# Who Saves More, the Naive or the Sophisticated Agent?*

Max Groneck† Alexander Ludwig‡ Alexander Zimper§

September 30, 2019

**Abstract**

We construct and solve a dynamically inconsistent Choquet expected utility life-cycle model for naive and sophisticated agents, respectively. Pollak (1968) shows that the realized saving behavior of naive and sophisticated agents becomes identical for a logarithmic period-utility function. As a generalization of Pollak's analysis, we compare the saving behavior of both types of agents for the family of isoelastic utility functions. We show that a sophisticated agent saves more in every period than her naive counterpart if and only if her period-utility function is more concave than the logarithmic function. This relationship holds for arbitrary survival beliefs and time-discount factors. Quantitatively, the difference in saving behavior across the two types can be large.

*JEL Classification: D91, D83, E21.*

*Keywords: Survival beliefs; Ambiguity; Choquet expected utility; Dynamic inconsistency*

*We thank Menusch Khadjavi, Ulrich Schmidt, Peter Wakker and Mirko Wiederholt for valuable comments and suggestions. Alex Ludwig thanks the Universitat Autonoma de Barcelona for its hospitality during a sabbatical year and gratefully acknowledges financial support by Research Center SAFE, funded by the State of Hessen initiative for research LOEWE.

†Department of Economics, Econometrics and Finance; University of Groningen; Netspar; P.O. Box 800 9700AV Groningen, The Netherlands: E-mail m.groneck@rug.nl

‡SAFE, Goethe University Frankfurt; CMR; MEA; Netspar; House of Finance; Theodor-W.-Adorno-Platz 3; 60629 Frankfurt am Main, Germany; E-mail: ludwig@safe.uni-frankfurt.de

§Department of Economics; University of Pretoria; Private Bag X20; Hatfield 0028; South Africa; E-mail: alexander.zimper@up.ac.za

# 1 Introduction

How households consume and save over the life-cycle and how survival beliefs as well as time preferences affect these decisions are classical economic questions. The workhorse model to address this problem of inter-temporal allocation is the life-cycle model of Modigliani and Brumberg (1954) and Ando and Modigliani (1963). Standard life-cycle models consider an expected utility maximizing agent whose age-dependent survival beliefs are conditional additive probabilities.

This paper deviates from this standard by constructing and solving a life-cycle model under the assumption that the agent is a Choquet expected utility (CEU) decision maker. CEU preferences have first been axiomatized by Schmeidler (1989) for the Anscombe-Aumann (1963) framework and by Gilboa (1987) for the Savage (1954) framework. CEU preferences can accommodate Ellsberg (1961) paradoxes through non-additive beliefs whose properties are interpreted as expressions of the decision maker's ambiguity attitudes. On the gain domain CEU theory coincides with cumulative prospect theory (Tversky and Kahneman 1992; Wakker and Tversky 1993) which has proved in numerous experimental studies to be a powerful descriptive model of decision making under uncertainty (cf. Wakker 2010).

Because the Choquet operator integrates the Bernoulli utilities of a given act over a fixed order of states that corresponds to the ranking of the act's consequences, the CEU functional exhibits kinks whenever this order is about to change through changes in the considered act. As a consequence the optimal interior choices of a CEU decision maker might not always be characterized by first-order conditions. The fact that the CEU functional is not everywhere differentiable gives rise to interesting and relevant deviations from expected utility theory (see Billot et al. 2019 for a recent survey of this literature). On the downside, however, this technical difficulty can make it unappealing for economists to use CEU theory in their models.[1] Fortunately, this technical difficulty is not a problem for the application of CEU theory to life-cycle models as long as the decision maker's uncertainty only concerns her future survival chances. Under the natural assumption that the CEU agent prefers to live as long as possible the fixed order of states becomes the same for every consumption plan be-

[1] This has motivated the development of *smooth* ambiguity models (Klibanoff et al. 2005; 2009) as alternatives to CEU theory. Although these models have—arguably—no longer the descriptive power of CEU theory in experimental situations, they allow for first-order characterizations of interior optima in portfolio choice and asset-pricing. See, e.g., Collard et al. (2018) for an application of smooth ambiguity models to consumption-based asset pricing.

cause different states simply correspond to different periods in which the agent dies. Combined with an additive separable Bernoulli utility function defined over consumption streams the CEU life-cycle model takes on the form of a standard life-cycle model with general effective discount factors so that these discount factors are given as a combination of age-dependent non-additive survival beliefs and time-discount factors.

Although the optimal consumption plan of a CEU decision maker is thus characterized through intertemporal Euler equations, a major challenge arises from the fact that the CEU life-cycle model is, in general, not dynamically consistent. That is, the optimal consumption plan will, in general, not coincide with realized consumption over the life-cycle. More specifically, we show that the CEU life-cycle model is dynamically consistent for an exponential time-discounting agent if and only if this agent is also a Bayesian decision maker who uses Gilboa and Schmeidler's (1993) *optimistic* update rule to update her survival beliefs when she grows older.[2] For any other concepts of time-discounting or of age-dependent survival beliefs the future selfs of the agent will have strict incentives to deviate from the ex ante optimal consumption plan. To deal with this dynamic inconsistency, we solve the CEU life-cycle model for the realized consumption paths of a sophisticated and a naive agent, respectively. Despite the fact that the sophisticated agent and her naive counterpart share the same CEU preferences over consumption streams, the realized consumption paths of both agent types result from very different optimization problems. The sophisticated agent—who is aware of the deviating incentives of her future selfs—chooses her per-period consumption as if she plays a strategic game against her future selfs. In contrast, the naive agent chooses her per-period consumption under the misperception that her future selfs will stick to the consumption plan that is optimal from her ex ante perspective.

How does the saving behavior of the sophisticated agent compare with that of her naive counterpart? In models with a *presence bias*—induced by *hyperbolic* and *quasi-hyperbolic time-discounting* (cf. Phelps and Pollak 1968; Laibson 1997; O'Donoghue and Rabin 1999)—one would intuitively think that sophisticated agents save more than their naive counterparts. More generally, one would probably expect that the answer to the posed question depends on several model parameters like survival beliefs and time discount functions. However, this intuition is flawed. A remarkable result

[2] Instead of the unique update rule for additive beliefs, there exists a multitude of Bayesian update rules for CEU decision makers. Gilboa and Schmeidler (1993) describe one family of such update rules of which the *optimistic* and the *pessimistic* update rules are important special cases (cf. our discussion in Section 2).

by Pollak (1968) already shows that the realized consumption paths of naive and sophisticated agents coincide for a logarithmic period utility function irrespective of survival beliefs and time-discounting. That is, although both agent types solve very different life-cycle decision problems, the logarithmic utility function ensures that the actual outcome of realized consumption in every period coincides for the sophisticated and naive agent irrespective of the specification of their shared effective discount factors.

To understand this finding, it is useful to relate to the interpretation of the *generalized Euler equation* for the restricted class of quasi-hyperbolic discount functions in Harris and Laibson (2001). On the one hand, the quasi-hyperbolic sophisticated agent when making decisions at any age $h$ has a higher effective discount factor than the naive agent between periods $h$ and $h+1$. This induces the sophisticated agent to save more. On the other hand, the effective discount factor varies inversely with the marginal propensity to consume of her own future self. If the MPC at age $h+1$ increases the sophisticated agent reacts to this overconsumption by constraining her future self. She will accordingly save less. Thus, to relate back to Pollak (1968)'s finding, these two motives exactly offset each other for logarithmic utility.

How do these saving motives interplay for general isolastic utility functions under arbitrary discounting as in Pollak (1968)? To address this question, our CEU life-cycle generalizes Pollak (1968)'s analysis to isoelastic utility functions over per-period consumption so that the marginal utility function is $u'(c) = c^{-\theta}$ for concavity parameter $\theta > 0$. In a static decision situation, $\theta$ would correspond to the constant relative risk aversion coefficient so that greater values of $\theta$ express a greater aversion against risk. In the context of intertemporal consumption choices, $\theta$ measures the *resistance* to intertemporal substitution (Kimball and Weil 2009)[3] so that a greater value of $\theta$ describes a decision maker who is less willing to change the consumption allocation over time in response to a change of the inter-temporal price of consumption. Overall, a greater value of the concavity parameter $\theta$ means that the agent is more eager to smooth out consumption over different states of the world as well as over different time periods.

As our main finding we establish that, somewhat surprisingly, the value of the concavity parameter $\theta$ completely determines whether the naive or the sophisticated agent saves more in any given period (cf. Theorem 3):

[3]Its inverse is more commonly referred to as the *elasticity of inter-temporal substitution.*

- *The naive agent exhibits in every period a greater marginal propensity to consume than her sophisticated counterpart if and only $\theta > 1$.*

- *Conversely, the sophisticated agent exhibits in every period a greater marginal propensity to consume than her naive counterpart if and only $\theta < 1$.*

In other words, irrespective of the specification of survival beliefs and the time-discount factor, the sophisticated agent saves more in every period than her naive counterpart if and only if the isoelastic period-utility function is more concave than the logarithmic function. To be precise, we only prove this result analytically for effective discount factors that satisfy a specific linearity condition, which we also refer to as "constant patience" in all future periods after the next period. Constant patience means that the future marginal valuation of savings between the current self and the next period future self coincide for all periods after the next period. This linearity condition is, for example, satisfied in the deterministic version of our model for quasi-hyperbolic time-discounters (Laibson 1997). It is also satisfied for exponential time-discounters who are Bayesian decision makers that apply the Gilboa and Schmeidler (1993) *pessimistic* update rule to pessimistic *neo-additive* beliefs (Chateauneuf et al. 2007). Pessimistic neo-additive capacities are the simplest convex non-additive probability measures we can think of. By being convex these beliefs express pure ambiguity aversion and the CEU decision maker who applies the pessimistic update rule is formally equivalent to a maxmin expected utility decision maker of Gilboa and Schmeidler (1989) whose set of multiple priors is given as the core of the neo-additive belief.[4]

In order to compare the saving behavior of both agent types for effective discount factors that do not satisfy the linearity condition, we conduct two variants of Monte Carlo experiments. In our first variant, we generate sequences of arbitrary discount factors that are independently and identically distributed over age. As a second variant, we superimpose further structure by specifying discount functions as the combination of age-dependent non-additive survival beliefs and time-discount factors. We draw arbitrary prior survival beliefs and form posteriors by applying the *pessimistic* (Gilboa and Schmeidler 1993) and the *generalized Bayesian* (Eichberger et al. 2007) update rules. All our numerical experiments confirm the above sharp relationship between the concavity parameter and the saving behavior of naive versus

[4] The maxmin expected utility theory by Gilboa and Schmeidler (1989) has motivated the work on robustness in macroeconomic models by Hansen and Sargent (2011, 2016).

sophisticated agents. We therefore conclude that—irrespective of survival beliefs and time-discount-factors—the sophisticated agent saves more than her naive counterpart in our CEU life-cycle model if and only if the isoelastic period utility function is more concave than the logarithmic function. Our second variant also shows that differences in saving behavior across the two types of decision makers can be quantitatively large for empirically relevant parameterizations of $\theta$.

The remainder of our analysis proceeds as follows. Section 2 introduces our notion of survival beliefs. Section 3 constructs the CEU life-cycle model, which is then solved in Section 4 for the sophisticated and naive agent, respectively. Section 5 compares the saving behavior of both types of agents. Section 6 concludes. Our main analytical result, Theorem 3, is formally proved in Appendix A. Supplementary Appendix B contains details on our computational experiments.

# 2 Survival Beliefs

## 2.1 General Notion

Consider an agent of age $h \geq 0$ and fix some maximal $T \geq 2$ with the interpretation that the agent cannot survive beyond age $T$. For all ages $h$ we construct the probability spaces $\left(\Omega, \mathcal{F}, \nu^h\right)$ for a non-additive probability measure $\nu^h$ which describes the $h$-old agent's survival beliefs.[5] The state space is given as $\Omega = \{\omega_0, \ldots, \omega_T\}$ and the $\sigma$-algebra $\mathcal{F}$ is given as the powerset of $\Omega$. We interpret $D_t = \{\omega_t\}$ as the event in $\mathcal{F}$ that the agent dies at the end of age $t$. Observe that

$$D_t \cup \cdots \cup D_T \tag{1}$$

stands for the event in $\mathcal{F}$ that the agent of age $h < t$ survives until (at least) the beginning of age $t$. As a notational convention, we write for the $h$-old agent's belief to survive until (at least) the beginning of age $t > h$

$$\nu_{h,t} = \nu^h \left(D_t \cup \cdots \cup D_T\right).$$

[5]To be precise: when we speak of *non-additive probability measures* we actually mean *not necessarily additive probability measures* as we also allow for the possibility of additive probability measures.

**Definition 1** (Survival Beliefs)**.** *We consider a system of age-dependent non-additive probability measures* $\{\nu^h\}_{h=1,\ldots,T}$ *such that, for every* $h$*,* $\nu^h : \mathcal{F} \to [0,1]$ *satisfies the following conditions:*

*(i) Normalization:* $\nu_{h,t} = 0$ *for all* $t < h$*, and* $\nu_{h,h} = 1$*;*

*(ii) Monotonicity:* $\nu_{h,t} \geq \nu_{h,k}$ *for* $k > t \geq h$*;*

*(iii) Non-degeneracy:* $\nu_{h,t} > 0$ *for all* $t > h$*.*

The above notion of survival beliefs is as general as it gets. It encompasses, for example, survival beliefs derived from a fixed probability weighting function applied to conditional additive probabilities—as in the rank-dependent utility life-cycle models in Bleichrodt and Eeckhoudt (2006) and in Drouhin (2015)—as well as the calibrated survival beliefs in Ludwig and Zimper (2013) and in Groneck et al. (2016) that are derived from a Choquet Bayesian learning model. Let us emphasize at this point that all formal results in this paper will be derived for this general notion of survival beliefs unless additional restrictions are explicitly stated.

## 2.2 Bayesian Decision Maker

Out of additional consistency considerations it is common practice in the literature to consider a Bayesian decision maker, which imposes a stronger condition on survival beliefs than the above properties (i)-(iii). A Bayesian decision maker is characterized through a Bayesian update rule which generates from a prior belief conditional beliefs (i.e., posteriors) in the light of new information. The information filtration for our life-cycle model is simple: in each period the decision maker 'learns' whether she has survived or not whereby we are only interested in the updated beliefs of the surviving decision maker. That is, the relevant information in any given period $t$ is simply the survival event

$$D_h \cup \cdots \cup D_T$$

according to which the decision maker is $h$-old. Moreover, the only events that our decision maker cares about are her future survival events (1) for $t > h$. If the prior is some additive probability measure, denoted $\mu$, then there exists a unique Bayesian update rule (i.e., a unique definition of conditional beliefs) according to which

$$\begin{aligned}\mu_{h,t} &= \frac{\mu\left((D_t \cup \cdots \cup D_T) \cap (D_h \cup \cdots \cup D_T)\right)}{\mu\left(D_h \cup \cdots \cup D_T\right)} \\ &= \frac{\mu\left(D_t \cup \cdots \cup D_T\right)}{\mu\left(D_h \cup \cdots \cup D_T\right)},\end{aligned}$$

implying

$$\mu_{t,t+1} = \frac{\mu_{h,t+1}}{\mu_{h,t}}. \tag{2}$$

In contrast to this unique definition of Bayesian updating for additive probabilities, there exists a multitude of alternative Bayesian update rules for CEU decision makers with non-additive beliefs. Let us briefly explain why.[6] In order to explain Ellsberg paradoxes, CEU preferences must allow for the possibility that Savage's sure-thing principle is violated. Denote by $f_A h$ a Savage act (i.e., a mapping from the state space into the set of consequences) that gives the consequences of act $f$ in the event $A$ and the consequences of act $h$ in the complement event $\Omega \backslash A$. The sure thing principle states that, for all acts $f, g, h, h'$,

$$f_A h \succeq g_A h \Leftrightarrow f_A h' \succeq g_A h'.$$

A Bayesian decision maker is characterized by some rule that determines how her ex ante preferences $\succeq$ are updated to her ex post preferences $\succeq_A$ which are conditional on having observed the event $A$. If the sure-principle holds, we can unambiguously define, for any $h$,

$$f_A h \succeq g_A h \Rightarrow f \succeq_A g$$

as unique update rule. In violating the sure-thing principle, however, a CEU decision maker might have the ex ante preferences

$$f_A h \succeq g_A h \text{ and } g_A h' \succeq f_A h'.$$

Under the $h$-rule preferences would be updated to

$$f_A h \succeq g_A h \Rightarrow f \succeq_A g$$

[6] For an excellent introductory reading on this topic see Ghirardato (2002).

whereas we would obtain under the $h'$-rule the opposite ex post preferences

$$g_A h' \succeq f_A h' \Rightarrow g \succeq_A f.$$

In other words, a Bayesian update rule for a CEU decision maker has to specify some act $h^*$, possibly depending on $f, g$ and $A$,[7] whose consequences the decision maker associates with the outcomes in the now impossible complement event $\Omega \backslash A$. The fact that we can choose any such $h^*$ explains the multitude of possible update rules for CEU decision makers.

