TIME FOR MEMORABLE CONSUMPTION

Stefania Minardi
Andrei Savochkin

NES New Economic School

Working Paper
No 255

NES Working Paper series

October
2019

Time for Memorable Consumption*

Stefania Minardi† and Andrei Savochkin‡

October 8th, 2019

A consumption event is memorable if the memory of it affects well-being at times after the material consumption. We develop an axiomatic model of memorable consumption in a dynamic setting. Preferences are represented by the present value of the sum of utilities derived at each date from the current consumption and from recollecting the past. Our model accommodates well-known phenomena in psychology, such as the peak-end rule, duration neglect, and adaptation trends. We provide foundations for a prominent special case of memory that has the Markovian property. The model is illustrated in application to life-cycle consumption-savings decisions and asset pricing.

1 Introduction

Psychology and behavioral science have widely recognized that one's subjective well-being at any point in time is not determined simply by the consumption at that moment — the recollection of past experiences plays a crucial role. This idea is at the core of a well-known literature initiated by Kahneman and is supported by sizable experimental evidence.[1] Evoking early ideas of Bentham (1789) and Edgeworth (1881), Kahneman describes hedonic experiences as consisting of sequences of moments that give rise to two distinct measurements, so-called 'moment utility' and 'remembered utility.' The former expresses the instant degree of pain or pleasure associated with moments, while the latter refers to the judgement arising from the ex post recollection of the overall experience.

*We thank Fabio Maccheroni, Sujoy Mukerji, Pietro Ortoleva, Leonardo Pejsachowicz, Larry Samuelson, Aleh Tsyvinski, and especially Itzhak Gilboa for many valuable discussions, and participants of RUD 2018, 14th Csef-Igier Symposium, 42nd AMASES Meeting, TUS Conference 2018 (PSE), CHOp Workshop (PSE), Theory Days (Bocconi), SAET 2019, as well as seminars at Queen Mary University, Royal Holloway, and University College London for comments. The authors gratefully acknowledge the financial support of the Investissements d'Avenir program (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047) and CHOp (ANR-17-CE26-0003).

†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 Elster and Loewenstein (1992), Varey and Kahneman (1992), Diener, Suh, Lucas, and Smith (1999), Kahneman, Diener, and Schwarz (1999), Kahneman (2000a, 2000b), and Kahneman and Thaler (2006). The idea that past memories may influence well-being goes back, at least, to Smith (1759).

Date: 2019-10-08 18:48:36 Revision: 89c72c7

When viewed through the lens of modern economics, one problematic aspect of that discourse is that its central concepts of utility are not linked to choice behavior. Experienced 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 a methodology is common practice in psychology, the possibility that memories may affect well-being remains an elusive idea from an economic perspective: understanding the precise nature of the variables being measured is essential for making predictions and studying the implications for economic policy.[2]

This paper focuses on consumption events that we refer to as memorable. For instance, life achievements or exotic vacations can have enduring effects on a person long after the corresponding events have taken 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 one's well-being at a given point in time is affected not only by the current material consumption, but also by the recollection of past memorable events. Our agent recognizes that her current choices may generate valuable memories that will affect her future well-being. Through this channel, memorability also affects choice behavior.[3] Moreover, the effect of past memories may well depend on various features of one's consumption history, such as the intensity or the frequency of the experiences, thereby allowing for a rich dynamic. From an applied perspective, memorable effects may be relevant in different contexts, such as households' consumption patterns and financial planning, as well as employment decisions and the labor market in general.

The first contribution of this paper is to develop axiomatic foundations for a dynamic model of memorable consumption. Our axiomatization separates, behaviorally, the material effect of consumption in the present moment from its memorable effect experienced in subsequent periods. Moreover, it allows us to determine whether an agent perceives a particular consumption experience as memorable or ordinary. If consumption is represented by a bundle of distinct goods, we can identify which goods are memorable. Hence, the trait

[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.

[3] Kahneman and Krueger (2006) suggest that individuals' choices are often influenced by the way past consumption experiences are recollected (giving rise to remembered utilities) and not simply by the consumption profiles themselves. Weber and Johnson (2006) make a similar point by studying how memories may affect preferences.

of being memorable is endogenously derived and subjective, thus varying across individuals. Our second contribution is to develop a theory of Markovian memory as a special case of the general model. With this additional structure, we identify key features of the memory effect, such as longevity and strength, and make interpersonal comparisons along these dimensions. The Markovian specification is useful for reasoning about memory as a dynamic variable and is suitable for using standard dynamic programming methods in applications. We illustrate the broader economic relevance of the model in two classic contexts from macro and finance. First, we introduce memorable consumption into the standard linear-quadratic consumption-savings problem and examine its implications for life-cycle patterns of consumption and savings. Second, we add memorable effects to a Lucas tree economy and study the impact on the risk-free interest rate.

1.1 The model's essential components

We study memorable consumption in a dynamic framework of preferences over consumption streams of different finite length.[4] 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:

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

As in the standard theory of exponential discounting, the parameter $\beta \in (0,1)$ is a discount factor, and the value of $u(f_\tau)$ represents the agent's direct utility of consuming bundle f_τ at time τ. The novel component here is $M(f_{\tau-1}, \dots, f_0, 0, 0, \dots)$, which represents the agent's utility derived from the memory of the consumption history $(f_{\tau-1}, \dots, f_0)$. The expression $u(f_\tau) + M(f_{\tau-1}, \dots, f_0, 0, 0, \dots)$ captures the agent's total subjective well-being that can be attributed to time τ. The value of $M(f_{\tau-1}, \dots, f_0, 0, 0, \dots)$ 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 our representation

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

reduces to the standard exponential discounting. In this case, no consumption experience is memorable for the agent from the point of view of its 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 the analyst can learn from observed choices which goods in the bundles are memorable and which are "ordinary."

To better define the scope of this paper, we note that memorability is not the only potential reason for the past to affect the current utility. One striking example of history dependence is the well-known Mom's Treat, discussed by Machina (1989, p. 1643) and dating back to much earlier literature on interpersonal fairness. Suppose that a mom has a single indivisible treat that she can give either to her daughter or to her son. In principle, she is indifferent between giving the treat to either child. However, if her son got a treat just yesterday, she will strictly prefer to give the treat to her daughter today. Naturally, such a preference does not rely on whether 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,[5] habit formation, and anticipatory feelings. Therefore, we emphasize that our model is not a universal theory of history-dependent utility; rather, we are interested in the phenomenon of memorability, its effect on individual choices, and its relevance for economic analysis. Our focus manifests distinctly in the proposed axiomatization and in the corresponding uniqueness results.

1.2 Special cases and applied relevance

Representation (1) provides a general structure for analyzing different processes by which the memorable effect of past consumption may accrue over time. We look closely at three special cases of that functional form.

Peak-end rule Our first example provides a time-dependent specification of the memory function M that accommodates the so-called *peak-end rule* and *duration neglect* (see

[5] E.g., a person would prefer to accompany his popcorn with different movies on two consecutive nights, even if both movies are not memorable.

Fredrickson and Kahneman, 1993). In these experimentally observed phenomena, the recollection of a prolonged experience is driven by only two salient points — the peak of the intensity and the most recent moment — while neglecting the duration of the experience.

Adaptation trends Our second example proposes a time-dependent specification of the memory function M that captures the agent's adaptation to repeated similar experiences.[6] In particular, an experience becomes memorable and generates utility at later dates depending on its contrast with previous experiences. The proposed specification can be used in a wide range of contexts, from capturing the role of breaks in repeated consumption experiences to thinking about prevention of adaptation in the design of compensation schemes. Furthermore, our specification can be used in job-search models to account for psychological factors such as job satisfaction or to study the impact of unemployment history on future behavior.[7]

Markovian memory We study in more detail a special case of representation (1) in which memory evolves according to a time-invariant Markov law. A consumption stream $f = (f_0, \dots, f_{t-1})$ is evaluated according to

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

where m_τ is computed as $m_\tau = \psi(m_{\tau-1}, f_\tau)$ for $\tau = 0, \dots, t-2$ and $m_{-1} = 0$. The key feature of this specification 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 being used in macroeconomic applications. With the additional Markovian structure of memory, we provide a comparative statics analysis and identify two independent channels through which memories can affect the overall utility in representation (2). One channel is related to the persistence of memory in the agent's mind and, hence, to the rate at which past memories decay. The other one is related to

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

[7] Job satisfaction is often treated as a factor in job mobility, as supported by empirical evidence (Freeman, 1978). See Arulampalam (2001) for evidence on the correlation between past unemployment and future wages.

the sensitivity of the agent's memory to consumption and the ability of consumption to generate new memories.

We apply our Markovian specification to two well-known contexts in macroeconomics and finance. In macroeconomics, we introduce memorable consumption into a standard linear-quadratic consumption-savings problem. In equilibrium, the solution exhibits two key features: 1) the higher sensitivity of consumption to income shocks in comparison with the standard models; and 2) the negative dependence of the optimal consumption on the accumulated stock of memory, which highlights the potential relevance of memorable consumption in explaining some well-known puzzles about life-cycle dynamics of consumption and savings. In finance, we extend the classic Lucas tree asset pricing model to allow for memorable effects of consumption. With this extension, the memorability of consumption modifies the incentives to reallocate consumption across time periods in different states of the economy. We illustrate the impact of memorability on the levels, the volatility, and the dynamic profile of the risk-free interest rate.

The rest of the paper is organized as follows. We next discuss the related literature. Section 2 presents our two applications. Section 3 presents the general model and illustrates its applicability with two examples of time-dependent laws of motion for memory. Section 4 provides a foundation for the special case of time-invariant Markovian dynamics and a comparative statics analysis. Section 5 concludes with a brief summary.

1.3 Related Literature

Our theory entails a particular kind of violation of time separability, which places it in stark contrast with other history-dependent phenomena, most notably habit formation. Two forces drive decision making in our model — the joy of instantaneous consumption and the joy of memories. We assume that the agent's preferences may violate time separability solely in the memory component. Only the latter is influenced by the history; indeed, memories generated by a fine dining experience may depend on the reference point set by past experiences of that sort. Yet our agent's tastes do not change over time, and the direct utility obtained from the *current* consumption is not affected by the history, as happens, for example, in models of habit formation. Our axiomatization captures precisely that:

potential history dependence in memory and history independence in the direct value of consumption. It is worth noting that the differing assumptions about time separability between our model and habit formation lead to opposite predictions. Indeed, habit formation typically strengthens the desire for consumption smoothing. On the contrary, investments in memorable goods frequently generate lumpy patterns of consumption, as Hai, Krueger, and Postlewaite (2015) argue.

