, 0, \ldots), c) = M(c, y, g_{\ell(g)-1}, \ldots, g_0, 0, \ldots)$. By Lemma 10 again,

$$f|x|c \succsim^{\mathtt{m}-k} z \quad \Leftrightarrow \quad g|y|c \succsim^{\mathtt{m}-k} z \qquad \text{for all } z \in \mathcal{C} \text{ and } k > 0.$$

□

Proof of Proposition 4. We will prove the necessity by using the fact that Theorem 3 is a special case of the general representation in Theorem 1. Let (β, u, I, ψ) and $\left(\hat{\beta}, \hat{u}, \hat{I}, \hat{\psi}\right)$ represent the same binary relation $\succsim$ on $\mathcal{L}$ as in Theorem 3.

Define $M : \mathcal{C}_0^\infty \to \mathbb{R}$ recursively via ψ in the same way as in the proof of the "if" part of Theorem 3, and, similarly, $\hat{M}$ via $\hat{\psi}$. As pointed out in that proof, such functions M and $\hat{M}$ satisfy the properties of Theorem 1. By the uniqueness result for the general representation (Proposition 2), $\beta = \hat{\beta}$ and there exists $\lambda > 0$ such that $\hat{u} = \lambda u$ and $\hat{M} = \lambda M$. Now, fix arbitrary $r \in \hat{I}$ and $c \in \mathcal{C}$. Let $t \in \mathbb{N}$ and $f \in \mathcal{F}_t$ be such that $r = \hat{M}(f_{t-1}, \ldots, f_0, 0, \ldots)$, and note that $r = \lambda M(f_{t-1}, \ldots, f_0, 0, \ldots)$. Then, by the construction of the functions M and $\hat{M}$, we have $\hat{M}(c, f_{t-1}, \ldots, f_0, 0, \ldots) = \hat{\psi}(r, c)$ and $M(c, f_{t-1}, \ldots, f_0, 0, \ldots) = \psi\left(\frac{1}{\lambda} r, c\right)$, which gives $\hat{\psi}(r, c) = \lambda \psi\left(\frac{1}{\lambda} r, c\right)$.

The sufficiency of the conditions can be verified directly. □

Proof of Proposition 5. *Part (i).* Suppose that $\succsim$ satisfies Monotonicity in the Past Memory. Let $m_1, m_2 \in I$ such that $m_1 \geq m_2$. By construction of I, there exist $f, g \in \mathcal{F} \backslash \mathcal{F}_0$ such that $m_1 = M(f_{t-1}, \ldots, f_0, 0, \ldots)$ and $m_2 = M(g_{t'-1}, \ldots, g_0, 0, \ldots)$, where $t = \ell(f)$ and $t' = \ell(g)$. Then, $M(g_{t'-1}, \ldots, g_0, 0, \ldots) \geq ku(z)$ implies $M(f_{t-1}, \ldots, f_0, 0, \ldots) \geq ku(z)$ for all $z \in \mathcal{C}$ and $k > 0$, which, by Lemma 10, is equivalent to $f \; \mathcal{R}^{\succsim} \; g$. By Monotonicity in the Past Memory, we have $(f|c) \; \mathcal{R}^{\succsim} \; (g|c)$ for all $c \in \mathcal{C}$. Using Lemma 10, again, and the Markovian representation, it follows that $\psi(m_2, c) \geq ku(z) \Rightarrow \psi(m_1, c) \geq ku(z)$ for all $z \in \mathcal{C}$ and $k > 0$. Since z and k are arbitrary, we conclude that $\psi(m_1, c) \geq \psi(m_2, c)$.