Gilboa and Schmeidler (1993) consider a family of Bayesian update rules for CEU decision makers such that $h^*$ is the same for all $f, g$ and $A$. Two extreme rules out of this family come with straightforward psychological interpretations. According to the *optimistic* update rule, the act $h^*$ would always result in the worst possible consequences so that the decision maker feels relieved to observe event $A$ instead of the complement event $\Omega \backslash A$. Conversely, the *pessimistic* update rule associates $h^*$ with the best possible consequences so that the decision maker will be disappointed upon observing $A$.

Denote by

$$\nu^{Bayes} \left( D_t \cup \cdots \cup D_T \mid D_h \cup \cdots \cup D_T \right)$$

the conditional belief of the $h$-old decision maker to survive until (at least) the beginning of age $t$ such that this belief is formed in accordance with some update rule '*Bayes*'. Fix some update rule *Bayes*. We speak of a Bayesian decision maker if her system of age-dependent beliefs $\left\{ \nu^h \right\}_{h=1,\ldots,T}$ satisfies, for all $t > h$,

$$\nu_{h,t} = \nu^{Bayes} \left( D_t \cup \cdots \cup D_T \mid D_h \cup \cdots \cup D_T \right).$$

Next we apply Gilboa and Schmeidler's (1993) formal characterizations of the optimistic and pessimistic update rule, respectively, to survival beliefs.

**Optimistic versus pessimistic Bayesian updating of survival beliefs.**

**(i)** *Optimistic update rule:*

$$\nu_{h,t} = \frac{\nu \left( D_t \cup \cdots \cup D_T \right)}{\nu \left( D_h \cup \cdots \cup D_T \right)}.$$

[7]See, e.g., the update rule in Sarin and Wakker (1998) which guarantees dynamic consistency of CEU preferences for all possible information filtrations.

**(ii)** *Pessimistic update rule:*

$$\nu_{h,t} = \frac{\nu\left(D_0 \cup \cdots \cup D_{h-1} \cup D_t \cup \cdots \cup D_T\right) - \nu\left(D_0 \cup \cdots \cup D_{h-1}\right)}{1 - \nu\left(D_0 \cup \cdots \cup D_{h-1}\right)}. \tag{3}$$

Analogously to the additive case (2), the optimistic update rule implies

$$\nu_{t,t+1} = \frac{\nu_{h,t+1}}{\nu_{h,t}} \tag{4}$$

for all $t \geq h$. This will drive the fact that the CEU life-cycle model is dynamically consistent for an exponential time-discounter if and only if the agent is a Bayesian decision maker who applies the optimistic update rule (cf. Proposition 2).[8] We are going to use the pessimistic update rule for illustrative reasons when we discuss the 'linearity' condition for effective discount factors that enters our Theorem 3 as assumption. More precisely, we will consider a Bayesian decision maker who applies the pessimistic update rule to a *pessimistic neo-additive belief* (cf. Chateauneuf et al. 2007). Such a belief is given as

$$\nu\left(D_t \cup \cdots \cup D_T\right) = (1-\delta)\, \mu\left(D_t \cup \cdots \cup D_T\right) \tag{5}$$

where $\mu$ is an additive probability measure and the parameter $\delta \in (0,1)$ measures ambiguity. Applying the pessimistic update rule to (5) results in the new pessimistic neo-additive belief[9]

$$\nu_{h,t} = (1-\delta_h)\, \mu_{h,t} \tag{6}$$

with increased ambiguity parameter

$$\delta_h = \frac{\delta}{\delta + (1-\delta)\, \mu_{0,h}}.$$

[8] The optimistic updating rule violates—as the other update rules—the law of iterated Choquet expectations to the effect that this rule cannot guarantee dynamic inconsistency for all perceivable information filtration processes (cf. Zimper 2011; Lapied and Toquebeuf 2013). That the optimistic update rule gives rise to dynamic consistency of the CEU life-cycle model is therefore a consequence of its specific information filtration process.

[9] As it is, the use of another prominent update rule—namely, the *Generalized Bayesian* update rule (Eichberger et al. 2007)—results in the same posterior for a pessimistic neo-additive belief. For formal applications of update rules to neo-additive beliefs see Zimper (2011).

## 2.3 Relation to the Multiple Priors Approach

Gilboa and Schmeidler's (1989) maxmin expected utility theory is the classical multiple priors model that describes an ambiguity averse decision maker. Instead of resolving her uncertainty through a unique additive probability measure the maxmin expected utility decision maker holds a whole set of additive priors whereby she is extremely pessimistic in that she evaluates an act by its expected utility with respect to the worst belief in this set.

CEU theory and maxmin expected utility theory coincide for a *convex* $\nu$ (i.e., satisfying $\nu(A) + \nu(B) \leq \nu(A \cup B) + \nu(A \cap B)$ for all $A, B \in \mathcal{F}$) when the set of multiple priors is taken to be the *core* of $\nu$ defined as the following subset of all additive probability measures $\mu$ on $(\Omega, \mathcal{F})$

$$core\,(v) = \{\mu \mid \nu(A) \leq \mu(A) \text{ for all } A \in \mathcal{F}\}.$$

Suppose that this maxmin expected utility decision maker applies the Full Bayesian update rule according to which her set of multiple posteriors consists of all updated priors. Then this decision maker can be equivalently be characterized as a CEU decision maker who uses the *Generalized Bayesian* update rule (Eichberger et al. 2007; 2012). That is, for a convex $\nu$ we have that

$$\begin{aligned} core\,(\nu_{h,t}) &= \left\{\mu_{h,t} \mid \nu_{h,t}(A) \leq \mu_{h,t}(A) \text{ for all } A \in \mathcal{F}\right\} \\ &= \left\{\mu_{h,t} \mid \mu_{h,t} = \frac{\mu(D_t \cup \cdots \cup D_T)}{\mu(D_h \cup \cdots \cup D_T)} \text{ for all } \mu \in core\,(v)\right\} \end{aligned}$$

whereby the conditional survival belief $\nu_{h,t}$ is formed by the following Bayesian rule.

**Generalized Bayesian updating rule for survival beliefs:**

$$\nu_{h,t} = \frac{\nu(D_t \cup \cdots \cup D_T)}{\nu(D_t \cup \cdots \cup D_T) + 1 - \nu(D_0 \cup \cdots \cup D_{h-1} \cup D_t \cup \cdots \cup D_T)}. \tag{7}$$

Observe that (i) the pessimistic neo-additive belief (5) is convex and that (ii) an application of the Generalized Bayesian update rule to this belief results in the same conditional survival belief as (6). Consequently, the CEU agents whose conditional survival beliefs are given by (6) can be equivalently modelled as an ambiguity averse multiple priors decision maker who applies the Full Bayesian update rule.

## 3 The Choquet Expected Utility Life-Cycle Model

Denote by

$$\mathbf{c} = (c_h, c_{h+1}..., c_T) \tag{8}$$

a consumption plan such that $c_k \geq \eta > 0$, for all $k \in \{h, \ldots, T\}$ whereby the lower bound $\eta$ is chosen to be non-binding in an optimum. An agent who consumes in accordance with (8) and dies at the end of age $t \geq h$ obtains the truncated consumption stream $\mathbf{c}^t = (c_h, \ldots, c_t)$ as consequence. Defining the set of consequences $X$ as the set of all truncated consumption streams allows us to interpret a consumption plan (8) as a mapping from the relevant state space into the set of consequences, i.e., $\mathbf{c} : \Omega \backslash \{\omega_0, ..., \omega_{h-1}\} \rightarrow X$ such that

| | $\omega_h$ | $\omega_{h+1}$ | $\cdots$ | $\omega_T$ |
|---|---|---|---|---|
| $\mathbf{c} = (c_h, c_1..., c_T)$ | $\mathbf{c}^h = (c_h)$ | $\mathbf{c}^{h+1} = (c_h, c_{h+1})$ | $\cdots$ | $\mathbf{c}^T = (c_h, c_{h+1}..., c_T)$ |

That is, we interpret consumption plans as Savage (1954) acts whose deterministic consequences are truncated consumption streams. The states $\{\omega_0, ..., \omega_{h-1}\}$ are irrelevant to the utility of the $h$-old agent as they have become impossible. We assume that the decision maker prefers to live (i.e., to consume) longer, that is, we assume the following preference ranking over consequences for any $h$-old agent:

$$\mathbf{c}^T \succeq \cdots \succeq \mathbf{c}^h.$$

Denote by $\{\Omega_0, .., \Omega_m\} \subseteq \mathcal{F}$ a finite partition of the state space $\Omega$ such that we have for a measurable real-valued function $f$

$$f(\Omega_0) \geq \cdot \cdot \cdot \cdot \geq f(\Omega_m).$$

The *Choquet integral* of $f$ with respect to the non-additive probability measure $\nu^h$ on $(\Omega, \mathcal{F})$ becomes (Schmeidler 1986):

$$\int f d\nu^h = \sum_{j=0}^{m} f(\Omega_j) \left[ \nu^h (\Omega_0, .., \Omega_j) - \nu^h (\Omega_0, .., \Omega_{j-1}) \right]$$

where $\nu^h (\Omega_0, \Omega_{-1}) = 0$. Letting $f$ be a Bernoulli utility function defined over truncated consumption streams results in the following definition of Choquet expected utility over consumption plans.

**Definition 2.** *The Choquet expected utility (CEU) of consumption plan* $\mathbf{c}$ *of an h-old agent is given as*

$$U\left(\mathbf{c},\nu^{h}\right)=\sum_{j=0}^{T-h}U^{h}\left(\mathbf{c}^{T-j}\right)\left[\nu_{h,T-j}-\nu_{h,T-j+1}\right] \tag{9}$$

*where* $U^{h}\left(\cdot\right)$ *is a Bernoulli utility function over truncated consumption streams satisfying*

$$U^{h}\left(\mathbf{c}^{T}\right)\geq\cdots\geq U^{h}\left(\mathbf{c}^{0}\right). \tag{10}$$

We follow the literature (cf. Epper et al. 2011; Andreoni and Sprenger 2012) and distinguish between a pure time-discount factor and the agent's survival belief. Denote by $\beta_{h,k}\in(0,1]$, $k=h,...,T$, the pure time-discount factors of an $h$-old agent such that $\beta_{h,h}=1$ and $\beta_{h,k}\geq\beta_{h,k+1}$. Note that our notion of time-discount factors is extremely general so that it encompasses, e.g., all notions of *hyperbolic* and *quasi-hyperbolic time-discounting* (cf. Phelps and Pollak 1968; Laibson 1997; O'Donoghue and Rabin 1999).

**Assumption 1.** *The Bernoulli utility of a truncated consumption stream* $c^{h+t}$ *is additively separable with pure time-discount factors, i.e.,*

$$U^{h}\left(\mathbf{c}^{h+t}\right)=\sum_{k=h}^{h+t}\beta_{h,k}u\left(c_{k}\right)$$

*for a strictly increasing period-utility function* $u:[\eta,\infty)\rightarrow R_{+}$.

By Assumption 1, we can transform the CEU from a consumption plan as follows

$$
\begin{aligned}
& U\left(\mathbf{c},\nu^h\right) \\
&= \sum_{j=0}^{T} U^h\left(\mathbf{c}^{h+T-j}\right)\left[\nu_{h,T-j}-\nu_{h,T-j+1}\right] \\
&= \left(\sum_{k=h}^{T}\beta_{h,k}u\left(c_k\right)\right)\nu_{h,T} + \left(\sum_{k=h}^{T-1}\beta_{h,k}u\left(c_k\right)\right)\left[\nu_{h,T-1}-\nu_{h,T}\right]+\cdots \\
&= \left[\left(\sum_{k=h}^{T}\beta_{h,k}u\left(c_k\right)\right)-\left(\sum_{k=h}^{T-1}\beta_{h,k}u\left(c_k\right)\right)\right]\nu_{h,T} + \left[\left(\sum_{k=h}^{T-1}\beta_{h,k}u\left(c_k\right)\right)-\left(\sum_{k=h}^{T-2}\beta_{h,k}u\left(c_k\right)\right)\right]\nu_{h,T-1}+\cdots \\
&= \sum_{t=h}^{T}\beta_{h,t}\nu_{h,t}u\left(c_t\right).
\end{aligned}
$$

**Proposition 1.** *Under Assumption 1, the CEU (9) of the h-old agent from consumption plan* $\mathbf{c}=(c_h,...,c_T)$ *is equivalently given as*

$$U\left(\mathbf{c},\rho^h\right)=\sum_{t=h}^{T}\rho_{h,t}u\left(c_t\right) \tag{11}$$

*such that the effective discount factors* $\rho^h = \left(\rho_{h,t},...,\rho_{h,T}\right)$ *of the h-old agent are defined as*

$$\rho_{h,t}=\beta_{h,t}\nu_{h,t}.$$

To investigate parametric deviations from the logarithmic period utility function considered in Pollak (1968), we consider the family of isoelastic power per period utility functions:

**Assumption 2.** *The period-utility function satisfies*

$$u\left(c\right)=\chi+\begin{cases}\frac{c^{1-\theta}}{1-\theta} & \text{for } \theta\neq 1\\ \ln(c) & \text{for } \theta=1.\end{cases} \tag{12}$$

*for a given concavity parameter* $\theta>0$. *The constant* $\chi\geq 0$ *has to ensure that* $u\left(\eta\right)\geq 0$.[10]

[10] If $u\left(c\right)<0$, the ranking condition (10), which is crucial to the definition of CEU, could be violated. For $\theta<1$ the period utility function is positive so that $\chi$ can be set to zero. For $\theta\geq 1$ we can set $\chi=-\frac{\eta^{1-\theta}}{1-\theta}$ or $\chi=-\ln\left(\eta\right)$, respectively. This explains the role of the lower boundary $\eta>0$.

There exists an initial amount of total wealth $w_0 > 0$ that the agent can spend over her life-cycle. Total wealth includes financial assets and the discounted value of future labor income (human capital wealth). That is, our analysis encompasses models with a risk-free labor income stream in absence of any borrowing constraints. Without loss of generality we assume that the interest rate is zero and thus make the following

**Assumption 3.** *The budget constraint is given by*

$$w_{t+1} = w_t - c_t \geq 0 \quad \textit{for } t \in \{0, 1, \ldots, T-1\}\,. \tag{13}$$

Throughout this paper we refer to (11) with isoelastic period-utility function (12) and budget constraint (13) as the *CEU life-cycle model.*

# 4 Solving the Model

## 4.1 Optimal Consumption Plan and Dynamic Consistency

For fixed period consumption $c_t$ and wealth $w_t$ let

$$c_t = m_t w_t$$

where $m_t$ denotes the agent's *marginal propensity to consume* (MPC). Because the optimal period consumption is linear in total wealth for isoelastic period utility functions, it will be convenient to consider MPCs rather than absolute consumption levels when we analyze saving behavior.[11] Expressed in terms of MPCs for the periods $h+1, ..., T$ and period $h$ wealth the CEU (11) of the $h$-old agent from consumption plan $\mathbf{c} = (c_h, ..., c_T)$ becomes

$$U\left((c_h; m_{h+1}, ..., m_T, w_h), \rho^h\right) = u(c_h) + \sum_{t=h+1}^{T} \rho_{h,t} u\left((w_h - c_h)\, m_t \prod_{j=h+1}^{t-1} (1 - m_j)\right).$$

[11] Linearity of consumption policy functions in models with a deterministic labor income stream and no borrowing constraints is a well-established result in the consumption literature, cf., e.g., Deaton (1992).

Next we derive the MPCs that would maximize this utility function from the perspective of the $h$-old agent.