While the role of memories in consumption decisions was discussed much earlier (see, e.g., Elster and Loewenstein, 1992), the notion of a memorable good was first formalized by Hai et al. (2015) and Gilboa, Postlewaite, and Samuelson (2016). In these papers, the distinction between ordinary goods and memorable goods is exogenously given. The key feature of those models 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 issue of divisibility, which is key in studying durable goods. Hai et al.'s (2015) theoretical findings are complemented by an empirical analysis that points out stark differences among memorable goods, durable, and non-durable goods. Moreover, their empirical evidence indicates that 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. From a theoretical viewpoint, Gilboa et al. (2016) provide an axiomatic foundation of the static utility structure $u(x, y) + v(y, z)$ at the basis of their applied 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.

We pursue a different line of research. Our axiomatic model is cast in a temporal framework. We treat memorability as a subjective trait and let agents reveal through their choices whether consumption of a particular good is memorable or not.[8] Moreover, we allow for a richer set of laws of motion for the evolution of memory, which support various observations made in psychology. Our additional contribution is the study of the Markovian evolution, a

[8]There is also a lesser point of departure: in our model, memory starts to give the agent additional utility from the moment after the material consumption has taken place; in the other papers, instead, material consumption gives rise to memories and starts to generate utility exactly when the good is consumed.

highly tractable and particularly convenient specification for macroeconomic applications.

2 Memorable consumption in Macro and Finance: Two illustrations

Before presenting our theory in full detail, we illustrate the economic relevance of memorable consumption in two classic contexts — consumption-savings decisions and consumption-based asset pricing.

2.1 Consumption-savings decisions with memorable effects

This section introduces memorable effects of consumption that follow our Markovian specification (2) into a simple linear-quadratic consumption-saving problem. We then solve such a problem and analyze the implications of memorability for the optimal levels of consumption, responses to income shocks, and lifetime paths of consumption and savings.

Suppose that, in periods $t = 0, 1, 2 \ldots$, a consumer receives income y_t that is stochastic and i.i.d. across time. There are no borrowing constraints, hence she can reallocate income between periods by borrowing or saving at the gross interest rate $R > 0$. The time horizon is infinite, and the future is discounted using discount factor $\beta \in (0, 1)$. For simplicity, we assume that there is only one good ($N = 1$). Utility from physical consumption is given by $u(c) = c - \frac{1}{2}c^2$; utility from consuming memories conforms to a convenient special case of (2) where $m = v(\tilde{m})$ with $v(\tilde{m}) = b\tilde{m} - \frac{1}{2}a\tilde{m}^2$, $a, b > 0$, and the memory stock $\tilde{m}$ follows an AR(1)-type law.[9] Thus, the consumer faces the following maximization problem:

$$\underset{\{c_t\}_{t=0}^{\infty}, \{s_t\}_{t=0}^{\infty}, \{\tilde{m}_t\}_{t=0}^{\infty} \text{ adapted}}{\text{maximize}} \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \left(c_t - \frac{1}{2}c_t^2 + b\tilde{m}_{t-1} - \frac{1}{2}a\tilde{m}_{t-1}^2 \right) \right] \tag{3}$$

$$\begin{aligned} \text{s.t. } & c_t + s_t = y_t + Rs_{t-1} && \text{for } t = 0, 1, \ldots, \\ & \tilde{m}_t = \alpha\tilde{m}_{t-1} + (1 - \alpha)c_t && \text{for } t = 0, 1, \ldots, \\ & \tilde{m}_{-1} = 0, \\ & s_{-1} \text{ is given.} \end{aligned}$$

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

[9] We provide all details of this specification in Section 4.3.3 (Example 3). Note, also, that this example extends our specification (2) by considering infinite consumption streams.

Our goal here is to see how memorability of consumption changes the standard Permanent-Income-type solution of the model.[10]

The Lagrangian of the problem is

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

and the First-Order Conditions become

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

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

$$c_t = \left((1-\beta)y_t + \beta\mathbb{E}[y] + \frac{1-\beta}{\beta}s_{t-1}\right)(1+\kappa) - \kappa\tilde{m}_{t-1}, \tag{4}$$

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

$$\kappa = \frac{\sqrt{1 - 2\beta(\alpha^2 - a(1-\alpha)^2) + \beta^2(\alpha^2 + a(1-\alpha)^2)^2} - (1 - \beta(\alpha^2 - a(1-\alpha)^2))}{2\alpha(1-\beta\alpha)}.$$ [11]

Expression (4) 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 explain the well-known empirical evidence on excess sensitivity of consumption to income

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

[11] This consumption rule is supported by μ_t that depends on state variables also linearly, $\mu_t = \left((1-\beta)y_t + \beta\mathbb{E}[y] + \frac{1-\beta}{\beta}s_{t-1} - \tilde{m}_{t-1}\right)\kappa' + \frac{\beta}{1-\beta\alpha}(b - a\tilde{m}_{t-1})$, with a suitably chosen constant κ'.

changes.[12] 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. Such a negative relationship suggests that material consumption (of memorable goods) and consumption of memories behave as substitutes. This distinguishing feature contrasts typical patterns observed in habit formation models where the habit stock and the consumption level move in a complementary way, reinforcing each other.

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 $c_t = \bar{c}_t + \kappa(\bar{c}_t - \tilde{m}_{t-1})$, 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, these dynamics imply 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 reduce 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, these dynamics may represent one key source of support for the empirical evidence according to which individuals consume too little at retirement age compared to the predictions of the canonical model.

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

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

memorable effects lead to greater over-consumption at young ages, as well as stronger reactions 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. As α approaches zero, instead, memory loses its lasting effect, and the decision problem transforms into the standard question of consuming today versus tomorrow. The effect of memorability is the greatest at intermediate levels of α.

2.2 Consumption-based asset pricing with memorable effects

This section introduces memorable effects into a classic Lucas tree framework (Lucas, 1978). Consider an exchange economy with a representative consumer. There is a single productive unit that costlessly produces a stochastic output d_t at time $t = 0, 1, 2, \ldots$. We first analyze the case of i.i.d. dividends and then the more general Markovian case. The output is assumed to be not storable, so the only way to reallocate wealth across periods is given by holding assets. Let z_t denote the share holding at the beginning of period t; owning $z_t = 1$ is an entitlement to all output produced at t. After the dividend is paid, shares are traded at the price p_t that the consumer takes as given. Note that this is a standard setup, traditionally interpreted as a tree that produces a fruit in each period.

Utility from physical consumption in each time period is given by $u(c)$, where u is a strictly increasing, strictly concave and twice differentiable function. The consumer discounts the future with the discount factor $\beta \in (0,1)$. The stock of memory represented by the variable $\tilde{m}$ is assumed to be Markovian and evolves according to the linear ($AR(1)$-like) law: $\tilde{m}_t = \alpha\tilde{m}_{t-1} + (1-\alpha)c_t$. The contribution of the stock of memory $\tilde{m}_{t-1}$ to the consumer's time-t utility is $v(\tilde{m}_{t-1})$, where $v(\tilde{m}) = b\tilde{m} - \frac{1}{2}a\tilde{m}^2$ and $a, b > 0$ are parameters. Similar to the previous application, this specification corresponds to a special case of our Markovian model (2), in which we use $\tilde{m}$ as our state variable instead of $m = v(\tilde{m})$. We assume that the economy operates in a region in which we always have that $v'(\tilde{m}) > 0$. Our consumer faces the following maximization problem:

$$\underset{\{c_t\}_{t=0}^{\infty},\{z_t\}_{t=0}^{\infty},\{\tilde{m}_t\}_{t=0}^{\infty}\text{ adapted}}{\text{maximize}} \mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t\left(u(c_t)+b\tilde{m}_{t-1}-\frac{1}{2}a\tilde{m}_{t-1}^2\right)\right] \tag{5}$$

$$\begin{aligned}
\text{s.t. } & c_t+p_tz_t=(p_t+d_t)z_{t-1} \quad \text{for } t=0,1,\ldots, \\
& \tilde{m}_t=\alpha\tilde{m}_{t-1}+(1-\alpha)c_t \quad \text{for } t=0,1,\ldots, \\
& \tilde{m}_{-1}=0, \\
& z_{-1} \text{ is given.}
\end{aligned}$$

The Lagrangian of the problem is

$$\mathcal{L}=\mathbb{E}_0\Bigg[\sum_{t=0}^{\infty}\beta^t\Bigg(u(c_t)+b\tilde{m}_{t-1}-\frac{1}{2}a\tilde{m}_{t-1}^2 \\ -\lambda_t(c_t+p_tz_t-(p_t+d_t)z_{t-1})-\mu_t(\tilde{m}_t-\alpha\tilde{m}_{t-1}-(1-\alpha)c_t)\Bigg)\Bigg]$$

and the First-Order Conditions are

$$\begin{cases} u'(c_t)-\lambda_t+(1-\alpha)\mu_t=0 \\ -\lambda_tp_t+\beta\mathbb{E}_t[\lambda_{t+1}(p_{t+1}+d_{t+1})]=0 \\ \beta(b-a\tilde{m}_t)-\mu_t+\beta\alpha\mathbb{E}_t[\mu_{t+1}]=0. \end{cases}$$

As in the original Lucas model, the market clearing condition in the goods market means that $c_t=d_t$ for all t.

The pricing kernel (stochastic discount factor) that prices one-period-ahead returns at time t takes the familiar form

$$\mathcal{M}_{t+1}=\frac{\beta\lambda_{t+1}}{\lambda_t}.$$

Let us examine the equilibrium risk-free interest rate in this model. Given the first-order maximality and market clearing conditions, the risk-free interest rate at time t can be computed as

$$R_t^f=\frac{1}{\mathbb{E}_t[\mathcal{M}_{t+1}]}=\frac{1}{\beta}\frac{\lambda_t}{\mathbb{E}_t[\lambda_{t+1}]}=\frac{1}{\beta}\frac{u'(d_t)+(1-\alpha)\mu_t}{\mathbb{E}_t[u'(d_{t+1})]+(1-\alpha)\mathbb{E}_t[\mu_{t+1}]}. \tag{6}$$

Memorable consumption affects R_t^f through the values of the multiplier μ_t and $\mathbb{E}_t[\mu_{t+1}]$.