Part (ii). Suppose that $\succsim$ satisfies Monotonicity in the Current Consumption. Let $m \in I$ and $c_1, c_2 \in \mathcal{C}$ such that $u(c_1) \geq u(c_2)$. Then, there exists $f \in \mathcal{F} \backslash \mathcal{F}_0$ such that $m = M(f_{t-1}, \ldots, f_0, 0, \ldots)$ where $t = \ell(f)$. Since $(c_1) \succsim (c_2)$, we have $(f|c_1) \; \mathcal{R}^{\succsim} \; (f|c_2)$. By Lemma 10 and the Markovian representation, $\psi(m, c_2) \geq ku(z) \Rightarrow \psi(m, c_1) \geq ku(z)$ for all $z \in \mathcal{C}$ and $k > 0$. Since z and k are arbitrary, we conclude that $\psi(m, c_1) \geq \psi(m, c_2)$.

The sufficiency of the conditions for both parts can be verified directly. □

Proof of Proposition 6. Part (a). Assume that $\succsim_2$ exhibits longer effects of positive memory. Let $m \in I_1 \cap I_2 \cap \mathbb{R}_+$. By construction of I_1 and I_2, we can find $f, g \in \mathcal{F} \backslash \mathcal{F}_0$ such that $m = M_1(f_{t-1}, \ldots, f_0, 0, \ldots) = M_2(g_{t'-1}, \ldots, f_0, 0, \ldots) \geq 0$, where $t = \ell(f)$ and $t' = \ell(g)$. Then, $M_1(f_{t-1}, \ldots, f_0, 0, \ldots) \geq ku(z) \Leftrightarrow M_2(g_{t'-1}, \ldots, g_0, 0, \ldots) \geq ku(z)$ for all $z \in \mathcal{C}$ and $k > 0$. That is, $f \ {}^{\succsim_1}\mathcal{I}^{\succsim_2} \ g \ \mathcal{R}^{\succsim_2} \ (0)$. By Definition 5a, we have that $(g|0) \ {}^{\succsim_2}\mathcal{R}^{\succsim_1} \ (f|0) \ \mathcal{R}^{\succsim_1} \ (0)$. Using the Markovian representation, the latter pattern is equivalent to

$$0 \geq u(z) \Rightarrow \psi_1(m, 0) \geq ku(z) \Rightarrow \psi_2(m, 0) \geq ku(z) \quad \text{for all } z \in \mathcal{C} \text{ and } k > 0.$$

Since z and k are arbitrary, we conclude that $\psi_2(m, 0) \geq \psi_1(m, 0) \geq 0$.

Part (b). Assume $\succsim_2$ exhibits longer effects of negative memory. Let $m \in I_1 \cap I_2 \cap \mathbb{R}_-$. Similarly to the above argument, we can find $f, g \in \mathcal{F} \backslash \mathcal{F}_0$ such that $m = M_1(f_{t-1}, \ldots, f_0, 0, \ldots) = M_2(g_{t'-1}, \ldots, f_0, 0, \ldots) \leq 0$. Thus, we have that $(0) \ \mathcal{R}^{\succsim_1} \ f \ {}^{\succsim_1}\mathcal{I}^{\succsim_2} \ g$, and Definition 5b implies that $(0) \ \mathcal{R}^{\succsim_1} \ (f|0) \ {}^{\succsim_1}\mathcal{R}^{\succsim_2} \ (g|0)$. Using the Markovian representation, the latter pattern is equivalent to

$$\psi_2(m, 0) \geq ku(z) \Rightarrow \psi_1(m, 0) \geq ku(z) \Rightarrow 0 \geq u(z) \quad \text{for all } z \in \mathcal{C} \text{ and } k > 0.$$

Thus, it must be that $0 \geq \psi_1(m, 0) \geq \psi_2(m, 0)$.