For $h = T$, we trivially have as optimal consumption

$$c_T^{*,h}(w_T) = m_T^{*,h} w_T$$

with optimal MPC $m_T^{*,h} = 1$. For $h < T$, the optimal period $h$ consumption $c_h^{*,h}$ from the perspective of the $h$-old agent is pinned down by the following FOC:

$$\left.\frac{dU\left(\left(c_h; m_{h+1}, ..., m_T, w_h\right), \rho^h\right)}{dc_h}\right|_{c_h = c_h^{*,h}} = 0$$
$$\Leftrightarrow$$
$$u'(c_h^{*,h}) = \sum_{t=h+1}^{T} \rho_{h,t} u'\left(\left(w_h - c_h^{*,h}\right) m_t \prod_{j=h+1}^{t-1} (1 - m_j)\right)\left(m_t \prod_{j=h+1}^{t-1} (1 - m_j)\right),$$

which becomes for the power period utility function

$$\left(c_h^{*,h}\right)^{-\theta} = \left(w_h - c_h^{*,h}\right)^{-\theta} \sum_{t=h+1}^{T} \rho_{h,t} \left(m_t \prod_{j=h+1}^{t-1} (1 - m_j)\right)^{1-\theta}.$$

Solving for $c_h^{*,h}$ results in

$$c_h^{*,h}(m_{h+1}, ..., m_T, w_h) = m_h^{*,h}(m_{h+1}, ..., m_T) w_h$$

such that the optimal period $h$ MPC for fixed period $h+1, ..., T$ MPCs is given as

$$m_h^{*,h}(m_{h+1}, ..., m_T) = \frac{1}{1 + \left(\sum_{t=h+1}^{T} \rho_{h,t} \left(m_t \prod_{j=h+1}^{t-1} (1 - m_j)\right)^{1-\theta}\right)^{\frac{1}{\theta}}}.$$

More generally, by the envelope theorem, the optimal period $t \geq h$ consumption from the perspective of the $h$-old agent given fixed values of $m_{t+1}, ..., m_T$ and wealth $w_t$ is pinned down by

$$\rho_{h,t}\left(c_t^{*,h}\right)^{-\theta} = \left(w_t - c_t^{*,h}\right)^{-\theta} \sum_{s=t+1}^{T} \rho_{h,s} \left(m_t \prod_{j=t+1}^{s-1} (1 - m_j)\right)^{1-\theta}.$$

This gives us the following

**Lemma 1.** *The MPCs* $m_t^{*,h}$ *that are optimal from the perspective of the h-old agent for fixed* $m_{t+1}, ..., m_T$ *are given as*

$$m_t^{*,h}(m_{t+1},...,m_T) = \begin{cases} 1 & \text{for } t = T \\ \frac{1}{1+\left(\sum_{s=t+1}^{T} \frac{\rho_{h,s}}{\rho_{h,t}} \left(m_s \prod_{j=t+1}^{s-1}(1-m_j)\right)^{1-\theta}\right)^{\frac{1}{\theta}}} & \text{for } h \leq T-1 \end{cases}$$

If the CEU life-cycle model was dynamically consistent, the optimal MPCs of Lemma 1 would also characterize the realized MPCs over the agent's life cycle. From the first-order conditions

$$\frac{u'\left(c_t^{*,h}\right)}{u'\left(c_{t+1}^{*,h}\right)} = \frac{\rho_{h,t+1}}{\rho_{h,t}}$$

of the $h$-old agent and of the $t$-old agent

$$\frac{u'\left(c_t^{*,t}\right)}{u'\left(c_{t+1}^{*,t}\right)} = \rho_{t,t+1},$$

respectively, one can see that the CEU life-cycle model is dynamically consistent if and only if, for all $h$,

$$\frac{\rho_{h,t+1}}{\rho_{h,t}} = \rho_{t,t+1} \text{ for all } t > h. \tag{14}$$

Assuming exponential time discounting $\beta_{h,t} = \beta^{t-h}$ in (14) gives

$$\begin{aligned} \frac{\beta_{h,t+1}\nu_{h,t+1}}{\beta_{h,t}\nu_{h,t}} &= \beta_{t,t+1}\nu_{t,t+1} \\ &\Leftrightarrow \\ \frac{\nu_{h,t+1}}{\nu_{h,t}} &= \nu_{t,t+1}, \end{aligned}$$

which is exactly the optimistic update rule (4).

**Proposition 2.** *If time-discounting is exponential, then the CEU life-cycle model is dynamically consistent if and only if the agent is a Bayesian decision maker who applies the optimistic update rule.*

## 4.2 Dynamic Inconsistency: Naive Versus Sophisticated Agent

The CEU life-cycle model is dynamically inconsistent whenever there is some $t$-old agent with $t > h$ who would like to deviate in period $t$ from the $h$-old agent's optimal consumption plan. In that case the optimal MPCs of Lemma 1 will no longer characterize the realized consumption stream. To derive the MPCs that correspond to the realized consumption stream, we need to clarify how the agent deals with dynamic consistency. For this purpose, we follow the literature and distinguish between the two extreme cases of a naive versus a sophisticated agent (cf. O'Donoghue and Rabin 1999).

The sophisticated agent correctly anticipates at every age $h$ her future behavior. Denote by $m_t^s$ the realized MPC of the $t$-old sophisticated agent. Expressed in terms of the optimal MPCs of Lemma 1, the sophisticated agent solves at every age $h \geq 0$ the problem

$$m_h^s = m_h^{*,h}\left(m_{h+1}^s, ..., m_T^s\right),$$

which gives us immediately a recursive characterization of the realized MPCs of the sophisticated agent.

**Theorem 1.** *The realized MPCs of the sophisticated agent are recursively characterized as follows:*

$$m_h^s = \begin{cases} 1 & \text{for } h = T \\ \frac{1}{1+\left(\sum_{t=h+1}^{T} \rho_{h,t}\left(m_t^s \prod_{j=h+1}^{t-1}\left(1-m_j^s\right)\right)^{1-\theta}\right)^{\frac{1}{\theta}}} & \text{for } h \leq T-1. \end{cases}$$

Turn now to the naive agent who is completely unaware of her dynamically inconsistent preferences. Expressed in terms of the optimal MPCs, the $h$-old naive agent's *planned* MPCs for $t \geq h$ are characterized as

$$m_t^{n,h} = m_t^{*,h}\left(m_{t+1}^{n,h}, ..., m_T^{n,h}\right).$$

Equivalently, the 0-year old naive agent's *planned* MPCs are pinned down by the following FOCs for all $t$ such that $h \le t < T$:

$$\begin{aligned}
\rho_{h,t}(m_t^{n,h} w_t)^{-\theta} &= \rho_{h,t+1}\left(m_{t+1}^{n,h} w_{t+1}\right)^{-\theta} \\
&\Leftrightarrow \\
\rho_{h,t}(m_t^{n,h} w_t)^{-\theta} &= \rho_{h,t+1}\left(m_{t+1}^{n,h}\left(w_t - m_t^{n,h} w_t\right)\right)^{-\theta} \\
&\Leftrightarrow \\
m_t^{n,h} &= \frac{1}{1+\left(\frac{\rho_{h,t+1}}{\rho_{h,t}}\right)^{\frac{1}{\theta}}\left(m_{t+1}^{n,h}\right)^{-1}}.
\end{aligned} \tag{15}$$

Substituting

$$m_{t+1}^{n,h} = \frac{1}{1+\left(\frac{\rho_{h,t+2}}{\rho_{h,t+1}}\right)^{\frac{1}{\theta}}\left(m_{t+2}^{n,h}\right)^{-1}}$$

in (15) gives

$$m_t^{n,h} = \frac{1}{1+\left(\frac{\rho_{h,t+1}}{\rho_{h,t}}\right)^{\frac{1}{\theta}}+\left(\frac{\rho_{h,t+2}}{\rho_{h,t}}\right)^{\frac{1}{\theta}}\left(m_{t+2}^{n,h}\right)^{-1}}.$$

By repeating this argument until $m_T^{n,h} = 1$, we obtain the following closed form description of planned MPCs

$$m_t^{n,h} = \begin{cases} 1 & \text{for } t = T \\ \frac{1}{1+\sum_{k=t+1}^{T}\left(\frac{\rho_{h,k}}{\rho_{h,t}}\right)^{\frac{1}{\theta}}} & \text{for } t \le T-1. \end{cases}$$

Let us summarize the above argument, whereby we write $m_h^n = m_h^{n,h}$ for the realized MPCs of the $h$-old naive agent:

**Theorem 2.** *The realized MPCs of the naive agent are given as follows:*

*(i) Recursive characterization:*

$$m_h^n = \begin{cases} 1 & \textit{for } h = T \\ \frac{1}{1+\left(\sum_{t=h+1}^{T}\rho_{h,t}\left(m_t^{n,h}\prod_{j=h+1}^{t-1}\left(1-m_j^{n,h}\right)\right)^{1-\theta}\right)^{\frac{1}{\theta}}} & \textit{for } h \le T-1 \end{cases}$$

*with planned MPCs*

$$m_t^{n,h} = \begin{cases} 1 & \textit{for } t = T \\ \frac{1}{1+\sum_{k=t+1}^{T}\left(\frac{\rho_{h,k}}{\rho_{h,t}}\right)^{\frac{1}{\theta}}} & \textit{for } t \leq T-1 \end{cases}$$

*(ii) Closed form:*

$$m_h^n = \frac{1}{1+\sum_{t=h+1}^{T}\left(\rho_{h,t}\right)^{\frac{1}{\theta}}} \textit{ for } h \leq T-1.$$

## 4.3 Illustration: The Four-Period Model

We illustrate the respective solutions to our CEU life-cycle model for the four-period model, where $T = 3$:

**Corollary 1.** *For the sophisticated agent the realized MPCs become:*

$$\begin{aligned} m_3^s &= 1, \\ m_2^s &= \frac{1}{1+\left(\rho_{2,3}\right)^{\frac{1}{\theta}}}, \\ m_1^s &= \frac{1}{1+\left[\rho_{1,2}\left(m_2^s\right)^{1-\theta}+\rho_{1,3}\left(m_3^s\left(1-m_2^s\right)\right)^{1-\theta}\right]^{\frac{1}{\theta}}}, \\ m_0^s &= \frac{1}{1+\left[\rho_{0,1}\left(m_1^s\right)^{1-\theta}+\rho_{0,2}\left(m_2^s\left(1-m_1^s\right)\right)^{1-\theta}+\rho_{0,3}\left(m_3^s\left(1-m_2^s\right)\left(1-m_1^s\right)\right)^{1-\theta}\right]^{\frac{1}{\theta}}}. \end{aligned}$$

Solving the model for the sophisticated agent through backward induction is equivalent to solving an extensive form game for the unique subgame-perfect Nash equilibrium whereby the agents of different ages are different players who can choose MPCs at each information node. The only way how an agent can influence through her chosen MPC the future consumption path in her favor is by restricting the budget, i.e., wealth level, of her future agents. The MPC $m_0^s$—being a best response of the 0-old agent against the correctly anticipated MPCs of her future selfs—is therefore a function in $m_1^s$, $m_2^s$, and $m_3^s$. On the other hand, the MPCs of future agents do not depend on previously chosen MPCs. This is a consequence of the fact that optimal MPCs are independent of wealth levels for power period utility functions.

**Corollary 2.** *For the naive agent the realized MPCs become:*

$$m_3^n = 1,$$
$$m_2^n = \frac{1}{1+\left(\rho_{2,3}\right)^{\frac{1}{\theta}}},$$
$$m_1^n = \frac{1}{1+\left[\rho_{1,2}\left(m_2^{n,1}\right)^{1-\theta}+\rho_{1,3}\left(m_3^{n.1}\left(1-m_2^{n,1}\right)\right)^{1-\theta}\right]^{\frac{1}{\theta}}},$$
$$m_0^n = \frac{1}{1+\left[\rho_{0,1}\left(m_1^{n,0}\right)^{1-\theta}+\rho_{0,2}\left(m_2^{n,0}\left(1-m_1^{n,0}\right)\right)^{1-\theta}+\rho_{0,3}\left(m_3^{n,0}\left(1-m_2^{n,0}\right)\left(1-m_1^{n,0}\right)\right)^{1-\theta}\right]^{\frac{1}{\theta}}}.$$

*whereby the planned MPCs are: for the* 1*-old agent:*

$$m_3^{n.1} = 1,$$
$$m_2^{n,1} = \frac{1}{1+\left(\frac{\rho_{1,3}}{\rho_{1,2}}\right)^{\frac{1}{\theta}}};$$

*and for the* 0*-old agent:*

$$m_3^{n,0} = 1,$$
$$m_2^{n,0} = \frac{1}{1+\left(\frac{\rho_{0,3}}{\rho_{0,2}}\right)^{\frac{1}{\theta}}},$$
$$m_1^{n,0} = \frac{1}{1+\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}+\left(\frac{\rho_{0,3}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}}.$$

The realized versus planned MPCs illustrate how the dynamic inconsistency of the model plays out for the naive agent. Focus, e.g., on period 2 and observe that the realized MPC $m_2^n$ coincides with the previously planned MPCs $m_2^{n,1}$ and $m_2^{n,0}$ if and only if

$$\rho_{2,3} = \frac{\rho_{1,3}}{\rho_{1,2}} = \frac{\rho_{0,3}}{\rho_{0,2}}.$$

This is, in general, only the case if the agent is a Bayesian decision maker who applies the optimistic update rule and who discounts time exponentially. If we have instead that $\rho_{2,3} > \frac{\rho_{1,3}}{\rho_{1,2}}$, i.e., if discounting exhibits *increasing patience* so that the marginal valuation of saving increases as the agent ages, then the 2-old naive agent will be consuming strictly less than the 1-old agent had originally planned for period 2, i.e., $m_2^{n,1} > m_2^n$.

Turning back to the multi-period model, increasing patience means that the marginal valuation of savings increases when the decision maker ages, i.e., from $h$ to $h+1$, as in the quasi-hyperbolic time discounting model. In fact, it is straightforward to derive from the comparison of plans in Theorem 2 for the multi-period model a very similar condition. Observe that $m_{h+1}^{n,h} > m_{h+1}^n$ for all $h = 0, \ldots, T-2$ if and only if

$$\sum_{k=h+2}^{T} \left( \frac{\rho_{h,k}}{\rho_{h,h+1}} \right)^{\frac{1}{\theta}} < \sum_{k=h+2}^{T} \rho_{h+1,k}^{\frac{1}{\theta}} \tag{16}$$

for which a sufficient condition is that for all $k = h+2, \ldots, T$

$$\left( \frac{\rho_{h,k}}{\rho_{h,h+1}} \right)^{\frac{1}{\theta}} \leq \rho_{h+1,k}^{\frac{1}{\theta}} \qquad \Leftrightarrow \qquad \frac{\rho_{h,k}}{\rho_{h,h+1}} \leq \rho_{h+1,k}$$

with strict inequality for at least one period $k$. We summarize this insight by first defining impatience more generally in a subsequent proposition:

**Definition 3** (Increasing Patience)**.** *We say that a discount function exhibits strictly increasing patience between ages $h$ and $h+1$ with respect to future periods $k > t \geq h+1$ if and only if*

$$\frac{\rho_{h,k}}{\rho_{h,t}} < \frac{\rho_{h+1,k}}{\rho_{h+1,t}}.$$

**Proposition 3.** *If at all ages $h = 0, \ldots, T-2$ the discount function exhibits (weakly) increasing patience with respect to future periods $k > t = h+1$ so that*

$$\frac{\rho_{h,k}}{\rho_{h,h+1}} \leq \rho_{h+1,k} \tag{17}$$

*with the inequality being strict for at least one period $k$ then the naive decision maker consumes less at all ages $h+1$ than planned at ages $h$, i.e., $m_{h+1}^{n,h} > m_{h+1}^n$.*

Notice that condition (17) is only a sufficient condition for the downward adjustment of the MPC as the naive decision maker ages, which is easier to interpret in terms of increasing patience than the necessary and sufficient condition (16). Thus, the revisal of plans of the naive decision maker is mainly[12] driven by the behavior of the discount function. The next section, also for the multi-period model, compares the realized saving behavior of the sophisticated and the naive agent to show that this comparison of realizations is instead exclusively driven by $\theta$, the resistance to intertermporal substitution.

# 5 Naive Versus Sophisticated Savings Behavior

If the CEU life-cycle model is dynamically consistent, the sophisticated and the naive agent's saving behavior coincide, by definition, with the optimal consumption plan of Lemma 1. Remarkably, even if the CEU life-cycle model is dynamically inconsistent both types of agents exhibit the same saving behavior whenever the period-utility function is of the logarithmic form. This can be easily seen by setting $\theta = 1$ in the MPCs of Theorems 1 and 2.

**Proposition 4.** *If the period utility function is logarithmic, i.e., $\theta = 1$, then the realized MPCs of the sophisticated and the naive agent are identically given as*

$$m_h^s = m_h^n = \begin{cases} 1 & \text{for } h = T \\ \frac{1}{1+\sum_{t=h+1}^{T} \rho_{h,t}} & \text{for } h \leq T-1. \end{cases}$$

Formally, the finding of Proposition 4 is already implied by the seminal analysis in Pollak (1968) for deterministic life-cycle models with arbitrary discount factors. Applied to our CEU life-cycle model, Proposition 4 comes with the surprising implication that the beliefs of CEU decision makers have no impact on the fact that the sophisticated agent and her naive counterpart exhibit the same saving behavior when the period-utility function is logarithmic. As Pollak's (1968) original analysis only applies to a logarithmic period-utility function, the question arises what can be said about naive versus sophisticated saving behavior for general power period-utility functions with $\theta \neq 1$.

[12]"Mainly" in the sense that increasing patience is a sufficient condition in the multi-period model, whereas the necessary and sufficient condition (16) is expressed also in terms of $\theta$.