To formally study its effects, we first consider the special case of i.i.d. dividends. In this case, the risk-free rate can be written as[13]

$$R_t^f = \frac{1}{\beta}\frac{u'(d_t) + a\beta(1-\alpha)\Big(\frac{1}{1-\beta\alpha}\tilde{m}^* - \frac{1}{1-\beta\alpha^2}\tilde{m}_t - \frac{\beta\alpha(1-\alpha)}{(1-\beta\alpha)(1-\beta\alpha^2)}\mathbb{E}[d]\Big)}{\mathbb{E}[u'(d)] + a\beta(1-\alpha)\Big(\frac{1}{1-\beta\alpha}\tilde{m}^* - \frac{\alpha}{1-\beta\alpha^2}\tilde{m}_t - \frac{1-\alpha}{(1-\beta\alpha)(1-\beta\alpha^2)}\mathbb{E}[d]\Big)}, \tag{7}$$

where $\tilde{m}^* = \frac{b}{a}$ is the level of the stock of memory that gives the agent the highest utility (the "bliss point"), $\mathbb{E}[d]$ is the expected dividend ($\mathbb{E}[d] = \mathbb{E}[d_\tau]$ for all τ), and $\mathbb{E}[u'(d)]$ is the expected value of the marginal utility. As can be seen from the above expression, the interest rate depends on the current stock of memory $\tilde{m}_t$, as well as on the parameters a and α, which capture the strength of memory effects and the persistence of memory, respectively.

To interpret (7), fix, first, the stock of memory at its long-term average, $\tilde{m}_t = \mathbb{E}[d]$. Then, the expression for R_t^f becomes

$$R_t^f = \frac{1}{\beta}\frac{u'(d_t) + a\frac{\beta(1-\alpha)}{1-\beta\alpha}(\tilde{m}^* - \mathbb{E}[d])}{\mathbb{E}[u'(d)] + a\frac{\beta(1-\alpha)}{1-\beta\alpha}(\tilde{m}^* - \mathbb{E}[d])}.$$

In the absence of memory, the risk-free interest rate is high (above $\frac{1}{\beta}$) in "bad" states of the world — the ones in which $d_t < (u')^{-1}\big(\mathbb{E}[u'(d)]\big)$. Similarly, the rate is low (below $\frac{1}{\beta}$) in "good" states — the ones in which $d_t > (u')^{-1}\big(\mathbb{E}[u'(d)]\big)$. The presence of memory has a "moderating" effect on the interest rates: in comparison with the no-memory case, the interest rate decreases towards $\frac{1}{\beta}$ in bad states and increases towards $\frac{1}{\beta}$ in good states.

If the current stock of memory $\tilde{m}_t$ in (7) is allowed to vary, then additional effects will emerge. Suppose that the stock of memory is below its long-term average ($\tilde{m}_t < \mathbb{E}[d]$); then, the interest rate shifts upward. In comparison with the no-memory case, the interest rate increases even further in the good state, whereas in the bad state the effect of memory may be ambiguous. (See Figure 1.) Suppose, instead, that the current stock of memory is above its long-term average ($\tilde{m}_t > \mathbb{E}[d]$); then, the interest rate shifts downward. Compared to the no-memory case, in the bad state the interest rate decreases even further, whereas in the good state the effect of memory may be ambiguous.

In the general Markovian case, the above effects remain present.[14] Figure 1 illustrates a

[13] All computations concerning this application are relegated to Appendix C.

[14] See Appendix C, in particular Eq. (28), for an analytical expression for the interest rate in the general case.

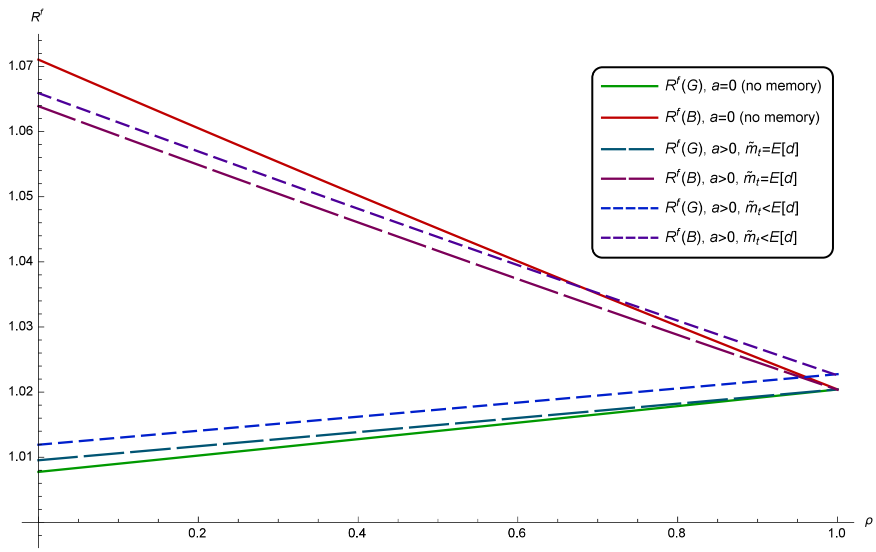

Figure 1.— Risk-free interest rates (gross) for two states. The interest rates are plotted as a function of the parameter ρ that captures the persistence of the state in the Markov process: $\rho = 0$ corresponds to the i.i.d. case, and $\rho = 1$ means that the state does not change with time (in which case, there is no uncertainty in the economy).

typical picture of the interest rates in a two-state process.

We conclude this section by fixing the observable implications of the model in this example.

- The presence of memorable consumption changes the level of the risk-free interest rate. Past memories become a substitute for material consumption, thus making agents behave as if they were more patient. This way, memorability may contribute to explaining the well-known risk-free rate puzzle.[15]
- Memorable consumption has an impact on the volatility of the interest rates and asset returns over time. Indeed, the effects of memorable consumption may have different signs in different states. In our two-state example, in good states the agents prefer to consume more and save less, which decreases the demand for saving and moderates the fall of the interest rate. By contrast, in bad states the agents are less eager to

[15] The puzzle stems from empirically observable risk-free rates that are too low. In the standard framework with time-separable power utility, these rates could be explained by discount factors that are unrealistically high (see, e.g., Campbell (2003) for a discussion).

borrow because memories from the past can alleviate low current consumption, and, hence, the interest rate does not rise that high.

- Finally, the dynamics of interest rates also change. With memorability, the interest rate depends not only on the current state of the economy, but also on the length of the runs of good or bad states in the past.[16] Thus, even if the stochastic process that governs dividends has the simplest Markovian structure, long (or double-dip) downturns in the economy have stronger impacts on the interest rates in comparison with ordinary cyclic downturns. In the former case, the stock of memory suffers a deeper depletion and agents become increasingly eager to consume.

The natural next question is to investigate the effect of memorable consumption on the price of risky assets. This lengthy analysis is outside of the scope of this paper.

3 The general model

This section presents our general model of memorable consumption. Our goal is to propose a minimal deviation from the standard exponential discounting paradigm that allows capturing the essential features of memorable consumption and distinguishing it from other forms of history-dependent consumption.

3.1 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 $c = (c^1, \dots, c^N)$. The set $\mathcal{F}_t = \mathcal{C}^t$ for $t \in \mathbb{N}$ represents the collection of consumption streams of finite length t, with the typical element given by $f = (f_0, f_1, \dots, f_{t-1})$. Also, let $\otimes$ denote the stream of length zero and let $\mathcal{F}_0 = \{\otimes\}$. We denote by $\mathcal{F} = \bigcup_{t=0}^{\infty} \mathcal{F}_t$ the collection of all consumption streams of finite length. 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 $\ell(f) \coloneqq t$ 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

[16] Although the uncertainty in the economy is Markovian, the interest rates now depend on two state variables — the state of Nature and the accumulated stock of memory.

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 weak-* topology: A net $\{P_\alpha\}_\alpha$ in $\mathcal{L}_t$ converges to $P \in \mathcal{L}_t$ iff, for any continuous and bounded function $U : \mathcal{F}_t \to \mathbb{R}$, we have $\int U\, dP_\alpha \to \int U\, dP$.

Throughout the paper, we use the following notation.

Notation. For any $f = (f_0, f_1, \dots, f_k)$ and $h = (h_0, h_1, \dots, h_m)$ in $\mathcal{F}$ and $P \in \mathcal{L}$,

- let $h|f \in \mathcal{F}$ denote the concatenated stream $(h_0, h_1, \dots, h_m, f_0, f_1, \dots, f_k)$.
- let $h|P$ denote the 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.

As usual, we identify a degenerate lottery that gives some stream $f \in \mathcal{F}$ with probability one with the stream itself.

3.2 Axioms

We next introduce the behavioral properties that characterize memorable consumption. They are organized into three groups: properties pertaining to the framework of lotteries over consumption streams; key axioms capturing memorability effects; and few technical assumptions.

Framework assumptions Our first three assumptions are standard for models that deal with both time and uncertainty.

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

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

The above axiom is closely related to the standard formulation of Koopmans (1960): it guarantees consistency and stability of the agent's tastes over time.

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

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

As usual, Impatience ensures that the present is more valuable than the future.[17] Note that Stationarity and Impatience guarantee that our agents make dynamically consistent choices, regardless of whether or not the consumption has any memorable effects.

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

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

Independence is the classic property that delivers the underlying expected utility form.

Overall, the above framework assumptions guarantee that our agent behaves in a perfectly standard way in terms of attitudes toward time and risk.

Axioms pertinent to memory The next two key axioms delineate the behavioral features of memorable consumption that distinguish it from other forms of history-dependent phenomena.

The first axiom asserts that 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 tastes remain unchanged after varying histories. One implication is that it rules out various types of backward-looking reference-dependent evaluations of the current consumption, such as those arising from habit formation, preferences for intrinsic variety, and exposure to experience goods.