The converse implication for both parts is routine. □

Proof of Proposition 7. Part (a). Assume that positive memory has stronger effects for $\succsim_2$ in comparison to $\succsim_1$. Let $m \in I_1 \cap I_2 \cap \mathbb{R}_+$. By construction of I_1 and I_2, we can find $f, g \in \mathcal{F} \backslash \mathcal{F}_0$ such that $m = M_1(f_{t-1}, \ldots, f_0, 0, \ldots) = M_2(g_{t'-1}, \ldots, f_0, 0, \ldots) \geq 0$, where $t = \ell(f)$ and $t' = \ell(g)$. Then, $f \ {}^{\succsim_1}\mathcal{I}^{\succsim_2} \ g \ \mathcal{R}^{\succsim_2} \ (0)$.

Pick arbitrary $r \in c_1^+(m)$ and $s \in c_2^+(m)$. If $r > s$ the conclusion immediately holds. Thus, assume $s \geq r$. By definition of $c_i^+(m)$ for $i = 1, 2$, there exist $x, y \in \mathcal{C}$ such that $u(x) = r$, $u(y) = s$, $\psi_1(m, x) \geq m$, and $\psi_2(m, y) \geq m$. This means that $(y) \succsim_2 (x)$, $(f|x) \ \mathcal{R}^{\succsim_1} \ f$, and $(g|y) \ \mathcal{R}^{\succsim_2} \ g$. The conditions of part (a) of Definition 6 are therefore satisfied. It follows that $(f|y) \ \mathcal{R}^{\succsim_1} \ f$ and $(g|x) \ \mathcal{R}^{\succsim_2} \ g$. Using the representation, the latter patterns are equivalent to have $\psi_1(m, y) \geq m$, and $\psi_2(m, x) \geq m$, respectively. We conclude that $s \in c_1^+(m)$ and $r \in c_2^+(m)$, that is, $c_1^+(m)$ dominates $c_2^+(m)$ in the strong set order. The proof that $\mathring{c}_1^+(m)$ dominates $\mathring{c}_2^+(m)$ is analogous and, hence, we omit it.

Date: 2018-07-09 14:54:50 Revision: 3abf17a

Vice versa, assume that $c_1^+(m)$ and $\mathring{c}_1^+(m)$ dominate $c_2^+(m)$ and $\mathring{c}_2^+(m)$ for all $m \in I_1 \cap I_2 \cap \mathbb{R}_+$. Let $f, g \in \mathcal{F} \backslash \mathcal{F}_0$ such that $f \,{}^{\succsim_1}\mathcal{I}^{\succsim_2}\, g \,\mathcal{R}^{\succsim_2}\, (0)$. Using the representation, this means that $M_1(f_{t-1}, \ldots, f_0, 0, \ldots) = M_2(g_{t'-1}, \ldots, f_0, 0, \ldots) = m \geq 0$ for some $m \in I_1 \cap I_2 \cap \mathbb{R}_+$, where $t = \ell(f)$ and $t' = \ell(g)$. Moreover, suppose that $(f|x) \,\mathcal{R}^{\succsim_1}\, f$, $(g|y) \,\mathcal{R}^{\succsim_2}\, g$, and $(y) \succsim_2 (x) \succsim_2 (0)$ for some $x, y \in \mathcal{C}$. Again, by the representation, we have $\psi_1(m, x) \geq m$, $\psi_2(m, y) \geq m$, and clearly $u(y) \geq u(x) \geq 0$. Thus, there exist $r, s \in \mathbb{R}_+$ such that $u(x) = r$, $u(y) = s$, and $r \in c_1^+(m)$, $s \in c_2^+(m)$. Since $c_1^+(m)$ dominates $c_2^+(m)$, it follows that $s \in c_1^+(m)$ and $r \in c_2^+(m)$. By definition of $c_i^+(m)$, we conclude that $\psi_1(m, y) \geq m$ and $\psi_2(m, x) \geq m$. Similarly, it can be shown that part (ii) of Definition 6 follows from the fact that $\mathring{c}_1^+(m)$ dominates $\mathring{c}_2^+(m)$.