### 5.1 Analytical Approach

To tackle this question analytically, it is insightful to rewrite the MPCs of the sophisticated agent from Theorem 1 as follows (see Appendix A.1 for details of the derivation):

**Lemma 2.** *The realized MPCs of the sophisticated agent are equivalently given as*

$$m_h^s = \frac{1}{1 + \frac{\left(\rho_{h,h+1}\lambda_h + \rho_{h,h+2}\xi_h\right)^{\frac{1}{\theta}}}{m_{h+1}^s}} \tag{18}$$

*where*

$$\begin{aligned}
\lambda_h &= m_{h+1}^s + \frac{\rho_{h,h+2}}{\rho_{h,h+1}\rho_{h+1,h+2}}\left(1 - m_{h+1}^s\right) \\
\xi_h &= \left(1 - m_{h+1}^s\right)^{1-\theta}\left(m_{h+1}^s\right)^{\theta}\left(\zeta_{h+2}^h - \zeta_{h+2}^{h+1}\right) \\
\zeta_t^h &= \begin{cases} 1 & \textit{for } t = T \\ \left(m_t^s\right)^{1-\theta} + \frac{\rho_{h,t+1}}{\rho_{h,t}}\left(1 - m_t^s\right)^{1-\theta}\zeta_{t+1}^h & \textit{else.} \end{cases}
\end{aligned}$$

To interpret these expressions, observe that the effective discount factor of the sophisticated agent is $\rho_{h,h+1}\lambda_h + \rho_{h,h+2}\xi_h$ and thus, in addition to the raw time discount factor $\rho_{h,h+1}$, features two adjustment terms $\lambda_h$ and $\xi_h$ (where the latter is discounted with $\rho_{h,h+2}$). The first factor $\lambda_h$ is analogous to the role of the generalized Euler equation in the quasi-hyperbolic time discounting model (Harris and Laibson 2001). With increasing patience, cf. Definition 3, the term $\frac{\rho_{h,h+2}}{\rho_{h,h+1}\rho_{h+1,h+2}} > 1$ so that $\lambda_h > 1$ which increases the marginal value of savings. At the same time $\lambda_h$ varies inversely with the marginal propensity to consume $m_{h+1}^s$ in the next period. If the next period's self's MPC increases then the current period sophisticated agent will respond by increasing impatience to constrain her own future self by consuming more today, thus saving less and shifting fewer resources to the next period. Term $\xi_h$ represents a closed form expression of the "adjustment factor" introduced by Groneck et al. (2016) for a stochastic model. It increases in the difference between the marginal valuation of wealth between self $h$ and self $h+1$ in period $h+2$. As the proof of Lemma 2 shows, the marginal valuation of wealth is reflected in the terms $\zeta_t^h$. If the difference $\zeta_{h+2}^h - \zeta_{h+2}^{h+1}$ is positive, then self $h$ values wealth in the future period $h+2$ more than self $h+1$, which induces self $h$ to save more.

Comparing the saving behavior of the sophisticated and the naive agent becomes analytically intractable whenever the "adjustment factor" $\xi_h$ does not vanish in (18). This adjustment term is only relevant for models with at least four periods because of $\zeta_2^0 - \zeta_2^1 = 1 - 1 = 0$ for $T = 2$. Suppose now that $T > 2$. Then observe that

$$\zeta_{h+2}^{h} = \left(m_{h+2}^{s}\right)^{1-\theta} + \frac{\rho_{h,h+3}}{\rho_{h,h+2}} \left(1 - m_{h+2}^{s}\right)^{1-\theta} \zeta_{h+3}^{h}$$
$$\zeta_{h+2}^{h+1} = \left(m_{h+2}^{s}\right)^{1-\theta} + \frac{\rho_{h+1,h+3}}{\rho_{h+1,h+2}} \left(1 - m_{h+2}^{s}\right)^{1-\theta} \zeta_{h+3}^{h+1}$$

and thus the difference in the marginal valuation of wealth in period $h + 2$ between selves $h$ and $h + 1$ is given by

$$\zeta_{h+2}^{h} - \zeta_{h+2}^{h+1} = \left(1 - m_{h+2}^{s}\right)^{1-\theta} \left(\frac{\rho_{h,h+3}}{\rho_{h,h+2}}\zeta_{h+3}^{h} - \frac{\rho_{h+1,h+3}}{\rho_{h+1,h+2}}\zeta_{h+3}^{h+1}\right).$$

Now suppose that $\frac{\rho_{h,h+3}}{\rho_{h,h+2}} = \frac{\rho_{h+1,h+3}}{\rho_{h+1,h+2}}$ so that we get in the above

$$\zeta_{h+2}^{h} - \zeta_{h+2}^{h+1} = \left(1 - m_{h+2}^{s}\right)^{1-\theta} \frac{\rho_{h,h+3}}{\rho_{h,h+2}} \left(\zeta_{h+3}^{h} - \zeta_{h+3}^{h+1}\right).$$

Next, notice that the difference in the marginal valuation of wealth in period $h + 3$ between selves $h$ and $h + 1$ is in turn given by

$$\zeta_{h+3}^{h} = \left(m_{h+3}^{s}\right)^{1-\theta} + \frac{\rho_{h,h+4}}{\rho_{h,h+3}} \left(1 - m_{h+3}^{s}\right)^{1-\theta} \zeta_{h+3}^{h}$$
$$\zeta_{h+3}^{h+1} = \left(m_{h+3}^{s}\right)^{1-\theta} + \frac{\rho_{h+1,h+4}}{\rho_{h+1,h+4}} \left(1 - m_{h+3}^{s}\right)^{1-\theta} \zeta_{h+3}^{h+1}$$

and so forth. Thus, we find that at any age $h$ the term $\xi_h$ becomes zero—leading to analytical tractability—if and only if

$$\frac{\rho_{h,t+1}}{\rho_{h,t}} = \frac{\rho_{h+1,t+1}}{\rho_{h+1,t}}$$

for all $t = h + 2, \ldots, T - 3$, i.e., if the discount function between all future periods $t$ and $t + 1$ with $t \geq h + 2$ features "constant patience", cf. Definition 3.

The proof of the next theorem analyzes how the MPCs of the sophisticated agent deviate from the naive's MPCs when we move the effective discount factors away—through the manipulation of $\epsilon$-parameters—from the benchmark case of dynamic consistency for which the MPCs of both agent types coincide. Because the formal proof is rather complex, we relegate it to the Appendix.[13]

**Theorem 3.** *Consider a dynamically inconsistent CEU life-cycle model such that either (i) $T = 2$ or (ii) $T > 2$ and the effective discount factors satisfy, for all $h = 0, ..., T-3$ the linearity condition (i.e., constant patience between all future periods $t$ and $t+1$)*

$$\frac{\rho_{h,t+1}}{\rho_{h,t}} = \frac{\rho_{h+1,t+1}}{\rho_{h+1,t}}. \tag{19}$$

*for all $t = h+1, \ldots, T-3$.*

1. *If $\theta > 1$, then the naive agent exhibits at every age $h = 0, .., T-2$ a strictly greater propensity to consume than her naive counterpart, i.e., $m_h^s < m_h^n$.*

2. *Conversely, if $\theta < 1$, then the sophisticated agent exhibits at every age $h = 0, .., T-2$ a strictly greater propensity to consume than her naive counterpart, i.e., $m_h^s > m_h^n$.*

We also refer to the restriction (19) on effective discount factors as a linearity condition because it is satisfied for the beliefs of an exponential time-discounter if

$$\nu_{h,t} = \gamma_h \mu_{h,t}$$

for some $h$-dependent constant $\gamma_h$. This is, for example the case if the agent is a Bayesian decision maker who applies the pessimistic or the Generalized Bayesian update rule to a pessimistic neo-additive belief so that (6) implies

$$\gamma_h = 1 - \frac{\delta}{\delta + (1-\delta)\,\mu_{0,h}}.$$

Also note that the restriction (19) is satisfied for the degenerate deterministic case, i.e., $\nu_{h,t} = 1$, if the agent is a quasi-hyperbolic time-discounter in the sense of Laibson (1997).

[13] The strict inequalities of Theorem 3 only apply to the ages $h = 0, ..., T-2$ because the MPCs of both agent types coincide for the final two periods, i.e., $h = T-1, T$.

## 5.2 Computational Approach

To gain insights about the marginal propensities of naive and sophisticated agents for general effective discount factors, we rely on Monte Carlo simulations. In particular, we analyze numerically whether Theorem 3 carries over to the multiperiod ($T > 2$) case when we drop the linearity restriction (19). We hence propose the following

**Conjecture 1.** *For all (arbitrary) specifications of the effective discount factors we have for the $T$-period model at every age $h$:*

*(i)* $m_h^n > m_h^s$ *if and only if* $\theta > 1$.

*(ii)* $m_h^n < m_h^s$ *if and only if* $\theta < 1$;

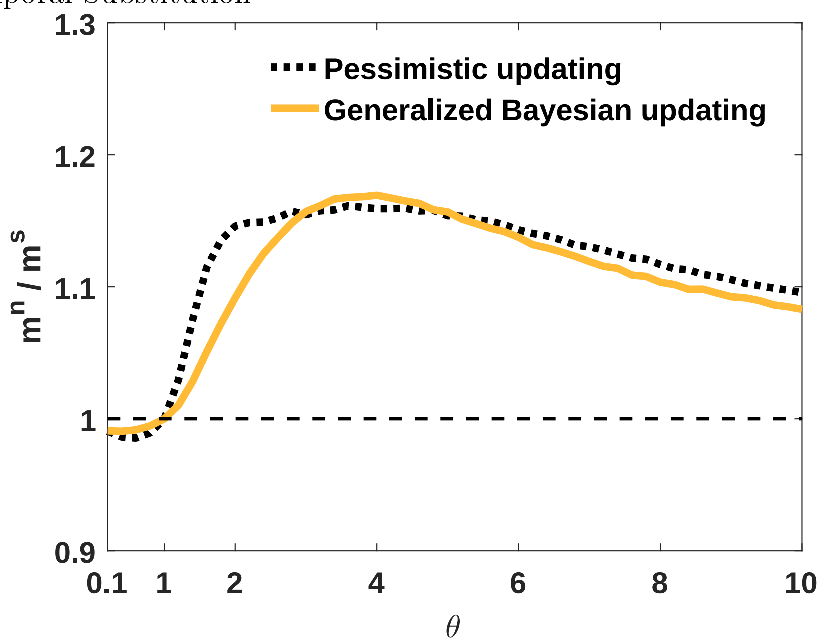
In order to verify if this conjecture is correct, we consider two variants of Monte Carlo simulations. In our first variant, we construct in each simulation a system of independently and identically distributed effective discount factors $\rho_{h,t}$ for all $h, t$. In our second variant, we consider discount functions as the product of a geometric time discount function $\beta^{t-h}$ and CEU survival beliefs $\nu_{h,t}$ that fulfill the set of assumptions of Definition 1 and are updated with the (i) pessimistic and the (ii) generalized Bayesian update rule, respectively. With regard to the time discount function we assume that $\beta \in \{0.5, 0.9, 1.0\}$ and construct $\rho_{h,t} = \beta^{t-h}\nu_{h,t}$. Details are described in Appendix B. In both variants we further take $T \in \{4, 10, 60\}$ and $\theta \in \{0.1, 0.5, 2, 4\}$. For all parameterizations we consider 10,000 Monte Carlo simulations and in each draw $c \in [1, \ldots, 10,000]$ we compute the age $h$ dependent MPCs, $m_{h,c}^n$ and $m_{h,c}^s$, cf. Theorems 1 and 2.

In all simulations (and both variants) our conjecture is confirmed: For $\theta > 1$ we always found that $m_{h,c}^n > m_{h,c}^s$ for all $c$ and for $\theta < 1$ we found the opposite. In Appendix B we show a graphical representation of the ratio $m_{0,c}^n / m_{0,c}^s$ of the marginal propensities in the first period $h = 0$ for both variants of Monte Carlo simulations for 1,000 draws. Although the ratio of marginal propensities fluctuates—depending on the draw of effective discount functions—it never crosses the threshold of one (either from below when $\theta < 1$ or above when $\theta > 1$).

Our next interest is in investigating whether differences in saving behavior across the two types of decision makers are quantitatively relevant. To this purpose we concentrate on the economically more interesting analyses in the second variant of Monte Carlo simulations and analyze the behavior of average MPCs (averaged across

Monte Carlo simulations) between the two types of decision makers. First, we compute the average ratio of the marginal propensities over a grid of $\theta \in [0.1, \ldots, 10]$. The results shown in Figure 1 highlight two asymmetric patterns. First, the choice of the update rule seems to matter; the generalized Bayesian update rule results in larger differences in the MPCs on average than the pessimistic update rule for most values of $\theta$. Second, the ratio differs quantitatively for values of $\theta$ below or above one. For values of $\theta$ below one, the MPCs of the two decision makers are closely aligned with the ratio being only slightly below one. On the contrary, for values of $\theta$ above one the MPC of the naive agent is on average more then 15 percent higher compared to the MPC of the sophisticated agent. Interestingly, the most pronounced differences quantitatively occur for values of $\theta$ around two, which are the empirically most relevant cases (cf., e.g., Hall 1988; Barsky et al. 1997).

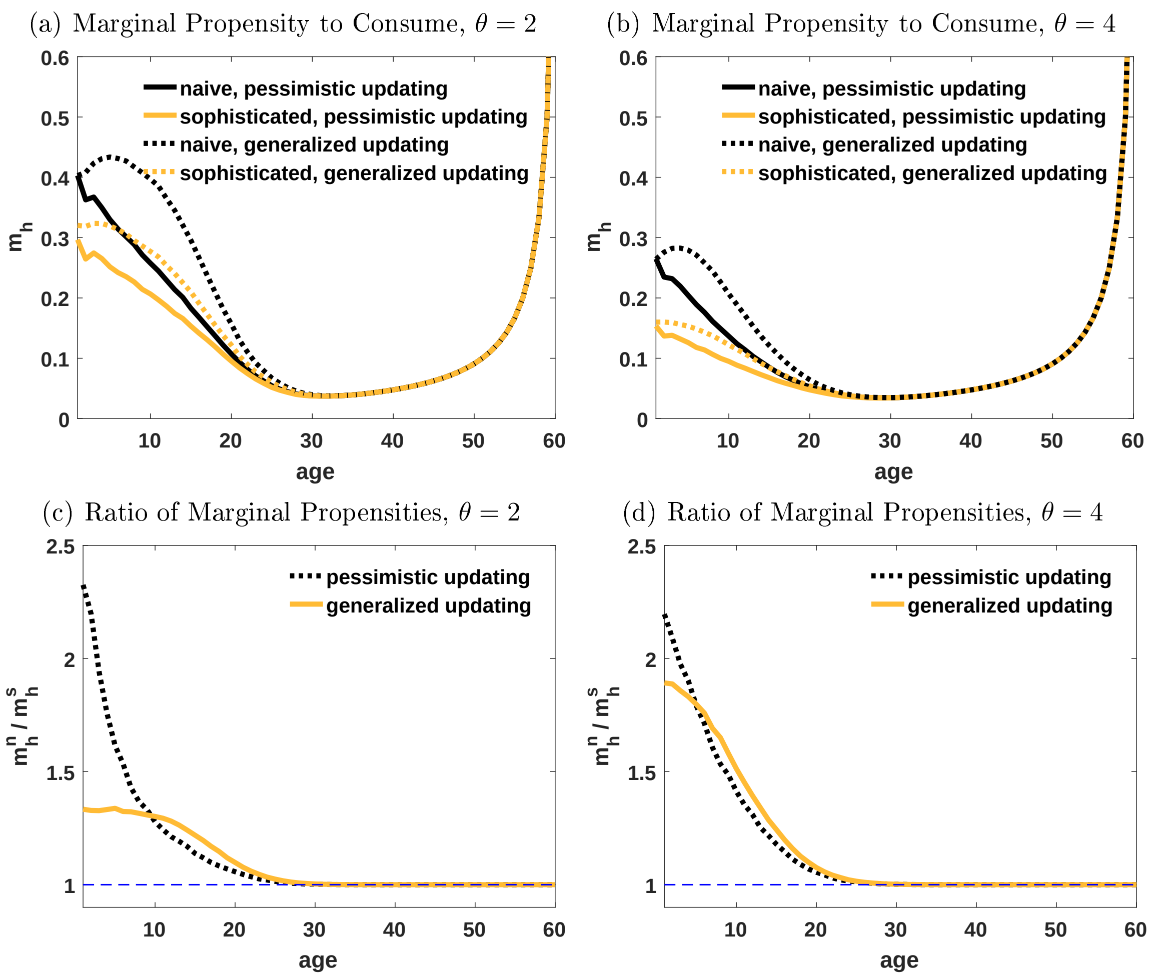
Figure 1: Average Ratio of Marginal Propensities to Consume: The Role of Resistance to Inter-temporal Substitution




Notes: Average ratio of MPCs over all ages and all simulations in the second variant of 10,000 Monte Carlo simulations where $\rho_{h,t} = \beta^{t-h}\nu_{h,t}$ for $\theta \in [0.1, \ldots, 10]$, $\beta = 1$, and $T = 60$.