Furthermore, since the streams f and g can have different lengths, the above axiom posulates that tastes and risk attitudes remain unchanged with the passage of time. In particular, this averts 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*

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

[17] In comparison to the standard formulation, our Impatience axiom involves comparisons of streams of different length; this is needed in order to isolate the time effects from the memorable effects.

Date: 2019-10-08 18:48:36 Revision: 89c72c7

To interpret this axiom, consider the following tradeoff between consumption in a fixed (the last) period and pleasant memories that are generated by the consumption stream from period zero to the second-to-last. Suppose that changes in the initial part of the consumption stream can be counterbalanced by replacing a consumption lottery p in the last period with a lottery that is a midpoint between p and some q. The axiom postulates 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 counterbalancing effect. Thus, it calibrates the relative effects of memory and consumption in quantitative terms. Note that, together with the other assumptions, this axiom implies the following simple property: for all $f, g \in \mathcal{F}_t$, and $p, q \in \mathcal{L}_1$, we have that $f|p \succsim g|p$ if and only if $f|q \succsim g|q$.[18] This property means 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 (say, positive anticipation of high consumption p versus negative anticipation of low consumption q).[19] The full-fledged Memory-Consumption Tradeoff Consistency rules out additional forms of forward-looking psychological effects.

The axioms of this section clearly hold in the standard model of discounted expected utility. It is noteworthy, however, that they express consistency properties that hold only with respect to the last-period consumption. The last period becomes significant in our theory (different from all preceding periods) because, effectively, the consumption in that period is never memorable, as 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. In turn, 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).

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

[19] Note that the simplest forms of anticipation effect — for instance, the case of a consumption p in a stream $f|p$, where $f \in \mathcal{F}_t$, giving a constant f-independent utility boost in each of the first t periods — is ruled out by the Risk Preference Consistency axiom. Indeed, that axiom asserts that p has the same value after the initial stream f of an arbitrary length t as after the stream $g = \oslash$ of length zero.

Technical requirements We conclude with few technical assumptions.

Axiom A6 (Continuity). *(i) 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. (ii) 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 $(c^*) \succ (0) \succ (c_*)$.*

3.3 Basic representation

The following notation will be useful for stating the results throughout.

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.

The space $\mathcal{C}_0^\infty$ is endowed with the following topology: a net $\{f^{(\alpha)}\}_\alpha$ converges to some f in $\mathcal{C}_0^\infty$ if and only if, for some $T \in \mathbb{N}$ such that $f_t = 0$ for all $t \geq T$, there exists 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. We also say that a function $\Phi : \mathcal{C}_0^\infty \to \mathbb{R}$ is *finite-horizon-bounded* if and only if, 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 $|\Phi(f)| \leq K$.

We are ready to provide a behavioral characterization of preferences that exhibit memorable effects of consumption.

Theorem 1. *Let $\succsim$ be a complete preorder on $\mathcal{L}$. The following statements are equivalent:*

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

(ii) 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 its range contains positive and negative numbers, and a continuous and finite-horizon-bounded function $M : \mathcal{C}_0^\infty \to \mathbb{R}$ with $M(0,0,\dots) = 0$, such that

$$V(P) = \sum_{f \in \operatorname{supp} P} P(f) \sum_{t=0}^{\ell(f)-1} \beta^t \big[u(f_t) + M(f_{t-1}, \dots, f_0, 0, 0, \dots)\big] \tag{8}$$

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

Theorem 1 delivers a preference representation that enriches the standard exponential discounting formula to accommodate the memorable effects of consumption. The usual parameters of the evaluation formula are the scalar β, which represents the discount factor, and the function u, which measures the utility of a bundle of goods at the time of material consumption. Besides its direct value, consumption generates additional utilities in the future, and their flow is measured by a novel object — the function M. Given a stream f, the overall utility that the agent obtains at time t is calculated as the sum $u(f_t) + M(f_{t-1}, \ldots, f_0, 0, 0, \ldots)$, in which the second term specifies the utility derived from the recollection of past memorable experiences. Thus, representation (8) can be interpreted as if the agent engages in two forms of consumption, the material one and the consumption of memories, giving rise to behaviorally distinct utilities. The notation for the arguments of the function M is backward-looking: first goes the most recent past consumption, then the second-to-most-recent, and so on. The sequence of arguments ends with an infinite sequence of zeroes since, at each point in time, the preceding history is assumed to be finite.[20] Note that, if the agent does not perceive any good as memorable, we have $M(\cdot) = 0$, 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 shown next. 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. Importantly, our uniqueness result ensures that our model cannot be reinterpreted in terms of other history-dependent phenomena.

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 $\lambda > 0$.*

3.4 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

[20] 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 $M_2(\cdot, 0) \equiv M_1(\cdot)$, and so on.

of moments that can 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 — namely, the *peak* — and the value experienced at its end — namely, the *end*.[21] One notable implication is that the duration of an experience has no impact on its recollection. For instance, a short, but rather exotic, vacation may generate more intense memories than a longer, but more ordinary, vacation. This pattern, dubbed *duration neglect*, is observed in numerous experimental studies suggesting that prolonging an unpleasant experience by adding some extra moments of diminished discomfort may mitigate the subsequent assessment of the overall experience.[22] 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 (Peak-end rule)**.** *Let the function M from representation (8) be defined as*

$$M(\underbrace{0,\dots,0}_{t}, f_k,\dots,f_{k-l+1},0,f_{k-l-1},\dots,f_0,0,0,\dots) = \beta^t\,\frac{1}{2}\big(u(f_{j^*})+u(f_k)\big)\,K, \tag{9}$$

for streams $f \in \mathcal{F}$ such that $u(f_j) > 0$ for all $j \in \mathbb{Z}_+$ with $k-l < j \le k$, where $t,k \in \mathbb{Z}_+$, $l \in \mathbb{N}$, and $l \le k+1$. Moreover, $j^* \in \mathbb{Z}_+$ satisfies $k-l < j^* \le k$ and $u(f(j^*)) \ge u(f(j))$ for all $j \in \mathbb{Z}_+$ with $k-l < j \le k$. For streams that do not conform to the above pattern, set M equal to zero.[23] *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 (9), the most recent memorable experience lasted l periods. The parameter $K > 0$ 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 (9) 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, such as a generalized "vacation good."

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

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

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

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

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

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.[24] The most relevant economic prediction of the theory is that repeated exposure to the same good experience will gradually attenuate the initial feeling of pleasure; similarly, persistent exposure to the same bad experience 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 the time of *recollection* may well depend on the past history of similar experiences and exhibit adaptation features, as shown next.

Example 2 (Adaptation)**.** *Let the function* M *from representation (8) be defined as*

$$M(f_t, \dots, f_0, 0, 0, \dots) = G(f_t, A(f_{t-1}, \dots, f_0, 0, 0, \dots)), \tag{10}$$

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

The function $A(f_{t-1}, \dots, f_0, 0, 0, \dots)$ *represents the adaptation level acquired from consumption up to time* $t-1$ *and sets the reference point for new memories at time* t*. The formula for* A *can be equivalently written as* $A(f_t, \dots, f_0, 0, 0, \dots) = \alpha f_t + (1-\alpha) A(f_{t-1}, \dots, f_0, 0, 0, \dots)$*, making it clear that the coefficient* α *is the weight attributed to the most recent experience in determining the new adaptation level. The function* G *measures the utility value of the memory from consuming bundle* x *after a history of consumption summarized by*

[24] 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) and Diener, Lucas, and Scollon (2006).

the reference level r. In Tversky and Griffin's (1991) terminology, $A(f_{t-1}, \dots, f_0, 0, 0, \dots)$ 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 memory utility is generated only if the most recent consumption exceeds the reference level. 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 very high standards for fine dining and, at the same time, enjoy pleasant memories from having coffee and pastries in some regular bakery.[25] *This is consistent with our subjective approach to memorability.*

Adaptation-level theory gives rise to a number of well-known patterns. For instance, it suggests that introducing an interval of lower consumption in a lengthy stream of positive consumption may make the agent appreciate it more.[26] Within our setup, we can exemplify this idea by considering the following preference of the agent over mixtures of consumption streams:

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

According to the standard discounted expected utility, the agent should be indifferent between them because, at each date, she consumes c with two thirds probability and zero with one third probability on both the left-hand and 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) than streams in which high consumption is interrupted. Our model can easily accommodate such a preference through a suitable choice of parameters. Furthermore, our model suggests a behaviorally founded reason for preferring intermittent consumption — if the utility from recollecting past experiences follows an adaptation process, then an intermittent profile will generate higher utility flows from memory than will an equivalent constant one. This observation has prescriptive implications, and our model

[25] In an empirical study, Diener et al. (2006) suggest that individuals may have multiple adaptation points. They also report evidence on the heterogeneity of the adaptation process across individuals.

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

may offer new insights into, for instance, designing optimal wage schemes, promotions, or unemployment benefits.[27]

4 Markovian memory

This section studies a prominent special case of our general representation that is particularly suitable for applications in macroeconomics and repeated games, as illustrated in Section 2. Specifically, we provide a behavioral characterization according to which the memory of past consumption follows a Markovian law of motion: the value of memories (i.e., the utility derived from them) at any time t is determined only by the corresponding 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 Markov process.

4.1 The Markovian property

We start 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

$$(f|0) \succsim \frac{1}{k+1} f + \frac{k}{k+1}(f|z). \tag{11}$$

We denote relationship (11) 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 left-hand side of (11), in comparison to the right-hand side, offers the agent a greater chance of enjoying the memory of f in the subsequent time period — on the left-hand side, the probability of enjoying such

[27] See, e.g., Kahneman and Thaler (1991) for a discussion of the implications of adaptation for job satisfaction.

a 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.[28] The additional consumption of z is available to the agent with probability $\frac{k}{k+1}$. Thus, the pattern in (11) 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 greater than the pleasure of consuming z.[29]

The notion of consumption-memory tradeoff allows us to compare consumption streams in terms of their value for generating future memories. As formally stated 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

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

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

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

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

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

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

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

[28] The use of the neutral element 0 on 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.