Part (b). The proof is similar to the proof of Part (a). For the sake of completeness, we next prove the "Only If" direction. Assume that negative memory has stronger effects for $\succsim_2$ in comparison to $\succsim_1$. Let $m \in I_1 \cap I_2 \cap \mathbb{R}_-$. Then we can find $f, g \in \mathcal{F} \backslash \mathcal{F}_0$ such that $m = M_1(f_{t-1}, \ldots, f_0, 0, \ldots) = M_2(g_{t'-1}, \ldots, f_0, 0, \ldots) \leq 0$. Then, $(0) \,\mathcal{R}^{\succsim_1}\, f \,{}^{\succsim_1}\mathcal{I}^{\succsim_2}\, g$.

Let $r \in c_1^-(m)$ and $s \in c_2^-(m)$. Without loss of generality, assume $r \geq s$. By definition of $c_i^-(m)$ for $i = 1, 2$, there exist $x, y \in \mathcal{C}$ such that $u(x) = r$, $u(y) = s$, $\psi_1(m, x) \leq m$, and $\psi_2(m, y) \leq m$. This means that $(x) \succsim_2 (y)$, $f \,\mathcal{R}^{\succsim_1}\, (f|x)$, and $g \,\mathcal{R}^{\succsim_2}\, (g|y)$. By part (b) of Definition 6, it follows that $f \,\mathcal{R}^{\succsim_1}\, (f|y)$ and $g \,\mathcal{R}^{\succsim_2}\, (g|x)$. Using the representation, we have $\psi_1(m, y) \leq m$, and $\psi_2(m, x) \leq m$. Thus, $s \in c_1^-(m)$ and $r \in c_2^-(m)$, that is, $c_2^-(m)$ dominates $c_1^-(m)$ in the strong set order. Similarly, it can be shown that $\mathring{c}_2^-(m)$ dominates $\mathring{c}_1^-(m)$. □

References

Ariely, D. and Z. Carmon (2000): "Gestalt characteristics of experiences: the defining features of summarized events," *Journal of Behavioral Decision Making*, 13, 191–201.

Ariely, D. and G. Zauberman (2000): "On the Making of an Experience: The Effects of Breaking and Combining Experiences on their Overall Evaluation," *Journal of Behavioral Decision Making*, 13, 219–232.

Attanasio, O. (1999): "Consumption," in *Handbook of Macroeconomics*, ed. by J. B. Taylor and M. Woodford, Elsevier, vol. 1, Part B, chap. 11, 741–812, 1 ed.

Bentham, J. (1789): *An Introduction to the Principle of Morals and Legislations*, Oxford: Blackwell; reprinted (1948).

Diener, E. (1984): "Subjective Well-Being," *Psychological Bulletin*, 95, 542–575.

Diener, E., E. M. Suh, R. E. Lucas, and H. L. Smith (1999): "Subjective well-being: Three decades of progress," *Psychological Bulletin*, 125, 276–302.

Edgeworth, F. Y. (1881): *Mathematical Psychics: An Essay on the Application of Mathematics to the Moral Sciences*, London C.K. Paul; reprinted (New York: M. Kelly, 1967).

Elster, J. and G. Loewenstein (1992): "Utility from memory and anticipation," in *Choice over time*, ed. by G. Loewenstein and J. E. Elster, New York: Russell Sage Foundation, 213–234.

Fernández-Villaverde, J. and D. Krueger (2007): "Consumption over the Life Cycle: Facts from Consumer Expenditure Survey Data," *Review of Economics and Statistics*, 89, 552–565.

Frederick, S. and G. Loewenstein (1999): "Hedonic adaptation," in *Well-being: The foundations of hedonic psychology*, ed. by D. Kahneman, E. Diener, and N. Schwarz, New York: Russell Sage Foundation, 302–329.

Fredrickson, B. and D. Kahneman (1993): "Duration neglect in retrospective evaluations of affective episodes," *Journal of Personality and Social Psychology*, 65, 45–55.