Finally, Figure 2 shows the age profile of average MPCs for two different settings with $\theta = 2$ in Panel (a) and $\theta = 4$ in Panel (b). Observe that average ratios of MPCs between naive and sophisticated agents—displayed in Panels (c) and (d) of the figure—are particularly pronounced in the first half of the life cycle and that the average ratio of MPCs approaches one at about age 30. For $\theta = 2$ the average ratio of MPCs at the beginning of the life-cycle ($h = 0$) is at about 2.3 under pessimistic updating and at about 1.4 for generalized Bayesian updating. For a higher resistance to inter-temporal substitution of $\theta = 4$ the average ratio at $h = 0$ is about 2 for both updating rules. We can thus conclude that the differences in saving behavior across the two types of decision makers can be quantitatively very large.

Figure 2: Average Marginal Propensities to Consume over the Life-Cycle




Notes: Average MPCs of naive and sophisticated agents for the second variant with 10,000 Monte Carlo simulations where $\rho_{h,t} = \beta^{t-h}\nu_{h,t}$ for pessimistic and generalized Bayesian updating of survival beliefs $\nu_{h,t}$ with $\theta \in \{2, 4\}$, $\beta = 1$, and $T = 60$.

# 6 Concluding Remarks

We develop a Choquet expected utility (CEU) (Gilboa 1987; Schmeidler 1989) life-cycle model of consumption and savings giving rise to dynamically inconsistent consumption decisions. We derive consumption policy functions in closed form to formally show, as our main result, that sophisticated decision makers save more than their naive counterparts if and only if the per period isolelastic utility function features a resistance to intertemporal substituion $\theta$ above one, independent of the exact form of the time discount function. Empirical evidence for values of $\theta$ smaller than one is rather weak (c.f., e.g., Hall 1988, Barsky et al. 1997), and most researchers accordingly calibrate life-cycle models with values of $\theta$ above one. Thus, the conventional calibration of these models should (also in extended models) give rise to the result that sophisticates save more than naives. In a quantitative evaluation of our stylized model we further show that the differences in saving behavior across the two types of decision makers can be quantitatively large for conventional empirical estimates of $\theta$.

We focus on the simplest model with a deterministic income stream, a zero interest rate and no borrowing constraint. Our formal results hinge on the linearity of consumption in total wealth in this setup. As known from Samuelson (1969), linear policy functions generally arise in these types of models also when the return is stochastic and when there is portfolio choice. Thus, our formal analysis extends to a large class of economic models encompassing Epstein-Zin-Weil (EZW) recursive preferences (Epstein and Zin 1989; Epstein and Zin 1991; Weil 1989) with portfolio choice as well as models with risky human capital and a linear human capital technology as in Krebs (2003).[14] We thus conjecture that our main result on the ranking of marginal propensities to consume (MPC) out of total wealth of a sophisticated and a naive decision maker—i.e., that the MPC of the sophisticated decision maker is lower for all ages if and only if the coefficient of resistance to inter-temporal substitution is above one—continues to hold. In our future research we plan to employ calibrated variants of such extended models to further quantitatively evaluate the differences in saving behavior between the two types of decision makers.

[14] In models with portfolio choice and deterministic labor income the consumption savings decision can be separated from the portfolio choice decision and thus both types of agents choose the same portfolio share. With EZW recursive utility the structure of the recursive expressions is unaltered with the certainty equivalent of the stochastic (portfolio) return replacing the expressions for the return. Results are available upon request.

# A Analytical Appendix

## A.1 Derivation of Lemma 2

Lemma 2 is derived from the first-order condition of a recursive representation of the decision problem as in Groneck et al. (2016), which we first restate for our CEU model in Proposition 5. We subsequently proof the Lemma using the recursive structure.

### A.1.1 Generalized Euler Equation with Adjustment Factor

We here restate from the quantitative work in Groneck et al. (2016) the first-order condition of the sophisticated CEU decision maker for a general concave per-period utility function, i.e., we assume that the per period utility function $u(c)$ is strictly increasing, concave and twice continuously differentiable, i.e., $u_c > 0$, $u_{cc} < 0$. The next proposition states the inter-temporal optimality conditions for the sophisticated agent, which are derived recursively:

**Proposition 5.** *The sophisticated CEU decision maker's agent's inter-temporal* generalized *Euler equation with* adjustment factor *for consumption in period $h$, is given by*

$$u_c(c_h^s) = \rho_{h,h+1}\lambda_h u_c(c_{h+1}^s) + \rho_{h,h+2}\Xi_h \tag{20}$$

*where*

$$\lambda_h \equiv \frac{\partial c_{h+1}^s(w_{h+1})}{\partial w_{h+1}} + \frac{\rho_{h,h+2}}{\rho_{h,h+1}\cdot\rho_{h+1,h+2}}\left(1 - \frac{\partial c_{h+1}^s(w_{h+1})}{\partial w_{h+1}}\right) \tag{21}$$

*and*

$$\Xi_h \equiv \left(1 - \frac{\partial c_{h+1}^s(w_{h+1})}{\partial w_{h+1}}\right)\left(\frac{\partial V_{h+2}^h(w_{h+2})}{\partial w_{h+2}} - \frac{\partial V_{h+2}^{h+1}(w_{h+2})}{\partial w_{h+2}}\right) \tag{22}$$

*where $\frac{\partial c_{h+1}^s(w_{h+1})}{\partial w_{h+1}} \in (0,1)$ is the derivative of the consumption policy function in period $h+1$ and $\frac{\partial V_{h+2}^h(w_{h+2})}{\partial w_{h+2}}$ is the derivative of the value function in the subsequent period $h+2$.*

*Proof.* (cf. Groneck et al. (2016), Proposition 4.) The value functions of self $h$ in periods $h$ and $h+1$ are given by

$$V_h(w_h) = \max_{c_h, w_{h+1}} \left\{ u(c_h) + \rho^h_{h,h+1} V^h_{h+1}(w_{h+1}) \right\}$$
$$V^h_{h+1}(w_{h+1}) = u(c_{h+1}) + \frac{\rho_{h,h+2}}{\rho_{h,h+1}} V^h_{h+2}(w_{h+2}),$$

where it is important to note that the maximization operator in the second line is missing because this is not a maximized object from the perspective of current self $h$. For self $h+1$ we accordingly have

$$V^{h+1}_{h+1}(w_{h+1}) = \max_{c_{h+1}, w_{h+2}} \left\{ u(c_{h+1}) + \rho_{h+1,h+2} V^{h+1}_{h+2}(w_{h+2}) \right\}.$$

The first-order conditions with respect to consumption for selves $h$ and $h+1$ are:

$$u_c(c_h) = \rho_{h,h+1} \frac{\partial V^h_{h+1}(\cdot)}{\partial w_{h+1}} \tag{23a}$$
$$u_c(c_{h+1}) = \rho_{h+1,h+2} \frac{\partial V^{h+1}_{h+2}(\cdot)}{\partial w_{h+2}}. \tag{23b}$$

The derivative of the value function is

$$\begin{aligned} \frac{\partial V^h_{h+1}(\cdot)}{\partial w_{h+1}} &= u_c(c_{h+1}) \frac{\partial c_{h+1}}{\partial w_{h+1}} + \frac{\rho_{h,h+2}}{\rho_{h,h+1}} \left(1 - \frac{\partial c_{h+1}}{\partial w_{h+1}}\right) \frac{\partial V^h_{h+2}(\cdot)}{\partial w_{h+2}} \\ &= \frac{\partial c_{h+1}}{\partial w_{h+1}} \underbrace{\left( u_c(c_{h+1}) - \frac{\rho_{h,h+2}}{\rho_{h,h+1}} \frac{\partial V^h_{h+2}(\cdot)}{\partial w_{h+2}} \right)}_{\neq 0 \text{, i.e., the envelope condition does not hold.}} + \frac{\rho_{h,h+2}}{\rho_{h,h+1}} \frac{\partial V^h_{h+2}(\cdot)}{\partial w_{h+2}}. \end{aligned} \tag{24}$$

Add and subtract terms to (23b) to get

$$u_c(c_{h+1}) = \rho_{h+1,h+2} \frac{\partial V^{h+1}_{h+2}(\cdot)}{\partial w_{h+2}} + \rho_{h+1,h+2} \left( \frac{\partial V^h_{h+2}(\cdot)}{\partial w_{h+2}} - \frac{\partial V^h_{h+2}(\cdot)}{\partial w_{h+2}} \right)$$

and then rewrite the above as

$$\frac{\partial V^h_{h+2}(\cdot)}{\partial w_{h+2}} = u_c(c_{h+1}) \frac{1}{\rho_{h+1,h+2}} + \Delta V^{h,h+1}_{h+2},$$

where $\Delta V^{h,h+1}_{h+2} \equiv \frac{\partial V^h_{h+2}(\cdot)}{\partial w_{h+2}} - \frac{\partial V^{h+1}_{h+2}(\cdot)}{\partial w_{h+2}}$.

Next, use (23b) in (24) to get

$$
\begin{aligned}
\frac{\partial V_{h+1}^{h}(\cdot)}{\partial w_{h+1}} &= u_c(c_{h+1})\frac{\partial c_{h+1}}{\partial w_{h+1}} + \frac{\rho_{h,h+2}}{\rho_{h,h+1}}\left(1-\frac{\partial c_{h+1}}{\partial w_{h+1}}\right)\left(u_c(c_{h+1})\frac{1}{\rho_{h+1,h+2}} + \Delta V_{h+2}^{h,h+1}\right) \\
&= u_c(c_{h+1})\left(\frac{\partial c_{h+1}}{\partial w_{h+1}} + \frac{\rho_{h,h+2}}{\rho_{h,h+1}\rho_{h+1,h+2}}\left(1-\frac{\partial c_{h+1}}{\partial w_{h+1}}\right)\right) \\
&\quad + \beta\frac{\rho_{h,h+2}}{\rho_{h,h+1}}\left(1-\frac{\partial c_{h+1}}{\partial w_{h+1}}\right)\Delta V_{h+2}^{h,h+1}.
\end{aligned}
$$

Using the above in (23a) and simplifying the resulting expression we finally get (20)–(22). □

### A.1.2 Proof of Lemma 2

*Proof.* 1. First, derive the closed form solution for the derivative of the value function. From (24) we have that for $c_t^s = m_t^s w_t$ the derivative of the value function for any period $t$, $h \le t < T$ is given by

$$
\frac{\partial V_t^h(\cdot)}{\partial w_t} = c_t^{s\,-\theta} m_t^s + \frac{\rho_{h,t+1}}{\rho_{h,t}}\left(1-m_t^s\right)\frac{\partial V_{t+1}^h(\cdot)}{\partial w_{t+1}} \tag{25}
$$

We now derive an expression for the derivative of the value function in terms of period $t$ wealth, again by backward induction.

(a) Claim: In any period $t \le T$ the derivative of the value function is

$$
\frac{\partial V_t^h(\cdot)}{\partial w_t} = \zeta_t^h w_t^{-\theta} \tag{26}
$$

for some $\zeta_t^h > 0$.

(b) Base case: In period $T$ we have $V_T^h = \frac{1}{1-\theta} c_T^{s\,1-\theta} = \frac{1}{1-\theta} w_T^{1-\theta}$ so that $\frac{\partial V_T^h(\cdot)}{\partial w_T} = w_T^{-\theta}$ and $\zeta_T^h = 1$.

(c) Induction step: Use (26) in (25) to get

$$
\frac{\partial V_t^h(\cdot)}{\partial w_t} = c_t^{s\,-\theta} m_t^s + \frac{\rho_{h,t+1}}{\rho_{h,t}}\left(1-m_t^s\right)\zeta_{t+1}^h w_{t+1}^{-\theta}.
$$

Using $c_t^s = m_t^s w_t$ and $w_{t+1} = (w_t - c_t^s) = (1 - m_t^s) w_t$ in the above we get

$$\frac{\partial V_t^h(\cdot)}{\partial w_t} = \left( {m_t^s}^{1-\theta} + \frac{\rho_{h,t+1}}{\rho_{h,t}} \left(1 - m_t^s\right)^{1-\theta} \zeta_{t+1}^h \right) w_t^{-\theta} = \zeta_t^h w_t^{-\theta}$$

so that we can generally write

$$\zeta_t^h = \begin{cases} 1 & \text{for } t = T \\ {m_t^s}^{1-\theta} + \frac{\rho_{h,t+1}}{\rho_{h,t}} \left(1 - m_t^s\right)^{1-\theta} \zeta_{t+1}^h & \text{otherwise.} \end{cases}$$

2. Next, use (26) to rewrite the distance of derivatives of value functions showing up in term $\Xi_h$ in equation (20) as

$$\Delta V_{h+2}^{h,h+1} = \frac{\partial V_{h+2}^h(\cdot)}{\partial w_{h+2}} - \frac{\partial V_{h+2}^{h+1}(\cdot)}{\partial w_{h+2}} = \left( \zeta_{h+2}^h - \zeta_{h+2}^{h+1} \right) w_{h+2}^{-\theta}.$$

Use the above to rewrite $\Xi_h$ as

$$\begin{aligned} \Xi_h &= \left(1 - m_{h+1}^s\right) \left( \frac{\partial V_{h+2}^h}{\partial w_{h+2}} - \frac{\partial V_{h+2}^{h+1}}{\partial w_{h+2}} \right) \\ &= \left(1 - m_{h+1}^s\right) \left( \zeta_{h+2}^h - \zeta_{h+2}^{h+1} \right) \left( (1 - m_{h+1}^s) w_{h+1} \right)^{-\theta} \\ &= \left(1 - m_{h+1}^s\right)^{1-\theta} \left( \zeta_{h+2}^h - \zeta_{h+2}^{h+1} \right) w_{h+1}^{-\theta} \end{aligned}$$

Dividing by $w_{h+1}^{-\theta}$ and multiplying with ${m_{h+1}^s}^{\theta}$ gives $\xi_h$ in equation (19).

3. Induction step: Use the above in the first-order condition (20) to get

$$\begin{aligned} & {c_h^s}^{-\theta} = \rho_{h,h+1} \lambda_h {c_{h+1}^s}^{-\theta} + \rho_{h,h+2} \xi_h {m_{h+1}^s}^{-\theta} w_{h+1}^{-\theta} \\ \Leftrightarrow \quad & \left(m_h^s w_h\right)^{-\theta} = \left( \rho_{h,h+1} \lambda_h + \rho_{h,h+2} \xi_h \right) {m_{h+1}^s}^{-\theta} \left( (1 - m_h^s) w_h \right)^{-\theta}. \end{aligned}$$

Rearranging the above gives (18).

□

## A.2 Derivation of Theorem 3

We derive Theorem 3 in three different stages which are building up on each other. At first, we prove Theorem 3 for the three-period model, which will give us a crucial formal argument in terms of first-order derivatives in a belief deviation parameter. In a second stage, we increase the complexity of the model by extending the analysis from the three-period model to the four period model whereby we have to take account of endogenous constraints for belief deviation parameters. Based on the structural insights for the four-period model, it is then easy to prove Theorem 3 for the general $T$-period model.