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

memory for the period t relative to another stream $(g_0, \dots, g_{t-1})$, regardless of the utilities that these streams generate for periods $0, \dots, t-1$. We will use this property shortly to set up the Markovian case. Second, these definitions enable the comparison of the value of memory for streams of different lengths. As a consequence, they can be used to verify that it is, indeed, behaviorally meaningful to attribute the memory utility $M(f_{t-1}, \dots, f_0, 0, 0, \dots)$ to date t in the general representation (8).

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 an additional period of identical consumption, then the value of memories remains the same. This captures the idea of a Markovian process: the memory generated by the extended stream depends only on the memory generated by the original stream and the last-period consumption.

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

For a function $\psi : I \times \mathcal{C} \to \mathbb{R}$, where $I \subseteq \mathbb{R}$ and $0 \in I$, we say that it is *normalized* if $\psi(0,0) = 0$, and it is *recursively bounded* if all sets I_t for $t \in \mathbb{N} \cup \{0\}$ defined recursively as $I_0 = \{0\}$ and $I_t = \psi(I_{t-1}, \mathcal{C})$ for $t \in \mathbb{N}$ are bounded.

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 $u(0) = 0$, an interval I of $\mathbb{R}$ that contains 0, and a normalized, continuous, and recursively bounded function $\psi : I \times \mathcal{C} \to I$ with $\operatorname{range} \psi = I$ 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

$$V(f) = \sum_{t=0}^{\ell(f)-1} \beta^t \big[u(f_t) + m_{t-1}\big], \tag{12}$$

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

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

According to representation (12), 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. Similar 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 $t-1, \ldots, 1, 0$ is encapsulated in the previous stock of memory, m_{t-1}. This is the nature of our Markovian evolution of memory. Note that such a 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 $\big(\hat{\beta}, \hat{u}, \hat{I}, \hat{\psi}\big)$ 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. These monotonicity properties have natural behavioral counterparts, as shown next.

Axiom A9 (Monotonicity in 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 one period of

Date: 2019-10-08 18:48:36 Revision: 89c72c7

extra consumption c.[30]

Axiom A10 (Monotonicity in 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 greater 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)–(A8), and let* (β, u, I, ψ) *be its representation as in Theorem 3. Then:*

(i) $\succsim$ *satisfies Monotonicity in 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 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 formal comparative statics analysis for our model. We propose two ways to compare agents to determine whether one of them is more sensitive to memorable experiences than the other. This inquiry is useful for the purpose of developing parametric examples of the memory evolution function.

As is standard in comparative statics analyses, we start by disentangling the effects of memorable consumption from other unrelated determinants of decision making. To this end, we consider agents who differ in their attitudes toward memorable consumption, but are identical in assessing atemporal risk (including deterministic consumption bundles).[31] The following routine definition formalizes these assumptions.

Definition 3. We say that $\succsim_1$ and $\succsim_2$ on $\mathcal{L}$ have the same risk attitude if

$$(p) \succsim_1 (q) \quad \Leftrightarrow \quad (p) \succsim_2 (q) \qquad \text{for all } p, q \in \mathcal{L}_1. \tag{13}$$

[30]Incidentally, note that Monotonicity in Memory implies the Markovian property.

[31]The other relevant component of individual tastes is the time attitude. As will be clear later, our comparative statics' definitions allow the agents to differ in that dimension.

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 $u_2 = \lambda u_1$.

The next definition introduces a key tool for comparing individuals. It builds upon the notion of a tradeoff between immediate consumption and memories and extends Definition 2 from a single agent to comparisons across different agents.

Definition 4. Let $\succsim_1$ and $\succsim_2$ on $\mathcal{L}$ have the same risk attitude. 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 $f \;{}^{\succsim_1}\mathcal{I}^{\succsim_2}\; g$.

4.3.1 Comparative persistence of memory

We now study the comparative attitudes towards memorable consumption. Our first definition provides a way to infer from choice behavior whether memory has a more persistent effect for one agent compared to another.

Definition 5. Let $\succsim_1$ and $\succsim_2$ on $\mathcal{L}$ satisfy Monotonicity in Memory (A9) and have the same risk attitude.

(a) *Positive* memorable consumption has *longer effects* for $\succsim_2$ in comparison to $\succsim_1$ if, for any $f, g \in \mathcal{F}$ such that $f \;{}^{\succsim_1}\mathcal{I}^{\succsim_2}\; g$, we have that:

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

(b) *Negative* memorable consumption has *longer effects* for $\succsim_2$ in comparison to $\succsim_1$ if, for any $f, g \in \mathcal{F}$ such that $f \ {}^{\succsim_1}\mathcal{I}^{\succsim_2}\ g$, we have that:

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

Parts (a) and (b) provide symmetric definitions that distinguish between positive and negative experiences. Consider Part (a) first. We consider two consumption streams, f and g, that produce the same stock of memories for Agent 1 and 2, respectively.[32] The antecedent of the implication formula defines the sign of the stock of memory: if 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. Thus, Part (a) restricts attention to comparisons of agents with equivalent baggage of positive experiences. Now, extend streams f and g by adding one last period of zero consumption. That is, consider the pair of streams $f|0$ and $g|0$ such that its only difference from pair f and g is that the extended streams allow both agents to enjoy their memories for one extra period. (Note that Monotonicity in Memory guarantees that the extended streams $f|0$ and $g|0$ keep generating positive memories like the initial streams f and g.[33]) Then, we say that a positive memory has longer effects for Agent 2 than for 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 in Agent 1's.

Part (b) presents a symmetric notion for unpleasant experiences. Here, 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.

Proposition 6 provides a characterization of the behavioral concept developed above in terms of the Markovian representation and shows the way in which comparative memory persistence is determined by the Markovian function ψ.

[32]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 memories of different value for different agents. By considering different streams and letting individuals express their preferences, we take this aspect fully into account.

[33]Monotonicity in Memory can be dropped in this definition. Then, to preserve the same sign for the memories of the extended streams, we would need to strengthen the consequent by adding $(f|0)\ \mathcal{R}^{\succsim_1}\ (0)$ in Part (a) and $(0)\ \mathcal{R}^{\succsim_1}\ (f|0)$ in Part (b).

Proposition 6. *Suppose that* $\succsim_1$ *and* $\succsim_2$ *are complete preorders on* $\mathcal{L}$ *that satisfy Axioms (A1)–(A8) and Monotonicity in Memory (A9), and have the same risk attitude. Let* $(\beta_1, u, I_1, \psi_1)$ *and* $(\beta_2, u, I_2, \psi_2)$ *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* $m \leq 0$.

4.3.2 Comparative strength of memory

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 sufficient to maintain the stock of memory at a particular level. The ranking is going to reflect how sensitive the agent's memory is to her consumption. At the same time, 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 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 $\succsim_2$ on $\mathcal{L}$ have the same risk attitude.

(a) *Positive* memorable consumption has *stronger effects* for $\succsim_2$ in comparison to $\succsim_1$ if, for any $f_1, f_2 \in \mathcal{F}$ such that $f_1 \,{}^{\succsim_1}\mathcal{I}^{\succsim_2}\, f_2$ and $f_1 \,\mathcal{R}^{\succsim_1}\, (0)$, we have that

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

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

(b) *Negative* memorable consumption has *stronger effects* for $\succsim_2$ in comparison to $\succsim_1$ if, for any $f_1, f_2 \in \mathcal{F}$ such that $f_1 \,{}^{\succsim_1}\mathcal{I}^{\succsim_2}\, f_2$ and $(0) \,\mathcal{R}^{\succsim_1}\, f_1$, we have that

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

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

Analogously to the previous notion of comparative persistence, Part (a) considers the case in which a stock of memory represents positive experiences for Agent 2 (and, hence, for Agent 1, as well). Suppose that two streams, f_1 and f_2, produce the same stock of memory for Agents 1 and 2, respectively, as revealed by the comparative relation ${}^{\succsim_1}\mathcal{I}^{\succsim_2}$. Suppose, also, 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 streams f_1 and f_2 for them. Importantly, suppose that the bundle y is weakly preferred to x (by both agents, clearly). Then, the axiom prescribes that x and y are interchangeable, in that they both increase the stock of memory for both individuals. This requirement rules out two situations that conflict with the intuitive idea of sensitivity of memory to consumption and the strength of memory effects. The first situation is that a bundle x increases the stock of memory for less sensitive Agent 1 (weaker memory effects), while it fails to do so for more sensitive Agent 2, for whom memory gets increased only by a better bundle y. The second situation is that a bundle y increases the stock of memory for more sensitive Agent 2 (stronger memory effects), while it fails to do so for less sensitive Agent 1, for whom memory gets increased by a worse bundle x.

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 again to capture 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)–(A8) and have the same risk attitude. Let $(\beta_1, u, I_1, \psi_1)$ and $(\beta_2, u, I_2, \psi_2)$ 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 $c_i^-(m) = \{r \leq 0 : u(c) = r \text{ for some } c \in \mathcal{C} \text{ and } \psi_i(m,c) \leq m\}$ 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, 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 $m \in I_1 \cap I_2 \cap \mathbb{R}_-$.

4.3.3 Examples

As an illustration, 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:*

$$\psi_i(m,c) = \alpha_i m + (1-\alpha_i)K_i u(c), \tag{14}$$

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 a longer effect if and only if $\alpha_2 \geq \alpha_1$ (and if and only if Agent 2's negative memory has a 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 $\alpha_i m + (1-\alpha_i)K_i u(c) \geq m$ for $i = 1,2$ — that is, $u(c) \geq \frac{1}{K_i} m$ — for all non-negative values of m, as well as for the set of solutions of similar strict inequalities. Such a set for Agent 2 is dominated if and only if $\frac{1}{K_2} \leq \frac{1}{K_1}$. Hence, Agent 2's positive memory has a stronger effect if and only if $K_2 \geq K_1$. 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.