Fredrickson, B. L. (2000): "Extracting meaning from past affective experiences: The importance of peaks, ends, and specific emotions," *Cognition and Emotion*, 14, 577–606.

Gilboa, I., A. Postlewaite, and L. Samuelson (2016): "Memorable Consumption," *Journal of Economic Theory*, 165, 414–455.

Hai, R., D. Krueger, and A. Postlewaite (2015): "On the Welfare Cost of Consumption Fluctuations in the Presence of Memorable Goods," Becker Friedman Institute for Research in Economics Working Paper No. 2553231.

Hall, R. E. (1978): "Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence," *Journal of Political Economy*, 86, 971–987.

Helson, H. (1947): "Adaptation-Level as Frame of Reference for Prediction of Psychophysical Data," *The American Journal of Psychology*, 60, 1–29.

——— (1948): "Adaptation-level as a basis for a quantitative theory of frames of reference," *Psychological Review*, 55, 297–313.

Jappelli, T. and L. Pistaferri (2010): "The Consumption Response to Income Changes," *Annual Review of Economics*, 2, 479–506.

Kahneman, D. (1999): "Objective happiness," in *Well-being: The foundations of hedonic psychology*, ed. by D. Kahneman, E. Diener, and N. E. Schwarz, New York: Russell Sage Foundation, 3–25.

——— (2000a): "Evaluation by Moments: Past and Future," in *Choices, values, and frames*, ed. by D. Kahneman and A. Tversky, New York: Cambridge University Press, 693–708.

——— (2000b): "Experienced Utility and Objective Happiness: A Moment-Based Approach," in *Choices, values, and frames*, ed. by D. Kahneman and A. Tversky, New York: Cambridge University Press, 673–692.

Kahneman, D., E. Diener, and N. Schwarz (1999): *Well-Being: Foundations of Hedonic Psychology*, New York: Russell Sage Foundation.

Kahneman, D., B. L. Fredrickson, C. A. Schreiber, and D. A. Redelmeier (1993): "When More Pain Is Preferred to Less: Adding a Better End," *Psychological Science*, 4, 401–405.

Kahneman, D., P. P. Wakker, and R. Sarin (1997): "Back to Bentham? Explorations of Experienced Utility," *The Quarterly Journal of Economics*, 112, 375–406.

Loewenstein, G. and D. Prelec (1991): "Negative Time Preference," *American Economic Review, Papers and Proceedings*, 81, 347–352.

——— (1993): "Preferences for Sequences of Outcomes," *Psychological Review*, 100, 91–108.

Loewenstein, G. and N. Sicherman (1991): "Do Workers Prefer Increasing Wage Profiles?" *Journal of Labor Economics*, 9, 67–84.

Machina, M. J. (1989): "Dynamic Consistency and Non-expected Utility Models of Choice under Uncertainty," *Journal of Economic Literature*, 27, 1622–68.

Nelson, L. D. and T. Meyvis (2008): "Interrupted Consumption: Disrupting Adaptation to Hedonic Experiences," *Journal of Marketing Research*, 45, 654–664.

Smith, A. (1759): *The Theory of Moral Sentiments*, Printed for A. Millar, in the Strand; and A. Kincaid and J. Bell, in Edinburgh.

Tversky, A. and D. Griffin (1991): "Endowment and contrast in judgments of well-being," in *International series in experimental social psychology, Vol. 21. Subjective well-being: An interdisciplinary perspective*, ed. by F. Strack, M. Argyle, and N. Schwarz, Elmsford, NY: Pergamon Press, 101–118.

Varey, C. and D. Kahneman (1992): "Experiences extended across time: Evaluation of moments and episodes," *Journal of Behavioral Decision Making*, 5, 169–185.

Wakker, P. P. (1989): *Additive Representations of Preferences: A New Foundation of Decision Analysis*, Kluwer Academic Publishers.

Date: 2018-07-09 14:54:50 Revision: 3abf17a