### A.2.1 Three-Period Model

Fix $T = 2$ so that we only have to compare the MPCs for period 0. By Theorem 2, we have for the naive agent

$$\begin{aligned} m_1^{n,0} &= \frac{1}{1+\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}} \\ m_0^n &= \frac{1}{1+\frac{\left(\rho_{0,1}\right)^{\frac{1}{\theta}}}{m_1^{n,0}}} = \frac{1}{1+\rho_{0,1}^{\frac{1}{\theta}}\left(1+\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\right)} \end{aligned} \tag{27}$$

whereas, by Lemma 2, we have for the sophisticated agent

$$\begin{aligned} m_1^s &= \frac{1}{1+\frac{\left(\rho_{1,2}\right)^{\frac{1}{\theta}}}{m_2^s}} = \frac{1}{1+\left(\rho_{1,2}\right)^{\frac{1}{\theta}}} \\ m_0^s &= \frac{1}{1+\frac{\left(\rho_{0,1}\lambda_1\right)^{\frac{1}{\theta}}}{m_1^s}} = \frac{1}{1+\rho_{0,1}^{\frac{1}{\theta}}\lambda^{\frac{1}{\theta}}\left(1+\left(\rho_{1,2}\right)^{\frac{1}{\theta}}\right)} \end{aligned} \tag{28}$$

where

$$\begin{aligned} \lambda = m_1^s + \frac{\rho_{0,2}}{\rho_{0,1}\rho_{1,2}}(1-m_1^s) &= \frac{1}{1+\left(\rho_{1,2}\right)^{\frac{1}{\theta}}} + \frac{\rho_{0,2}}{\rho_{0,1}\rho_{1,2}}\frac{\left(\rho_{1,2}\right)^{\frac{1}{\theta}}}{1+\left(\rho_{1,2}\right)^{\frac{1}{\theta}}} \\ = \frac{1+\frac{\rho_{0,2}}{\rho_{0,1}\rho_{1,2}}\left(\rho_{1,2}\right)^{\frac{1}{\theta}}}{1+\left(\rho_{1,2}\right)^{\frac{1}{\theta}}} &= \frac{1+\frac{\rho_{0,2}}{\rho_{1,2}}\left(\rho_{1,2}\right)^{\frac{1}{\theta}-1}}{1+\left(\rho_{1,2}\right)^{\frac{1}{\theta}}} \end{aligned} \tag{29}$$

Observe that dynamic inconsistency is here equivalent to

$$\rho_{1,2} = \frac{\rho_{0,2}}{\rho_{0,1}}(1+\epsilon)$$

for some $\epsilon \neq 0$ such that $\epsilon \in \left(-1, \frac{\rho_{0,1}}{\rho_{0,2}} - 1\right)$. Use this in (28) and (29) to get

$$m_0^s = \frac{1}{1 + \left(\rho_{0,1}\right)^{\frac{1}{\theta}} \lambda_1(\epsilon)^{\frac{1}{\theta}} \left(1 + \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}} (1+\epsilon)^{\frac{1}{\theta}}\right)} \tag{30}$$

where

$$\lambda(\epsilon) = \frac{1 + \frac{\rho_{0,2}}{\rho_{1,2}} \left(\rho_{1,2}\right)^{\frac{1}{\theta}-1}}{1 + \left(\rho_{1,2}\right)^{\frac{1}{\theta}}} = \frac{1 + \left(\frac{\rho_{0,2}}{\rho_{1,2}}\right)^{\frac{1}{\theta}} (1+\epsilon)^{\frac{1}{\theta}-1}}{1 + \left(\frac{\rho_{0,2}}{\rho_{1,2}}\right)^{\frac{1}{\theta}} (1+\epsilon)^{\frac{1}{\theta}}} \tag{31}$$

By comparing (27) and (30) we observe that the structure of the equations is identical. We can thus define

$$\begin{aligned} m_0^s &= \frac{1}{1 + \left(\rho_{0,1}\right)^{\frac{1}{\theta}} g_1(\epsilon)}, \\ g\left(\epsilon\right) &= \lambda(\epsilon)^{\frac{1}{\theta}} \left(1 + \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}} (1+\epsilon)^{\frac{1}{\theta}}\right). \end{aligned} \tag{32}$$

and we can think of $g(\epsilon)$ as a measure of the saving intensity of the sophisticated agent.

We know that $m_0^n$ and $m_0^s$ must coincide for the benchmark case of dynamic consistency. To verify this, set $\epsilon = 0$ for dynamic consistency to obtain

$$\begin{aligned} \lambda(0) &= 1, \\ g\left(0\right) &= 1 + \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}, \\ m_0^n &= m_0^s = \frac{1}{1 + \left(\rho_{0,1}\right)^{\frac{1}{\theta}} \left(1 + \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\right)}. \end{aligned}$$

In what follows, we deviate from $\epsilon = 0$ to show that (i) $g(\epsilon)$ is globally u-shaped in $\epsilon$ for $\theta > 1$ whereas it is globally inverse-u-shaped for $\theta < 1$ while (ii) $g(\epsilon)$ becomes an extreme point at $\epsilon = 0$. In other words, $g(\epsilon)$ takes on a global minimum at $\epsilon = 0$ for $\theta > 1$ whereas it takes on at $\epsilon = 0$ a global maximum for $\theta > 1$. This gives us the desired relationship

$$m_0^s \lesseqgtr m_0^n \qquad \Leftrightarrow \qquad 1 \gtreqless \theta.$$

We thus prove Theorem 3 for the three-period model by proving the following

**Observation 1.** *(i) The function (32) is equivalently given as*

$$g(\epsilon) = \left(1 + (1+\epsilon)^{\frac{1}{\theta}-1} \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\right)^{\frac{1}{\theta}} \left(1 + (1+\epsilon)^{\frac{1}{\theta}} \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\right)^{1-\frac{1}{\theta}}. \qquad (33)$$

*(ii) Taking the derivative of (33) establishes the desired u- and inverse-u shape, respectively:*

$$\begin{aligned}
\frac{\partial g(\epsilon)}{\partial \epsilon} &= 0 \text{ for } \epsilon = 0,\\
\frac{\partial g(\epsilon)}{\partial \epsilon} &> 0 \text{ for } \theta > 1, \epsilon > 0, \text{ and for } \theta < 1, \epsilon < 0,\\
\frac{\partial g(\epsilon)}{\partial \epsilon} &< 0 \text{ for } \theta > 1, \epsilon < 0, \text{ and for } \theta < 1, \epsilon > 0.
\end{aligned}$$

*Proof.* Use (31) to rewrite $g(\epsilon)$ as

$$\begin{aligned}
g(\epsilon) &= \left(\frac{1 + \left(\frac{\rho_{0,2}}{\rho_{1,2}}\right)^{\frac{1}{\theta}} (1+\epsilon)^{\frac{1}{\theta}-1}}{1 + \left(\frac{\rho_{0,2}}{\rho_{1,2}}\right)^{\frac{1}{\theta}} (1+\epsilon)^{\frac{1}{\theta}}}\right)^{\theta} \left(1 + (1+\epsilon)^{\frac{1}{\theta}} \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\right)\\
&= \left(1 + (1+\epsilon)^{\frac{1}{\theta}-1} \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\right)^{\frac{1}{\theta}} \left(1 + (1+\epsilon)^{\frac{1}{\theta}} \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\right)^{1-\frac{1}{\theta}}.
\end{aligned}$$

Taking the derivative gives

$$\begin{aligned}
\frac{\partial g(\epsilon)}{\partial \epsilon} = g(\epsilon) \Bigg( &\frac{1}{\theta}\left(\frac{1}{\theta} - 1\right) (1+\epsilon)^{\frac{1}{\theta}-2} \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}} \left(1 + (1+\epsilon)^{\frac{1}{\theta}-1} \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\right)^{-1} +\\
&\left(1 - \frac{1}{\theta}\right) \frac{1}{\theta} (1+\epsilon)^{\frac{1}{\theta}-1} \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}} \left(1 + (1+\epsilon)^{\frac{1}{\theta}} \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\right)^{-1} \Bigg)
\end{aligned}$$

so that

$$
\begin{aligned}
& \frac{\partial g(\epsilon)}{\partial \epsilon} > 0 \\
\Leftrightarrow \quad & \left(\frac{1}{\theta}-1\right)(1+\epsilon)^{-1}\left(1+(1+\epsilon)^{\frac{1}{\theta}-1}\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\right)^{-1} > \left(\frac{1}{\theta}-1\right)\left(1+(1+\epsilon)^{\frac{1}{\theta}}\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\right)^{-1}
\end{aligned}
$$

which gives rise to the following case distinction:

Suppose that $\frac{1}{\theta} - 1 > 0 \quad \Leftrightarrow \quad \theta < 1$. Then

$$
\begin{aligned}
& \frac{\partial g(\epsilon)}{\partial \epsilon} > 0 \\
\Leftrightarrow \quad & (1+\epsilon)\left(1+(1+\epsilon)^{\frac{1}{\theta}-1}\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\right) < 1+(1+\epsilon)^{\frac{1}{\theta}}\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}} \\
\Leftrightarrow \quad & 1+\epsilon+(1+\epsilon)^{\frac{1}{\theta}}\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}} < 1+(1+\epsilon)^{\frac{1}{\theta}}\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}} \\
\Leftrightarrow \quad & \epsilon < 0
\end{aligned}
$$

Consequently, $g(\epsilon)$ is inverse-u-shaped with maximum at $\epsilon = 0$.

Now suppose that $\frac{1}{\theta} - 1 < 0 \quad \Leftrightarrow \quad \theta > 1$. Then

$$
\begin{aligned}
& \frac{\partial g(\epsilon)}{\partial \epsilon} > 0 \\
\Leftrightarrow \quad & \epsilon > 0
\end{aligned}
$$

and thus $g(\epsilon)$ is u-shaped with minimum at $\epsilon = 0$. □

### A.2.2 Four-Period Model

Fix $T = 3$ so that we have to compare the MPCs for periods 0 and 1. Let us start with the analysis for period 0 whereby we consider the case for period 1 as we go along.

Recursive substitution in Theorem 2 gives us for the naive agent

$$
m_0^{n,0} = \frac{1}{1+\left(\rho_{0,1}\right)^{\frac{1}{\theta}}\left(1+\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\left(1+\left(\frac{\rho_{0,3}}{\rho_{0,2}}\right)^{\frac{1}{\theta}}\right)\right)}.
$$

Similarly, by Lemma 2, we obtain for the sophisticated agent

$$m_0^s = \frac{1}{1 + \left(\rho_{0,1}\right)^{\frac{1}{\theta}} \lambda_0^{\frac{1}{\theta}} \left(1 + \rho_{1,2}^{\frac{1}{\theta}} \lambda_1^{\frac{1}{\theta}} \left(1 + \rho_{2,3}^{\frac{1}{\theta}}\right)\right)}, \tag{34}$$

where

$$\lambda_0 = m_1^s + \frac{\rho_{0,2}}{\rho_{0,1}\rho_{1,2}}(1 - m_1^s), \qquad \lambda_1 = m_2^s + \frac{\rho_{1,3}}{\rho_{1,2}\rho_{2,3}}(1 - m_2^s).$$

Next, define

$$\rho_{1,2} = \frac{\rho_{0,2}}{\rho_{0,1}}\left(1 + \epsilon_{1,2}^0\right), \qquad \rho_{2,3} = \frac{\rho_{0,3}}{\rho_{0,2}}(1 + \epsilon_{2,3}^0),$$

for $\epsilon_{1,2}^0 \in \left(-1, \frac{\rho_{0,1}}{\rho_{0,2}} - 1\right)$, $\epsilon_{2,3}^0 \in \left(-1, \frac{\rho_{0,2}}{\rho_{0,3}} - 1\right)$, to rewrite (34) as

$$m_0^s = \frac{1}{1 + \left(\rho_{0,1}\right)^{\frac{1}{\theta}} \lambda_0(\epsilon_{1,2}^0)^{\frac{1}{\theta}} \left(1 + \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}} \left(1 + \epsilon_{1,2}^0\right)^{\frac{1}{\theta}} \lambda_1(\epsilon_{2,3}^0)^{\frac{1}{\theta}} \left(1 + \left(\frac{\rho_{0,3}}{\rho_{0,2}}\right)^{\frac{1}{\theta}} \left(1 + \epsilon_{2,3}^0\right)^{\frac{1}{\theta}}\right)\right)} \tag{35}$$

where

$$\lambda_0\left(\epsilon_{1,2}^0\right) = m_1^s + \frac{1}{1 + \epsilon_{1,2}^0}(1 - m_1^s), \qquad \lambda_1\left(\epsilon_{2,3}^0\right) = m_2^s + \frac{\rho_{1,3}}{\rho_{1,2}}\frac{\rho_{0,2}}{\rho_{0,3}}\frac{1}{1 + \epsilon_{2,3}^0}(1 - m_2^s),$$

which has the same structure as for the naive agent.

Note that the discount factors must satisfy

$$\rho_{2,3} = \frac{\rho_{0,3}}{\rho_{0,2}}(1 + \epsilon_{2,3}^0) \text{ and } \rho_{2,3} = \frac{\rho_{1,3}}{\rho_{1,2}}(1 + \epsilon_{2,3}^1)$$

for $\epsilon_{2,3}^1 \in \left(-1, \frac{\rho_{1,2}}{\rho_{1,3}} - 1\right)$ implying the (endogenous) constraint

$$\frac{(1 + \epsilon_{2,3}^0)}{(1 + \epsilon_{2,3}^1)} = \frac{\rho_{0,2}}{\rho_{0,3}}\frac{\rho_{1,3}}{\rho_{1,2}}. \tag{36}$$

By an application of (36), the MPC for period 0 (35) is equivalently given as

$$m_0^s = \frac{1}{1 + \left(\rho_{0,1}\right)^{\frac{1}{\theta}} g_0\left(\epsilon_{1,2}^0, g_1\left(\epsilon_{2,3}^1\right)\right)}$$

where

$$\begin{aligned} g_1\left(\epsilon^1_{2,3}\right) &= \lambda_1\left(\epsilon^1_{2,3}\right)^{\frac{1}{\theta}}\left(1+\left(\frac{\rho_{1,3}}{\rho_{1,2}}\right)^{\frac{1}{\theta}}\left(1+\epsilon^1_{2,3}\right)^{\frac{1}{\theta}}\right) \\ g_0\left(\epsilon^0_{1,2}, g\left(\epsilon^1_{2,3}\right)\right) &= \lambda_0\left(\epsilon^0_{1,2}\right)^{\frac{1}{\theta}}\left(1+\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\left(1+\epsilon^0_{1,2}\right)^{\frac{1}{\theta}} g\left(\epsilon^1_{2,3}\right)\right) \end{aligned}$$

and

$$\lambda_t\left(\epsilon^t_{t+1,t+2}\right) = m^s_{t+1} + \frac{1}{1+\epsilon^t_{t+1,t+2}}\left(1-m^s_{t+1}\right) \text{ for } t = 0, 1.$$

Setting $\epsilon^t_{t+1,t+2} = 0$ for $t = 0, 1$ gives us the benchmark case of dynamic consistency in which the realized MPCs of the sophisticated and the naive agent coincide. To see that $\epsilon^t_{t+1,t+2} = 0$ for $t = 0, 1$ implies $m^s_1 = m^n_1$, note that

$$\begin{aligned} \lambda_t(0) &= 1 \text{ for } t = 0, 1, \\ g_1(0) &= 1+\left(\frac{\rho_{1,3}}{\rho_{1,2}}\right)^{\frac{1}{\theta}}, \\ g_0\left(0, g_1(0)\right) &= \left(1+\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\left(1+\left(\frac{\rho_{1,3}}{\rho_{1,2}}\right)^{\frac{1}{\theta}}\right)\right). \end{aligned}$$

Moreover, the MPC for period 1 is given as

$$m^s_1 = \frac{1}{1+\left(\rho_{1,2}\right)^{\frac{1}{\theta}} g\left(\epsilon^1_{2,3}\right)}.$$

Consequently, we have $m^s_1 = m^n_1$ if and only if $\epsilon^1_{2,3} = 0$ so that $m^s_1 = m^n_1$ is implied by $m^s_0 = m^n_0$.

Let us consider deviations from this dynamic consistency benchmark case whereby we start with the MPCs for period 1. As in the three-period model, we must have that

$$m^s_1 < m^n_1$$

whenever $g_1\left(\epsilon^1_{2,3}\right)$ is u-shaped in $\epsilon^1_{2,3}$ with a minimum at $\epsilon^1_{2,3} = 0$. Conversely, we have

$$m^s_1 > m^n_1$$

if $g_2(\epsilon^1_{2,3})$ is inverse-u-shaped in $\epsilon^1_{2,3}$ with a maximum at $\epsilon^1_{2,3} = 0$. Note that $g_1\left(\epsilon^1_{2,3}\right)$

has exactly the same structure as the function (32) that we used in the analysis of the three-period model. By an identical formal argument as for Observation 1, we thus obtain the following observation which gives us the desired result for the period 1 MPCs

$$m_1^s \lesseqgtr m_1^n \quad \Leftrightarrow \quad \theta \gtreqless 1.$$

**Observation 2.** *(i) The function $g_1\left(\epsilon_{2,3}^1\right)$is equivalently given as*

$$g_1\left(\epsilon_{2,3}^1\right) = \left(1 + \left(1+\epsilon_{1,2}^0\right)^{\frac{1}{\theta}-1}\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\right)^{\frac{1}{\theta}}\left(1 + \left(1+\epsilon_{1,2}^0\right)^{\frac{1}{\theta}}\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\right)^{1-\frac{1}{\theta}}. \tag{37}$$

*(ii) Taking the derivative of (37) establishes the desired u- and inverse-u shape, respectively:*

$$\frac{\partial g_1\left(\epsilon_{2,3}^1\right)}{\partial \epsilon_{2,3}^1} = 0 \textit{ for } \epsilon_{2,3}^1 = 0,$$

$$\frac{\partial g_1\left(\epsilon_{2,3}^1\right)}{\partial \epsilon_{2,3}^1} > 0 \textit{ for } \theta > 1, \epsilon_{2,3}^1 > 0, \textit{ and for } \theta < 1, \epsilon_{2,3}^1 < 0,$$

$$\frac{\partial g_1\left(\epsilon_{2,3}^1\right)}{\partial \epsilon_{2,3}^1} < 0 \textit{ for } \theta > 1, \epsilon_{2,3}^1 < 0, \textit{ and for } \theta < 1, \epsilon_{2,3}^1 > 0,$$