Note that the specifications of the evolution of memory in the applications of Section 2 can be thought of as belonging to the same class as (14). Indeed, consider the following generalization of (14). Let $v : \mathbb{R} \to \mathbb{R}$ be a strictly increasing function with $v(0) = 0$, $w : \mathcal{C} \to \mathbb{R}$ be another function, and let the law for the evolution of memory be defined as

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

where $\alpha \in (0,1)$.[34] We can then change variables by defining $\tilde{m}_t = v^{-1}(m_t)$ and rewrite the Markovian representation (12) as

$$V(f) = \sum_{t=0}^{\ell(f)-1} \beta^t \big[u(f_t) + v(\tilde{m}_{t-1})\big], \tag{15}$$
$$\text{where} \quad \tilde{m}_\tau = \alpha \tilde{m}_{\tau-1} + (1-\alpha)w(f_\tau) \quad \text{for } \tau = 0,\dots,\ell(f)-2,$$
$$\tilde{m}_{-1} = 0.$$

Here, the stock of memory $\tilde{m}_t$ is measured in different "units" in comparison with (12), which leads to a specification convenient for applications: $AR(1)$-type law for memory and (potentially) non-linear function in the agent's objective. In Section 2, the utility specification is, indeed, given by (15) with $v(\tilde{m}) = b\tilde{m} - \frac{1}{2}a\tilde{m}^2$, $a,b > 0$, and $w(c) = c$.

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:*

$$\psi_i(m,c) = \alpha_i \max\{m, K_i u(c)\} + (1-\alpha_i)K_i u(c), \tag{16}$$

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 a 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 agents are compared: Agent 2's positive experience has stronger memory effects if and only if $K_2 \geq K_1$.

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

5 Conclusion

We have presented a model of memorable consumption in temporal setup: the agent makes choice over dated consumption streams, as well as lotteries over them. The model allows to elicit from observables whether consumption is memorable or not for the agent. Due to the uniqueness properties, we can distinguish between the utility derived from material consumption and the utility derived from the "consumption of memories;" moreover, those utilities can be rightfully attributed to time periods. Our axioms consist of testable properties that distinguish memorable effects from other types of history-dependent behavior that violate time separability.

We devote special attention to the case of Markovian evolution of memory: we propose a novel Markovian axiom and characterize an analytically tractable representation that should be particularly useful for macro and other dynamic applications. The Markovian setting brings to the fore two independent channels through which agents can be compared — according to the strength of memory effects and to the longevity of their memory. The model is illustrated with two examples in the classical settings of the consumption-savings problem and asset pricing in a Lucas tree economy.

Appendix

A Proof of the basic representation

Lemma 8. *Suppose that $\succsim$ is a preference relation on $\mathcal{L}$ 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 $p, q \in \mathcal{L}_1$,

$$f|p \succsim g|(\tfrac{n-1}{n}p + \tfrac{1}{n}q) \quad \Leftrightarrow \quad f|(\tfrac{1}{n}p + \tfrac{n-1}{n}q) \succsim g|q. \tag{17}$$

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

Date: 2019-10-08 18:48:36 Revision: 89c72c7

$\frac{n-1}{n}p + \frac{1}{n}q' = \frac{n}{n+1}p + \frac{1}{n+1}q$ 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 (17) 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 $a, b, c \in X$ and $u, v, w \in Y$ are such that $(b, u) \asymp (a, v)$[35] and $(c, u) \asymp (b, v) \asymp (a, w)$. Our goal is to show that $(c, v) \asymp (b, w)$. Let $r = \frac{1}{2}u + \frac{1}{2}w$. We claim that

$$(a, r) \asymp (a, v). \tag{18}$$

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

If $(b,u) \succ (a,r)$ then, on the one hand, $(a,v) \succ (a,r)$. On the other hand, $(b,u) \succ (a,r)$ implies by (iv) that $(b,r) \succ (a,w)$, which means that $(b,r) \succ (b,v)$. We obtained a contradiction with (iii). The situation $(a,r) \succ (b,u)$ similarly leads to a contradiction. We conclude that $(a,r) \asymp (b,u) \asymp (a,v)$. Then, observe that $(c,u) \asymp (b,v) \asymp (b,r)$, where the first part holds by assumption and the second follows from (18) and (iii). Then, $(c,r) \asymp (b,w)$ by (iv). As follows from (18) and (iii), we also have $(c,r) \asymp (c,v)$. The desired relationship follows by transitivity.

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

$$(x,y) \succcurlyeq (x',y') \quad \Leftrightarrow \quad U_x(x) + U_y(y) \geq U_x(x') + U_y(y'). \tag{19}$$

It remains to show that U_y must be affine. Indeed, (19) and property (iv) imply that, for any $y, y' \in Y$,

$$U_y(y) - U_y(\tfrac{1}{2}y + \tfrac{1}{2}y') \geq U_x(x') - U_x(x) \Leftrightarrow U_y(\tfrac{1}{2}y + \tfrac{1}{2}y') - U_y(y') \geq U_x(x') - U_x(x) \quad \text{for all } x, x' \in X. \tag{20}$$

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 (20) gives that, for any $y, y' \in Y$ such that $|U_y(y) - U_y(y')| \leq \varepsilon$,

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

Applying it repeatedly, this equation can be extended to all $y, y' \in Y$. Moreover, it can be rewritten 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). Throughout the proof, we will write z_t for $t \in \mathbb{N}$ to denote an element of $\mathcal{F}_t$ such that $z_t = (0, 0, \ldots, 0)$.

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 the range of u admits both positive and negative values.

Step 2. Independence, Part (ii) of Continuity, and Nondegeneracy ensure that the conditions of the mixture space theorem (Herstein and Milnor, 1953) 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 $p, q \in \mathcal{L}_1$.

If, for some $t \in \mathbb{N}$, we have $f|p \sim z_t|p$ for all $f \in \mathcal{F}_t$ and $p \in \mathcal{L}_1$, then, as follows from Stationarity, we can let $W_t(f) = 0$ for all $f \in \mathcal{F}_t$.

Otherwise, we obtain the claim by Lemma 9: Assumption (iv) holds due to Memory-Consumption Tradeoff Consistency, (iii) by Risk Preference Consistency, (ii) by Lemma 8, and (i) with respect to the second coordinate holds by Nondegeneracy and Stationarity. 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, by the uniqueness of ordinal representations, 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, $\zeta_t(\cdot)$ must be positive affine functions for all $t \in \mathbb{N}$: $\zeta_t(x) = A_t + B_t x$ for some $A_t \in \mathbb{R}$ and $B_t \in \mathbb{R}_{++}$. If we let $\tilde{W}_t(f) \coloneqq A_t + B_t W_t(f)$ for all $t \in \mathbb{N}$ and $f \in \mathcal{F}_t$, we obtain:

$$V(f|p) = \tilde{W}_t(f) + B_t\mathbb{E}_p[u] \qquad \text{for all } t \in \mathbb{N},\ f \in \mathcal{F}_t, \text{ and } p \in \mathcal{L}_1. \tag{21}$$

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 (21), we also have $\tilde{W}_t(z_t) = 0$.

Let $\beta \coloneqq B_1$ and note that Impatience implies that $\beta < 1$. For any $c \in \mathcal{C}$, let $p_c \in \mathcal{L}_1$ be defined as $p_c \coloneqq \beta\delta_c + (1-\beta)\delta_0$, and observe that $V(0,c) = \beta u(c) = \mathbb{E}_{p_c}[u] = V(p_c)$, where the first equality holds by (21) and the last equality by construction of V in Step 2. For any $t \in \mathbb{N}$, Stationarity gives that $z_t|0|c \sim z_t|p_c$ and, hence, by (21), $B_{t+1}u(c) = B_t\mathbb{E}_{p_c}[u] = B_t\beta u(c)$. Since c was arbitrarily chosen, we have that $B_t = \beta^t$ for all $t \in \mathbb{N}$ and

$$V(f|p) = \tilde{W}_t(f) + \beta^t\mathbb{E}_p[u] \qquad \text{for all } t \in \mathbb{N},\ f \in \mathcal{F}_t, \text{ and } p \in \mathcal{L}_1. \tag{22}$$

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

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

$$M_t(f_{t-1},\dots,f_0) \coloneqq \beta^{-t}(\tilde{W}_t(f_0,\dots,f_{t-1}) - V(f_0,\dots,f_{t-1})). \tag{23}$$

Using this definition in (22), we obtain that, for all $t \in \mathbb{N}$ and $f \in \mathcal{F}_{t+1}$,

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

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

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

Step 8. We claim that, for all $t \in \mathbb{N}$ and $P \in \mathcal{L}$, $V(z_t|P) = \beta^t V(P)$. First, recall that it was shown in Step 6 that $V(z_t|0) = 0$ for all $t \in \mathbb{N}$. Now, 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 $\mathcal{L}$. Hence, by uniqueness of affine representations, there exists $b > 0$ such that $V((0)|P) = bV(P)$ for all $P \in \mathcal{L}$. As follows from (22), it must be that $b = \beta$. The claim now follows by induction. Note that, by (22), we also have $\tilde{W}_{\ell(f)+t}(z_t|f) = \beta^t \tilde{W}_{\ell(f)}(f)$ 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 $t \in \mathbb{N} \cup \{0\}$,

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

Step 10. Observe that M is continuous in the specified topology: for any $t \in \mathbb{N}$, $M(f)$ coincides with $M_t(f)$ for all f that are zero starting from time t. Functions M_t for $t \in \mathbb{N}$ are defined through V and W_t that are continuous (Steps 2 and 4).

Furthermore, M is finite-horizon-bounded: for that, it is sufficient to show that M_t for all $t \in \mathbb{N}$ are bounded. Indeed, note that we can rewrite (23) in Step 7 as $M_t(f_{t-1}, \dots, f_0) = \beta^{-t}(V(f_0, \dots, f_{t-1}, 0) - V(f_0, \dots, f_{t-1}))$ for all $f \in \mathcal{F}_t$. Then, the function V restricted to $\mathcal{F}_t$ is bounded by the von Neumann-Morgenstern expected utility theorem because $\succsim$ restricted to $\mathcal{L}_t$ is a continuous preference relation. Representation (8) is now proven.

If part. Suppose that $\succsim$ admits a utility representation via a function V, as specified in (8). 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 (8) 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 (8), $u(c) > 0$. 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 (8) 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

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

Then, for any $t \in \mathbb{N}$, $f, g \in \mathcal{F}_t$, and $p, q \in \mathcal{L}_1$, we have by (8) 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 $f \mapsto \sum_{\tau=0}^{t-1} \beta^\tau \big[u(f_\tau) + M(f_{\tau-1}, \ldots, f_0, 0, 0, \ldots)\big]$ is continuous and bounded by the corresponding properties of u and M. Hence, when restricted to $\mathcal{L}_t$, V defined by (8) is continuous in the weak-* topology, which establishes the first part of the axiom. The second part follows immediately from the expected utility structure of V.

Nondegeneracy. The property follows directly from the fact that the range of u contains 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 $k > 0$,*

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

Proof. Let $t := \ell(f)$. Using representation (8), we obtain

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

□

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

Step 1. Let $V : \mathcal{L} \to \mathbb{R}$ be a utility representation of $\gtrsim$ as in (8), with β, u, and M as specified in Theorem 1. Let $I := \{M(f_{t-1}, \dots, f_0, 0, 0, \dots) \mid f \in \mathcal{F}_t, t \in \mathbb{N}\}$ and note that I contains 0 bacause $M(0, 0, \dots) = 0$. Define $\psi : I \times \mathcal{C} \to I$ as follows: For $r \in \mathbb{R}$ and $c \in \mathcal{C}$, $\psi(r, c) := M(c, f_{t-1}, \dots, f_0, 0, 0, \dots)$, where $f \in \mathcal{F}_t$ for some $t \in \mathbb{N}$ is an arbitrary act such that $M(f_{t-1}, \dots, f_0, 0, 0, \dots) = r$.

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 $M(f_{t-1}, \dots, f_0, 0, 0, \dots) = M(f'_{t'-1}, \dots, f'_0, 0, 0, \dots)$, 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

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

and, therefore, by Lemma 10,

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

By Axiom A8, we have

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

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

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

Step 4. We claim that ψ is recursively bounded. Indeed, the sets I_t, $t \in \mathbb{N} \cup \{0\}$ defined recursively as $I_0 = \{0\}$ and $I_t = \psi(I_{t-1}, \mathcal{C})$ for $t \in \mathbb{N}$ coincide with $\{M(f_{t-1}, \dots, f_0, 0, \dots) \mid f \in \mathcal{F}_t\}$ and, hence, are bounded for all $t \in \mathbb{N} \cup \{0\}$ because M is finite-horizon bounded.

Step 5. 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 $I_t = \{M(f_{t-1}, \dots, f_0, 0, 0, \dots) \mid f \in \mathcal{F}_t\}$. 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 $M\left(f_{t-1}^{(1)}, \dots, f_0^{(1)}, 0, 0, \dots\right) = r_1$ and $M\left(f_{t-1}, \dots, f_0, 0, 0, \dots\right) = r$. 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 $M\left(f_{t-1}^{(n)}, \dots, f_0^{(n)}, 0, 0, \dots\right) = r_n$, 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 $r = \lim_{n\to\infty} r_n = \lim_{n\to\infty} M\left(f_{t-1}^{(n)}, \dots, f_0^{(n)}, 0, 0, \dots\right) = M\left(f_{t-1}^{(\infty)}, \dots, f_0^{(\infty)}, 0, 0, \dots\right)$ by the continuity of M. By the result of Step 2, we have $\psi(r_n, c_n) = M\left(c_n, f_{t-1}^{(n)}, \dots, f_0^{(n)}, 0, 0, \dots\right)$ for all $n \in \mathbb{N}$ and $\psi(r, c) = M\left(c, f_{t-1}^{(\infty)}, \dots, f_0^{(\infty)}, 0, 0, \dots\right)$; by the continuity of M, we have $\lim_{n\to\infty} M\left(c_n, f_{t-1}^{(n)}, \dots, f_0^{(n)}, 0, 0, \dots\right) = M\left(c, f_{t-1}^{(\infty)}, \dots, f_0^{(\infty)}, 0, 0, \dots\right)$; and we obtain that $\lim_{n\to\infty} \psi(r_n, c_n) = \psi(r, c)$, 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 (12).

Let $M : \mathcal{C}_0^{\infty} \to \mathbb{R}$ be defined as follows. For $h = (h_0, h_1, \dots, h_{t-1}, 0, 0, \dots) \in \mathcal{C}_0^{\infty}$, where $t \in \mathbb{N} \cup \{0\}$, 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 $\tau \geq t$.

Clearly, M satisfies the normalization condition $M(0, 0, \dots) = 0$.

Next, observe that it is also finite-horizon-bounded: for any $T \in \mathbb{N} \cup \{0\}$, the range of M when restricted to the set $\{f \in \mathcal{C}_0^{\infty} : f_t = 0 \text{ for all } t \geq T\}$ can be computed recursively as $I_0 = \{0\}$ and $I_t = \psi(I_{t-1}, \mathcal{C})$ for $t \in \mathbb{N}$ and is bounded since ψ is recursively bounded.

Finally, 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 exists 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, $M\left(h^{(\alpha)}\right) = \psi\left(\psi\left(\ldots\psi\left(0, h_{T-1}^{(\alpha)}\right), \ldots, h_1^{(\alpha)}\right), h_0^{(\alpha)}\right) \to M(h) = \psi(\psi(\ldots\psi\left(0, h_{T-1}\right), \ldots, h_1), h_0)$ because of the continuity of ψ.

Thus, we can apply the converse direction of Theorem 1 to conclude that Axioms (A1)–(A7) hold. 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

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

By Lemma 10, this gives

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

Due to the arbitrariness of z and k and the fact that range u takes both positive and negative values, it must be that $M(x, f_{\ell(f)-1}, \ldots, f_0, 0, 0, \ldots) = M(y, g_{\ell(g)-1}, \ldots, g_0, 0, 0, \ldots)$. Fix an arbitrary $c \in \mathcal{C}$. Then, $M(c, x, f_{\ell(f)-1}, \ldots, f_0, 0, 0, \ldots) = \psi(M(x, f_{\ell(f)-1}, \ldots, f_0, 0, 0, \ldots), c) = \psi(M(y, g_{\ell(g)-1}, \ldots, g_0, 0, 0, \ldots), c) = M(c, y, g_{\ell(g)-1}, \ldots, g_0, 0, 0, \ldots)$. By Lemma 10, again,

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

□

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 $r = \hat{M}(f_{t-1}, \ldots, f_0, 0, 0, \ldots)$, and note that $r = \lambda M(f_{t-1}, \ldots, f_0, 0, 0, \ldots)$. Then, by the construction of the functions M and $\hat{M}$, we have $\hat{M}(c, f_{t-1}, \ldots, f_0, 0, 0, \ldots) = \hat{\psi}(r, c)$ and $M(c, f_{t-1}, \ldots, f_0, 0, 0, \ldots) = \psi\left(\frac{1}{\lambda}r, c\right)$, 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 Memory. Let $m_1, m_2 \in I$ such that $m_1 \geq m_2$. By construction of I, there exist $f, g \in \mathcal{F}$ such that $m_1 = M(f_{t-1}, \ldots, f_0, 0, 0, \ldots)$ and $m_2 = M(g_{t'-1}, \ldots, g_0, 0, 0, \ldots)$, where $t = \ell(f)$ and $t' = \ell(g)$. Then, $M(g_{t'-1}, \ldots, g_0, 0, 0, \ldots) \geq ku(z)$ implies $M(f_{t-1}, \ldots, f_0, 0, 0, \ldots) \geq ku(z)$ for all $z \in \mathcal{C}$ and $k > 0$, which, by Lemma 10, is equivalent to $f \ \mathcal{R}^{\succsim} \ g$. By Monotonicity in 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 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 $f \in \mathcal{F}$ such that $m = M(f_{t-1}, \ldots, f_0, 0, 0, \ldots)$ 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. Fix an arbitrary $m \in I_1 \cap I_2 \cap \mathbb{R}_+$. By construction of I_1 and I_2, we can find $f, g \in \mathcal{F}$ such that $m = M_1(f_{t-1}, \ldots, f_0, 0, 0, \ldots) = M_2(g_{t'-1}, \ldots, g_0, 0, 0, \ldots) \geq 0$, where $t = \ell(f)$ and $t' = \ell(g)$. Then, $M_1(f_{t-1}, \ldots, f_0, 0, 0, \ldots) \geq ku(z) \Leftrightarrow M_2(g_{t'-1}, \ldots, g_0, 0, 0, \ldots) \geq ku(z)$ for all $z \in \mathcal{C}$ and $k > 0$. Hence, by definition, $f \ {}^{\succsim_1}\mathcal{I}^{\succsim_2} \ g$. By the definition of longer effects of positive memory, this implies that $(g|0) \ {}^{\succsim_2}\mathcal{R}^{\succsim_1} \ (f|0)$. As follows from Lemma 10 and the Markovian formula for the M_1 and M_2 functions, this relationship is equivalent to

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

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

Part (b). Assume that $\succsim_2$ 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 $f, g \in \mathcal{F}$ such that $m = M_1(f_{t-1}, \ldots, f_0, 0, 0, \ldots) = M_2(g_{t'-1}, \ldots, g_0, 0, 0, \ldots) \leq 0$. Thus, we have that $(0) \ \mathcal{R}^{\succsim_1} \ f \ {}^{\succsim_1}\mathcal{I}^{\succsim_2} \ g$, and Definition 5b implies that $(f|0) \ {}^{\succsim_1}\mathcal{R}^{\succsim_2} \ (g|0)$. Monotoncity in Memory ensures that $(0) \ \mathcal{R}^{\succsim_1} \ (f|0)$. Using the Markovian representation, we obtain

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

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 $f, g \in \mathcal{F}$ such that $m = M_1(f_{t-1},\dots,f_0,0,0,\dots) = M_2(g_{t'-1},\dots,g_0,0,0,\dots) \geq 0$, 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 that $s \geq r$. 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 having $\psi_1(m,y) \geq m$, 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.