Next turn to the MPCs for period 0. Suppose, at first, that $g_1\left(\epsilon_{2,3}^1\right)$ is u-shaped in $\epsilon_{2,3}^1$ with a minimum at $\epsilon_{2,3}^1 = 0$. If, in addition, $g_0\left(\epsilon_{1,2}^0, g\left(\epsilon_{2,3}^1\right)\right)$ is also u-shaped in $\epsilon_{1,2}^0$ with a minimum at $\epsilon_{1,2}^0 = 0$ for any fixed value of $g_1\left(\epsilon_{2,3}^1\right)$, we obtain

$$m_0^s < m_0^n.$$

Conversely, if $g_1\left(\epsilon_{2,3}^1\right)$ is inverse-u-shaped in $\epsilon_{2,3}^1$ with a maximum at $\epsilon_{2,3}^1 = 0$ while $g_0\left(\epsilon_{1,2}^0, g\left(\epsilon_{2,3}^1\right)\right)$ is also inverse-u-shaped in $\epsilon_{1,2}^0$ with a maximum at $\epsilon_{1,2}^0 = 0$ for any fixed value of $g_1\left(\epsilon_{2,3}^1\right)$, we have

$$m_0^s > m_0^n.$$

In investigating the shape of $g_0\left(\epsilon_{1,2}^0, g\left(\epsilon_{2,3}^1\right)\right)$ we will again utilize the analysis from the three-period model by proving the following

**Observation 3.** *(i) The function $g_0\left(\epsilon_{1,2}^0, g\left(\epsilon_{2,3}^1\right)\right)$ is equivalently given as*

$$g_0\left(\epsilon_{1,2}^0, g\left(\epsilon_{2,3}^1\right)\right) = \left(1 + \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}} \left(1+\epsilon_{1,2}^0\right)^{\frac{1}{\theta}-1} g\left(\epsilon_{2,3}^1\right)\right)^{\frac{1}{\theta}} \left(1 + \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}} \left(1+\epsilon_{1,2}^0\right)^{\frac{1}{\theta}} g\left(\epsilon_{2,3}^1\right)\right)^{1-\frac{1}{\theta}}. \tag{38}$$

*(ii) Taking the derivative of (38) establishes the desired u- and inverse-u shape, respectively:*

$$\frac{\partial g_0\left(\epsilon_{1,2}^0, g\left(\epsilon_{2,3}^1\right)\right)}{\partial \epsilon_{1,2}^0} = 0 \text{ for } \epsilon_{1,2}^0 = 0,$$

$$\frac{\partial g_0\left(\epsilon_{1,2}^0, g\left(\epsilon_{2,3}^1\right)\right)}{\partial \epsilon_{1,2}^0} > 0 \text{ for } \theta > 1, \epsilon_{1,2}^0 > 0, \text{ and for } \theta < 1, \epsilon_{1,2}^0 < 0,$$

$$\frac{\partial g_0\left(\epsilon_{1,2}^0, g\left(\epsilon_{2,3}^1\right)\right)}{\partial \epsilon_{1,2}^0} < 0 \text{ for } \theta > 1, \epsilon_{1,2}^0 < 0, \text{ and for } \theta < 1, \epsilon_{1,2}^0 > 0,$$

*Proof.* Recall that

$$\lambda_0\left(\epsilon_{1,2}^0\right) = m_1^s + \frac{1}{1+\epsilon_{1,2}^0}(1-m_1^s)$$

whereby

$$m_1^s = \frac{1}{1 + \left(\frac{\rho_{0,2}}{\rho_{0,1}}\left(1+\epsilon_{1,2}^0\right)\right)^{\frac{1}{\theta}} g\left(\epsilon_{2,3}^1\right)}$$

so that

$$\begin{aligned}
\lambda_0\left(\epsilon_{1,2}^0\right) &= \frac{1}{1 + \left(\frac{\rho_{0,2}}{\rho_{0,1}}\left(1+\epsilon_{1,2}^0\right)\right)^{\frac{1}{\theta}} g\left(\epsilon_{2,3}^1\right)} + \frac{1}{1+\epsilon_{1,2}^0} \frac{\left(\frac{\rho_{0,2}}{\rho_{0,1}}\left(1+\epsilon_{1,2}^0\right)\right)^{\frac{1}{\theta}} g\left(\epsilon_{2,3}^1\right)}{1 + \left(\frac{\rho_{0,2}}{\rho_{0,1}}\left(1+\epsilon_{1,2}^0\right)\right)^{\frac{1}{\theta}} g\left(\epsilon_{2,3}^1\right)} \\
&= \frac{1 + \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}} \left(1+\epsilon_{1,2}^0\right)^{\frac{1}{\theta}-1} g\left(\epsilon_{2,3}^1\right)}{1 + \left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}} \left(1+\epsilon_{1,2}^0\right)^{\frac{1}{\theta}} g\left(\epsilon_{2,3}^1\right)}.
\end{aligned}$$

Substitution results in

$$\begin{aligned} g_0\left(\epsilon_{1,2}^0, g\left(\epsilon_{2,3}^1\right)\right) &= \lambda_0\left(\epsilon_{1,2}^0\right)^{\frac{1}{\theta}}\left(1+\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\left(1+\epsilon_{1,2}^0\right)^{\frac{1}{\theta}} g\left(\epsilon_{2,3}^1\right)\right) \\ &= \left(\frac{1+\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\left(1+\epsilon_{1,2}^0\right)^{\frac{1}{\theta}-1} g\left(\epsilon_{2,3}^1\right)}{1+\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\left(1+\epsilon_{1,2}^0\right)^{\frac{1}{\theta}} g\left(\epsilon_{2,3}^1\right)}\right)^{\frac{1}{\theta}}\left(1+\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\left(1+\epsilon_{1,2}^0\right)^{\frac{1}{\theta}} g\left(\epsilon_{2,3}^1\right)\right) \\ &= \left(1+\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\left(1+\epsilon_{1,2}^0\right)^{\frac{1}{\theta}-1} g\left(\epsilon_{2,3}^1\right)\right)^{\frac{1}{\theta}}\left(1+\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\left(1+\epsilon_{1,2}^0\right)^{\frac{1}{\theta}} g\left(\epsilon_{2,3}^1\right)\right)^{1-\frac{1}{\theta}}. \end{aligned}$$

Observe that $g_0\left(\epsilon_{1,2}^0, g\left(\epsilon_{2,3}^1\right)\right)$ has the same structure as the function (33) in the three-period model, with term $\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}} g\left(\epsilon_{2,3}^1\right)$ instead of $\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}$. □

Finally, by combining Observations 2 and 3, we obtain the desired result for the MPCs for period 1

$$m_0^s \lesseqgtr m_0^n \quad \Leftrightarrow \quad \theta \gtreqless 1.$$

### A.2.3 Proof for Multi-Period Model

*Proof.* Fix an arbitrary $T > 3$. We now have to compare the MPCs for all periods $0, .., T-2$. To focus thoughts, let us consider period 0 whereby the analysis can be identically applied to the other periods (we could actually write $h$ for 0). Our analysis closely mirrors the analysis for the four-period model.

Upon recursive substitution of the naives consumption policy function we get

$$m_0^{n,0} = \frac{1}{1+\left(\rho_{0,1}\right)^{\frac{1}{\theta}}\left(1+\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\left(1+\left(\frac{\rho_{0,3}}{\rho_{0,2}}\right)^{\frac{1}{\theta}}\left(1+\ldots\left(\frac{\rho_{0,T}}{\rho_{0,T-1}}\right)^{\frac{1}{\theta}}\right)\right)\right)}.$$

Likewise, we get for the sophisticated agent

$$m_0^s = \frac{1}{1+\left(\rho_{0,1}\right)^{\frac{1}{\theta}}\lambda_0(\epsilon_{1,2}^0)^{\frac{1}{\theta}}\left(1+\left(\frac{\rho_{0,2}}{\rho_{0,1}}\right)^{\frac{1}{\theta}}\left(1+\epsilon_{1,2}^0\right)^{\frac{1}{\theta}}\lambda_1(\epsilon_{2,3}^0)^{\frac{1}{\theta}}\left(1+\ldots\left(\frac{\rho_{0,T}}{\rho_{0,T-1}}\right)^{\frac{1}{\theta}}\left(1+\epsilon_{T-1,T}^0\right)^{\frac{1}{\theta}}\right)\right)}. \tag{39}$$

Use now the endogenous constraints on the beliefs

$$\rho_{t+1,t+2} = \frac{\rho_{0,t+2}}{\rho_{0,t+1}}(1+\epsilon^0_{t+1,t+2}) \text{ and } \rho_{t+1,t+2} = \frac{\rho_{t,t+2}}{\rho_{t,t+1}}(1+\epsilon^t_{t+1,t+2})$$

to rewrite (39) as

$$m_0^s = \frac{1}{1+\left(\rho_{0,1}\right)^{\frac{1}{\theta}} g_0\left(\epsilon^0_{1,2}, g_1\left(\epsilon^1_{2,3}, ..., g_{T-3}\left(\epsilon^{T-3}_{T-2,T-1}, g_{T-2}\left(\epsilon^{T-2}_{T-1,T}\right)\right)\right)\right)}$$

such that

$$g_{T-2}\left(\epsilon^{T-2}_{T-1,T}\right) = \lambda_{T-2}\left(\epsilon^{T-2}_{T-1,T}\right)^{\frac{1}{\theta}}\left(1+\left(\frac{\rho_{T-2,T}}{\rho_{T-2,T-1}}\right)^{\frac{1}{\theta}}\left(1+\epsilon^{T-2}_{T-1,T}\right)^{\frac{1}{\theta}}\right)$$

and, for all $t < T-2$,

$$g_t\left(\epsilon^t_{t+1,t+2}, g\left(\epsilon^{t+1}_{t+2,t+3}, ...\right)\right) = \lambda_t\left(\epsilon^t_{t+1,t+2}\right)^{\frac{1}{\theta}}\left(1+\left(\frac{\rho_{t,t+2}}{\rho_{t,t+1}}\right)^{\frac{1}{\theta}}\left(1+\epsilon^t_{t+1,t+2}\right)^{\frac{1}{\theta}} g\left(\epsilon^{t+1}_{t+2,t+3}, ...\right)\right)$$

where

$$\lambda_t\left(\epsilon^t_{t+1,t+2}\right) = m^s_{t+1} + \frac{1}{1+\epsilon^t_{t+1,t+2}}(1-m^s_{t+1}) \text{ for } t = 0, ..., T-2.$$

We have dynamic consistency—implying $m_0^{n,0} = m_0^s$—if and only if $\epsilon^t_{t+1,t+2} = 0$ for $t = 0, ..., T-2$. As in our previous analysis, we can establish the desired result

$$m_0^s \lesseqqgtr m_0^n \quad \Leftrightarrow \quad \theta \gtreqqless 1$$

by showing that all the functions $g_{T-2}\left(\epsilon^{T-2}_{T-1,T}\right)$ and $g_t\left(\epsilon^t_{t+1,t+2}, g\left(\epsilon^{t+1}_{t+2,t+3}, ...\right)\right)$ for $t = 0, ..., T-2$, have the desired u- or inverse-u-shape, respectively, with an extreme point at zero.

The according u-shape analysis follows now easily from the corresponding analysis of the four-period model. To see this, note that $g_{T-2}\left(\epsilon^{T-2}_{T-1,T}\right)$ has the same structural form as the function (37) so that we can employ Observation 2. To be precise, we obtain by an application of Observation 2 that

$$\frac{\partial g_{T-2}\left(\epsilon_{T-1,T}^{T-2}\right)}{\partial \epsilon_{T-1,T}^{T-2}} = 0 \text{ for } \epsilon_{T-1,T}^{T-2} = 0,$$
$$\frac{\partial g_{T-2}\left(\epsilon_{T-1,T}^{T-2}\right)}{\partial \epsilon_{T-1,T}^{T-2}} > 0 \text{ for } \theta > 1, \epsilon_{T-1,T}^{T-2} > 0, \text{ and for } \theta < 1, \epsilon_{T-1,T}^{T-2} < 0,$$
$$\frac{\partial g_{T-2}\left(\epsilon_{T-1,T}^{T-2}\right)}{\partial \epsilon_{T-1,T}^{T-2}} < 0 \text{ for } \theta > 1, \epsilon_{T-1,T}^{T-2} < 0, \text{ and for } \theta < 1, \epsilon_{T-1,T}^{T-2} > 0.$$

Similarly, any function $g_t\left(\epsilon_{t+1,t+2}^{t}, g\left(\epsilon_{t+2,t+3}^{t+1}, ...\right)\right)$ with $t = 0, ..., T-2$ has the structural form of the function (38) so that the desired result follows from Observation 3. That is, for all $t = 0, ..., T-2$,

$$\frac{\partial g_t\left(\epsilon_{t+1,t+2}^{t}, g\left(\epsilon_{t+2,t+3}^{t+1}, ...\right)\right)}{\partial \epsilon_{t+1,t+2}^{t}} = 0 \text{ for } \epsilon_{t+1,t+2}^{t} = 0,$$
$$\frac{\partial g_t\left(\epsilon_{t+1,t+2}^{t}, g\left(\epsilon_{t+2,t+3}^{t+1}, ...\right)\right)}{\partial \epsilon_{t+1,t+2}^{t}} > 0 \text{ for } \theta > 1, \epsilon_{t+1,t+2}^{t} > 0, \text{ and for } \theta < 1, \epsilon_{t+1,t+2}^{t} < 0,$$
$$\frac{\partial g_t\left(\epsilon_{t+1,t+2}^{t}, g\left(\epsilon_{t+2,t+3}^{t+1}, ...\right)\right)}{\partial \epsilon_{t+1,t+2}^{t}} < 0 \text{ for } \theta > 1, \epsilon_{t+1,t+2}^{t} < 0, \text{ and for } \theta < 1, \epsilon_{t+1,t+2}^{t} > 0.$$

This finally proves

$$m_0^s \lesseqqgtr m_0^n \quad \Leftrightarrow \quad \theta \gtreqqless 1$$

which gives us Theorem 3. □

# B Supplementary Online Appendix (Not for Publication)

## B.1 Monte Carlo Simulation Variant I

In the simulations of the first variant of Monte Carlo simulations we construct a system of i.i.d. effective discount factors as follows:

1. Set up an equidistant grid $\mathcal{G}^{\rho} = [\varepsilon, ..., 1]$, where $\varepsilon > 0$ is a small number.
2. Iterate forward from $h = 0$ to $h = T - 1$: For each age $h$, draw randomly from $\mathcal{G}^{\rho}$ the sequence of i.i.d. discount factors $\rho_{h,t}$, for $t = h + 1, \ldots, T$.

The results of the ratio of MPCs of the naive to the sophisticated decision maker at age $h = 0$ for a subset of $1,000$ Monte Carlo simulations is depicted in Figure 3.

## B.2 Monte Carlo Simulation Variant II

To construct the system of survival beliefs, we proceed as follows:

1. Set up an equidistant grid $\mathcal{G}^{\nu_1} = (\varepsilon, \nu(D_{t-1} \cup \cdots \cup D_T))$, where $\varepsilon > 0$ is a small number, and draw a prior $\nu(D_t \cup \cdots \cup D_T)$ for all $t = 0, ..., T$.[15]
2. Set up an equidistant grid $\mathcal{G}^{\nu_2} = (\nu(D_0 \cup \cdots \cup D_{h-1}), 1 - \varepsilon)$, $\nu(D_0 \cup \cdots \cup D_{h-1})$ and draw a prior $\nu(D_0 \cup D_1 \cup \cdots \cup D_h)$ for all $h = 0, ..., T$.[16]
3. For all $h$, draw values for all $t > h - 1$

$$\nu(D_0 \cup D_1 \cup \cdots \cup D_{h-1} \cup D_t \cup \cdots \cup D_T)$$

from equidistant grids

$$\mathcal{G}^{\nu_3} = (x(t) + \varepsilon, \nu(D_0 \cup D_1 \cup \cdots \cup D_{h-1} \cup D_{t-1} \cup \cdots \cup D_T))$$

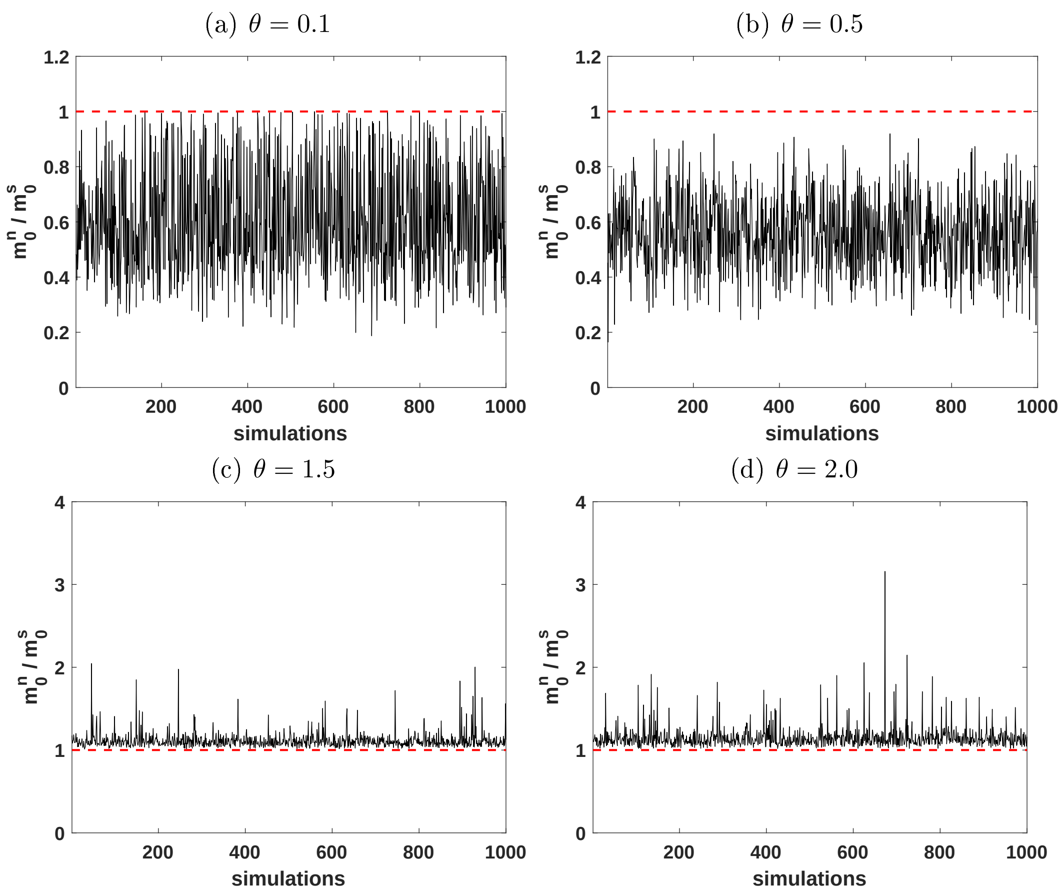
such that

$$x(t) = \max\{\nu(D_0 \cup D_1 \cup \cdots \cup D_{h-1}), \nu(D_0 \cup D_1 \cup \cdots \cup D_{h-2} \cup D_t \cup \cdots \cup D_T)\}$$

[15]Note that $\nu(D_t \cup \cdots \cup D_T) = 1$ in $t = 0$.

[16]Note that $\nu(D_0 \cup D_1 \cup \cdots \cup D_{h-1}) = 0$ at $h = 0$.

Figure 3: Ratio of Marginal Propensities: i.i.d. Effective Discount Factors




Notes: Ratio of MPCs of the naive and the sophisticated agent $m^n_{h,c}/m^s_{h,c}$ in 1,000 Monte Carlo simulations for $\beta = 1$ and $T = 10$.

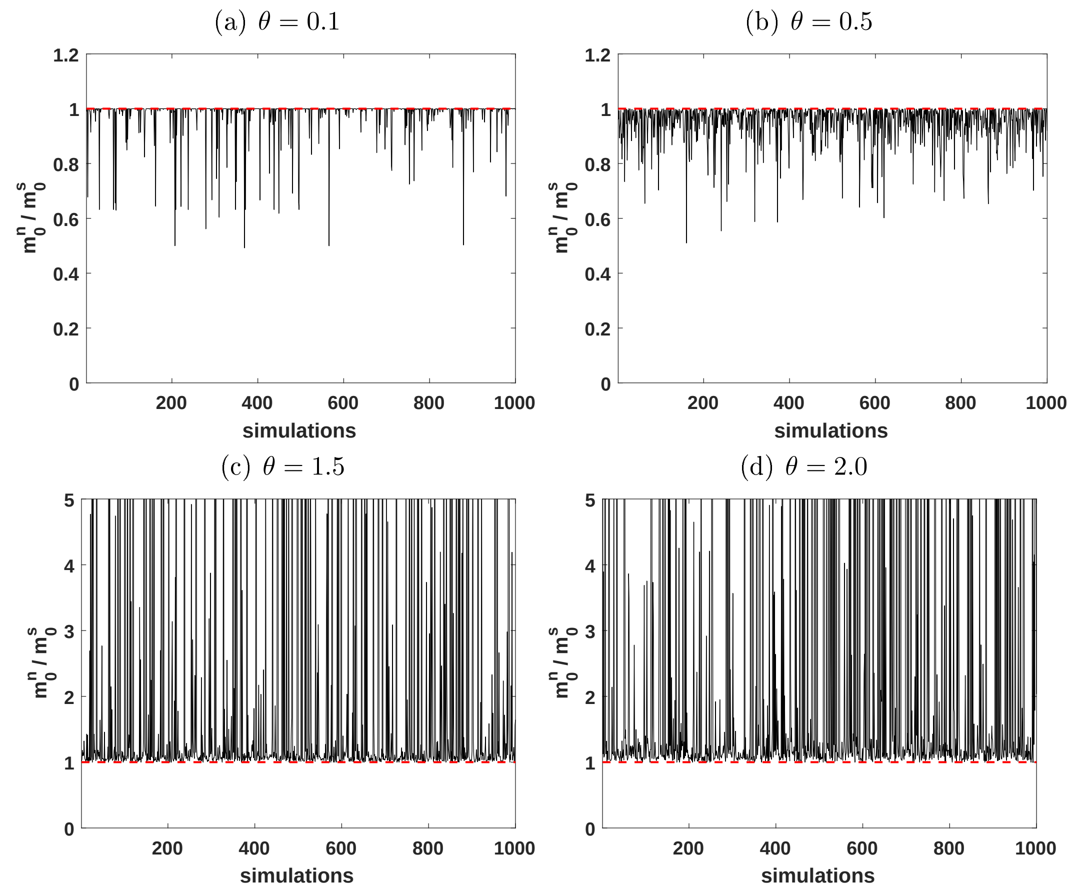
where $\nu(D_0 \cup D_1 \cup \cdots \cup D_{h-2} \cup D_t \cup \cdots \cup D_T)$ has already been drawn at the previous stage for $h-1$.[17]

4. With these constructed priors apply the updating rules (3), and (7), respectively, to construct two matrices of survival beliefs, $\nu^p_{h,t}$ for the pessimistic and $\nu^g_{h,t}$ for the generalized Bayesian update rule.

The resulting ratio of MPCs of the naive relative to the sophisticated decision maker at age $h = 0$ for a subset of $1,000$ Monte Carlo simulations is depicted in Figure 4 for the pessimistic update and in Figure 5 for the generalized Bayesian update rule.

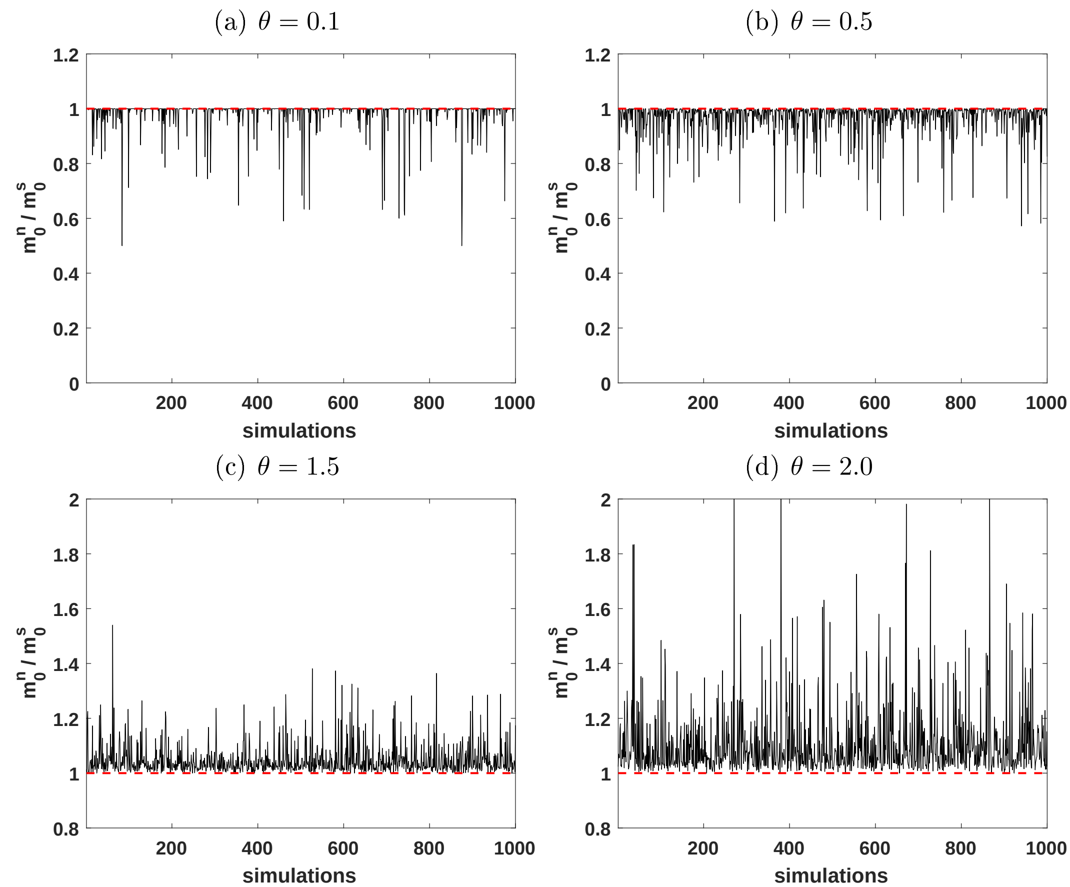
[17] Note, that $\nu(D_0 \cup \cdots \cup D_{h-1} \cup D_t \cup \cdots \cup D_T) = \nu(D_t \cup \cdots \cup D_T)$ in $h = 0$.

Figure 4: Ratio of Marginal Propensities: Pessimistic Update Rule




Notes: Ratio of MPCs of the naive and the sophisticated agent for the pessimistic update rule of survival beliefs, $m_{h,c}^{n,p}/m_{h,c}^{s,p}$, in 1,000 Monte Carlo simulations for $\beta = 1$ and $T = 10$.

Figure 5: Ratio of Marginal Propensities: Generalized Bayesian Update Rule




Notes: Ratio of MPCs of the naive and the sophisticated agent for the generalized Bayesian update rule of survival beliefs, $m_{h,c}^{n,g}/m_{h,c}^{s,g}$, in 1,000 Monte Carlo simulations for $\beta = 1$ and $T = 10$.

# References


Ando, A. and F. Modigliani (1963). The ‘Life-Cycle‘ Hypothesis of Saving: Aggregate Implications and Tests. American Economic Review 53(1), 55–84.

Andreoni, J. and C. Sprenger (2012). Risk Preferences Are Not Time Preferences. American Economic Review 102(7), 3357–3376.

Anscombe, F. J. and R. J. Aumann (1963). A Definition of Subjective Probability. Annals of American Statistics 34(1), 199-205.

Barsky, R.B., F.T. Juster, M.S. Kimball, and M.D. Shapiro (1997). Preference Parameters and Behavioral Heterogeneity: An Experimental Approach in the Health and Retirement Study. Quarterly Journal of Economics, 112, 537–79.

Billot, A., Tallon, J.-M. , and S. Mukerji (2019). Market Allocations under Ambiguity: A Survey. Working Paper.

Bleichrodt, H. and L. Eeckhoudt (2006). Survival Risks, Intertemporal Consumption, and Insurance: The Case of Distorted Probabilities. Insurance: Mathematics and Economics 38(2), 335–346.

Chateauneuf, A., J. Eichberger, and S. Grant (2007). Choice under Uncertainty with the Best and Worst in Mind: Neo-Additive Capacities. Journal of Economic Theory 137(1), 538–567.

Collard., F., Mukerji, S., Sheppard, K., and J.-M. Tallon (2018). Ambiguity and the Historical Equity Premium. Quantitative Economics 9, 945–993.

Deaton, A. (1992). Understanding Consumption. Oxford: Clarendon Press.

Drouhin, N. (2015). A Rank-dependent Utility Model of Uncertain Lifetime. Journal of Economic Dynamics and Control 53, 208–224.

Eichberger, J., S. Grant, and D. Kelsey (2007). Updating Choquet Beliefs. Journal of Mathematical Economics 43(7), 888–899.

Eichberger, J., S. Grant, and D. Kelsey (2012). When Is Ambiguity–Attitude Constant? Journal of Risk and Uncertainty 45(3), 239–263.

Ellsberg, D. (1961). Risk, Ambiguity, and the Savage Axioms. Quarterly Journal of Economics 75, 643–669.

Epper, T., H. Fehr-Duda, and A. Bruhin (2011). Viewing the Future through a Warped Lens: Why Uncertainty Generates Hyperbolic Discounting. Journal of Risk and Uncertainty 43, 169–203.

Ghirardato, P. (2002). Revisiting Savage in a Conditional World. Economic Theory 20(1), 83–92.

Gilboa, I. (1987). Expected Utility with Purely Subjective Non-Additive Probabilities. Journal of Mathematical Economics 16(1), 65–88.

Gilboa, I. and D. Schmeidler (1989). Maxmin expected utility with non-unique priors. Journal of Mathematical Economics 18(2), 141-153.

Gilboa, I. and D. Schmeidler (1993). Updating Ambiguous Beliefs. Journal of Economic Theory 59(1), 33–49.

Groneck, M., A. Ludwig, and A. Zimper (2016). A Life-Cycle Model with Ambiguous Survival Beliefs. Journal of Economic Theory 162, 137–180.

Hall, R. (1988). Intertemporal Substitution in Consumption. Journal of Political Economy 96, 339–57.

Hansen, L.P. and T. Sargent (2011). Wanting Robustness in Macroeconomics. in: B.M. Friedman and M. Woodford (eds.): Handbook of Monetary Economics, 3B, Ch. 20, pp. 1097–1157.

Hansen, L.P. and T. Sargent (2016). Robustness. Princeton University Press.

Harris, C. and D. Laibson (2001). Dynamic Choices of Hyperbolic Consumers. Econometrica 69(4), 935–957.

Kimball, M. and P. Weil (2009). Precautionary Saving and Consumption Smoothing. Journal of Money, Credit and Banking 41(2-3), 245–284.

Klibanoff, P., M. Marinacci, and S. Mukerji (2005). A Smooth Model of Decision Making under Ambiguity. Econometrica, 73(6), 1849–1892.

Klibanoff, P., M. Marinacci, and S. Mukerji (2009). Recursive Smooth Ambiguity Preferences. Journal of Economic Theory, 144(3), 930–976.

Laibson, D. I. (1997). Golden Eggs and Hyperbolic Discounting. Quarterly Journal of Economics 112(2), 443–477.

Lapied, A. and P. Toquebeuf (2013). A Note on "Re-examining the Law of Iterated Expectations for Choquet Decision Makers". Theory and Deccison 74, 439–445.

Ludwig, A. and A. Zimper (2013). A Parsimonious Model of Subjective Life Expectancy. Theory and Decision 75, 519–542.

Modigliani, F. and R. Brumberg (1954). Utility Analysis and the Consumption Function: An Interpretation of Cross-Section Data. In K. K. Kurihara (ed), Post-Keynesian Economics. New Brunswick, NJ.: Rutgers University Press.

O'Donoghue, T. and M. Rabin (1999). Doing It Now or Later. American Economic Review 89(1), 103–124.

Peracchi, F. and V. Perotti (2010). Subjective Survival Probabilities and Life Tables: Evidence from Europe. Working Paper.

Phelps, E.S. and R. Pollak (1968). On Second-Best National Saving and Game-Equilibrium Growth. Review of Economic Studies, 35, 185–199.

Samuelson, P. A. (1969). Lifetime Portfolio Selection by Dynamic Stochastic Programming. Review of Economics and Statistics 51(3), 239–246.

Sarin, R. and P.P. Wakker (1998). Dynamic Choice and Nonexpected Utility. Journal of Risk and Uncertainty 17(2) 87–119.

Savage, L. J. (1954). The Foundations of Statistics. New York, London, Sydney: John Wiley and Sons, Inc.

Schmeidler, D. (1986). Integral Representation Without Additivity. Proceedings of the American Mathematical Society 97, 255–261.

Schmeidler, D. (1989). Subjective Probability and Expected Utility Without Additivity. Econometrica 57(3), 571–587.

Strotz, R.H. (1955). Myopia and Inconsistency in Dynamic Utility Maximization. Review of Economic Studies 23(3), 165–180.

Tversky, A. and D. Kahneman (1992). Advances in Prospect Theory: Cumulative Representations of Uncertainty. Journal of Risk and Uncertainty 5(4) 297-323.

Wakker, P.P. (2010). Prospect Theory: For Risk and Ambiguity. Cambridge, UK: Cambridge University Press.

Wakker, P.P. and A. Tversky (1993). An Axiomatization of Cumulative Prospect Theory. Journal of Risk and Uncertainty, 7(2) 147-176.

Zimper, A. (2011a). Re-examining the Law of Iterated Expectations for Choquet Decision Makers. Theory and Decision, 71, 669-677.

Zimper, A. (2011b). Do Bayesians Learn Their Way out of Ambiguity? Decision Analysis, 8, 269-285.