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 $f, g \in \mathcal{F}$ such that $f \ {}^{\succsim_1}\mathcal{I}^{\succsim_2}\ g\ \mathcal{R}^{\succsim_2}\ (0)$. Using the representation, this means that $M_1(f_{t-1},\dots,f_0,0,0,\dots) = M_2(g_{t'-1},\dots,f_0,0,0,\dots) = m \geq 0$ 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 $f, g \in$ such that $m = M_1(f_{t-1},\dots,f_0,0,0,\dots) = M_2(g_{t'-1},\dots,g_0,0,0,\dots) \leq 0$. 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

$\psi_1(m,y) \le m$, and $\psi_2(m,x) \le m$. 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 $\mathring{c}_1^-(m)$. $\square$

C For Online Publication: Calculations for Section 2.2

From the First-Order Conditions, we have $\mu_t = \beta(b - a\tilde{m}_t) + \beta\alpha\mathbb{E}_t[\mu_{t+1}]$. Then:

$$\begin{aligned}\mathbb{E}_t[\mu_{t+1}] &= \mathbb{E}_t[\beta(b - a\tilde{m}_{t+1})] + \beta\alpha\mathbb{E}_{t+1}[\mu_{t+2}] \qquad (24)\\ &= \sum_{\tau=1}^{T}\mathbb{E}_t\left[\beta(\beta\alpha)^{\tau-1}(b - a\tilde{m}_{t+\tau})\right] + (\beta\alpha)^T\mathbb{E}_t[\mu_{t+T+1}] = \sum_{\tau=1}^{\infty}\mathbb{E}_t\left[\beta(\beta\alpha)^{\tau-1}(b - a\tilde{m}_{t+\tau})\right]\\ &= \frac{\beta b}{1-\beta\alpha} - \beta a\sum_{\tau=0}^{\infty}(\beta\alpha)^{\tau}\mathbb{E}_t[\tilde{m}_{t+\tau+1}]. \qquad (25)\end{aligned}$$

Iteratively using the constraint $\tilde{m}_t = \alpha\tilde{m}_{t-1} + (1-\alpha)c_t$ yields $\tilde{m}_{t+\tau+1} = \alpha^{\tau+1}\tilde{m}_t + (1-\alpha)\sum_{j=1}^{\tau+1}\alpha^{\tau+1-j}c_{t+j}$. Recall, also, that in equilibrium $c_t = d_t$ for all t. Thus, we can rewrite the last component of (24) as

$$\begin{aligned}\sum_{\tau=0}^{\infty}(\beta\alpha)^{\tau}\mathbb{E}_t[\tilde{m}_{t+\tau+1}] &= \sum_{\tau=0}^{\infty}(\beta\alpha)^{\tau}\alpha^{\tau+1}\tilde{m}_t + \sum_{\tau=0}^{\infty}(\beta\alpha)^{\tau}(1-\alpha)\sum_{j=1}^{\tau+1}\alpha^{\tau+1-j}\mathbb{E}_t[d_{t+j}] \qquad (26)\\ &= \frac{\alpha}{1-\beta\alpha^2}\tilde{m}_t + (1-\alpha)\sum_{j=1}^{\infty}\sum_{\tau=j-1}^{\infty}(\beta\alpha)^{\tau}\alpha^{\tau+1-j}\mathbb{E}_t[d_{t+j}]\\ &= \frac{\alpha}{1-\beta\alpha^2}\tilde{m}_t + (1-\alpha)\sum_{j=1}^{\infty}\alpha^{1-j}\frac{(\beta\alpha^2)^{j-1}}{1-\beta\alpha^2}\mathbb{E}_t[d_{t+j}]\\ &= \frac{\alpha}{1-\beta\alpha^2}\tilde{m}_t + \frac{1-\alpha}{1-\beta\alpha^2}\sum_{j=1}^{\infty}(\beta\alpha)^{j-1}\mathbb{E}_t[d_{t+j}]. \qquad (27)\end{aligned}$$

I.i.d. case Assume that the dividend is i.i.d with mean $\bar{d}$; then, $\mathbb{E}_t[d_{t+j}] = \mathbb{E}[d_{t+j}] = \bar{d}$. Thus, we can rewrite (27) as

$$\sum_{\tau=0}^{\infty}(\beta\alpha)^{\tau}\mathbb{E}_t[\tilde{m}_{t+\tau+1}] = \frac{\alpha}{1-\beta\alpha^2}\tilde{m}_t + \frac{1-\alpha}{1-\beta\alpha^2}\frac{\bar{d}}{1-\beta\alpha}.$$

Date: 2019-10-08 18:48:36 Revision: 89c72c7

The latter can be plugged into the expressions for $\mathbb{E}_t[\mu_{t+1}]$ and μ_t:

$$\mathbb{E}_t[\mu_{t+1}] = \frac{\beta b}{1-\beta\alpha} - \frac{\beta\alpha a}{1-\beta\alpha^2}\tilde{m}_t - \frac{\beta(1-\alpha)a}{1-\beta\alpha^2}\frac{\bar{d}}{1-\beta\alpha}$$
$$= \frac{\beta}{1-\beta\alpha}\left[b - a\left(\frac{\alpha(1-\beta\alpha)}{1-\beta\alpha^2}\tilde{m}_t + \frac{1-\alpha}{1-\beta\alpha^2}\bar{d}\right)\right];$$
$$\mu_t = \beta(b - a\tilde{m}_t) + \frac{\beta^2\alpha}{1-\beta\alpha}\left[b - a\left(\frac{\alpha(1-\beta\alpha)}{1-\beta\alpha^2}\tilde{m}_t + \frac{1-\alpha}{1-\beta\alpha^2}\bar{d}\right)\right]$$
$$= \frac{\beta}{1-\beta\alpha}\left[b - a\left(\frac{1-\beta\alpha}{1-\beta\alpha^2}\tilde{m}_t + \frac{\beta\alpha(1-\alpha)}{1-\beta\alpha^2}\bar{d}\right)\right].$$

Plugging the above expressions into (6), we obtain the following formulation of R_t^f:

$$R_t^f = \frac{1}{\beta}\frac{u'(d_t) + a\beta(1-\alpha)\left(\frac{1}{1-\beta\alpha}m^* - \frac{1}{1-\beta\alpha^2}m_t - \frac{\beta\alpha(1-\alpha)}{(1-\beta\alpha)(1-\beta\alpha^2)}\mathbb{E}[d_{t+1}]\right)}{\mathbb{E}[u'(d_{t+1})] + a\beta(1-\alpha)\left(\frac{1}{1-\beta\alpha}m^* - \frac{\alpha}{1-\beta\alpha^2}m_t - \frac{1-\alpha}{(1-\beta\alpha)(1-\beta\alpha^2)}\mathbb{E}[d_{t+1}]\right)},$$

where $m^* = \frac{b}{a}$.

Markovian case Assume that the dividend is represented by a vector Y with components corresponding to each state of a Markov process with transition matrix P. We have

$$\sum_{j=1}^{\infty}(\beta\alpha)^{j-1}\mathbb{E}_t[d_{t+j}] = \sum_{j=1}^{\infty}(\beta\alpha)^{j-1}P^jY = (I-\beta\alpha P)^{-1}PY;$$
$$\mathbb{E}_t[\mu_{t+1}] = \frac{\beta}{1-\beta\alpha}b - \frac{\beta\alpha}{1-\beta\alpha^2}a\tilde{m}_t - \frac{\beta(1-\alpha)}{1-\beta\alpha^2}a(I-\beta\alpha P)^{-1}PY;$$
$$= \beta a\left[\frac{1}{1-\beta\alpha}\tilde{m}^* - \frac{\alpha}{1-\beta\alpha^2}\tilde{m}_t - \frac{1-\alpha}{1-\beta\alpha^2}(I-\beta\alpha P)^{-1}PY\right];$$
$$\mu_t = \frac{\beta}{1-\beta\alpha}b - \frac{\beta}{1-\beta\alpha^2}a\tilde{m}_t - \frac{\beta^2\alpha(1-\alpha)}{1-\beta\alpha^2}a(I-\beta\alpha P)^{-1}PY.$$

Hence, the expression for the risk-free interest rate in the Markovian case is:

$$R_t^f = \frac{1}{\beta}\frac{u'(d_t) + a\beta(1-\alpha)\left(\frac{1}{1-\beta\alpha}m^* - \frac{1}{1-\beta\alpha^2}m_t - \frac{\beta\alpha(1-\alpha)}{(1-\beta\alpha^2)}(I-\beta\alpha P)^{-1}PY\right)}{\mathbb{E}[u'(d_{t+1})] + a\beta(1-\alpha)\left(\frac{1}{1-\beta\alpha}m^* - \frac{\alpha}{1-\beta\alpha^2}m_t - \frac{1-\alpha}{(1-\beta\alpha^2)}(I-\beta\alpha P)^{-1}PY\right)}. \quad (28)$$

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.

Arulampalam, W. (2001): “Is Unemployment Really Scarring? Effects of Unemployment Experiences on Wages,” *The Economic Journal*, 111, F585–F606.

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).

Campbell, J. (2003): “Consumption-based asset pricing,” in *Handbook of the Economics of Finance*, ed. by G. Constantinides, M. Harris, and R. M. Stulz, Elsevier, vol. 1, Part 2, chap. 13, 803–887, 1 ed.

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

Diener, E., R. E. Lucas, and C. N. Scollon (2006): “Beyond the hedonic treadmill: Revising the adaptation theory of well-being,” *American Psychologist*, 61, 305–314.

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.

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.

Freeman, R. B. (1978): “Job Satisfaction as an Economic Variable,” *American Economic Review, Papers and Proceedings of the Ninetieth Annual Meeting of the American Economic Association*, 68, 135–141.

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.

Herstein, I. N. and J. Milnor (1953): “An Axiomatic Approach to Measurable Utility,” *Econometrica*, 21, 291–297.

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. and A. B. Krueger (2006): "Developments in the Measurement of Subjective Well-Being," *Journal Economic Perspectives*, 20, 3–24.

Kahneman, D. and R. Thaler (1991): "Economic Analysis and the Psychology of Utility: Applications to Compensation Policy," *American Economic Review, Papers and Proceedings of the Hundred and Third Annual Meeting of the American Economic Association*, 81, 341–346.

——— (2006): "Anomalies: Utility Maximization and Experienced Utility," *Journal of Economic Perspectives*, 20, 221–234.

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

Koopmans, T. C. (1960): "Stationary Ordinal Utility and Impatience," *Econometrica*, 28, 287–309.

Lucas, R. (1978): "Asset Prices in an Exchange Economy," *Econometrica*, 46, 1429–45.

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.

Weber, E. U. and E. J. Johnson (2006): "Constructing Preferences from Memory," in *The Construction of Preference*, ed. by S. Lichtenstein and P. Slovic, New York, NY: Cambridge University Press, 302–329.