Investor Beliefs and Asset Prices Under Selective Memory

Maximilian Voigt*

December 8, 2023

Click here for the latest version.

— Job Market Paper —

Abstract

I present a consumption-based asset pricing model in which the representative agent selectively recalls past fundamentals that resemble current fundamentals and updates beliefs as if the recalled observations are all that occurred. This similarity-weighted selective memory jointly explains important facts about belief formation, survey data, and realized asset prices. Subjective expectations overreact and are procyclical, the subjective volatility is countercyclical, and the subjective risk premium has a low volatility. In contrast, realized returns are predictably countercyclical, highly volatile, and unrelated to variation of objective risk measures. My results suggest that human memory can simultaneously account for individual-level data and aggregate asset pricing facts.

*I especially thank Sebastian Ebert and Francesco Sangiorgi for their invaluable guidance and support. I also thank Nicholas Barberis, Giovanni Burro, Adrian Buss, Paula Cocoma, Chukwuma Dim, Paul Fontanier, Katrin Goedker, Andreas Grunewald, Jonas Happel, Bjoern Hoeppner, Jonas Jensen, Yigitcan Karabulut, Paul Karehnke, Heiko Karle, Emanuel Moench, Cameron Peng, Francesco Sannino, Paul Schmidt-Engelbertz, Frederik Schwerter, Kaushik Vasudevan, Kaspar Zimmermann, as well as seminar participants at the Frankfurt-Bonn-Mannheim PhD conference, the Finance Theory Group, Frankfurt School, and Yale for their helpful suggestions and feedback. I also thank Timothy Johnson and Lars Lochstoer for sharing code with me. Part of this research was conducted while I was at Heidelberg University and Yale, and I am grateful for the hospitality of both institutions. Contact: Frankfurt School of Finance & Management, Adickesallee 32-34, 60322 Frankfurt am Main, Germany. Mail: m.voigt@fs.de.

1 Introduction

Understanding belief formation is key to understanding asset prices. In any equilibrium, the price of an asset reflects investor beliefs about future dividends and prices. The predominant rational expectations approach assumes that investors understand the temporal fluctuation of dividends and prices, but survey evidence differs markedly and systematically from rational expectations (Adam and Nagel, 2023). Evidence from the economics (Zimmermann, 2020; Charles, 2022; Gödker et al., 2022; Enke et al., 2023; Jiang et al., 2023) and psychology literature (Schacter, 2008; Kahana, 2012) highlights the role of selective memory for belief formation and decision making. Although recent models incorporate memory biases to account for individual-level belief and choice puzzles (Mullainathan, 2002; Bordalo et al., 2020b, 2023a; Wachter and Kahana, 2023), the effect of selective memory on aggregate asset prices is largely unexplored (Malmendier and Wachter, 2022).

In this paper, I show that selective memory simultaneously explains important facts about belief formation, survey data, and asset prices. Consistent with evidence, I model the beliefs of a representative agent who is more likely to recall some observations than others and treats the recalled experiences as if they were all that ever occurred (naïvete).[1] Selective memory can generate a persistent wedge between the agent's subjective beliefs and rational expectations. I analyze the effect of this belief wedge on asset prices using a consumption-based asset pricing model in which the agent learns the parameters of the payoff process.

I focus on the implications of *similarity-weighted memory*—the selective recall of past observations that resemble today's observation—for beliefs and asset prices. Recent evidence finds that similarity-weighted memory is a key mechanism of individual belief formation (Kahana, 2012; Bordalo et al., 2020b; Enke et al., 2023; Jiang et al., 2023), and Kahana

[1]Theories of memory indicate that humans are more likely to recall some observations than others (selective recall, see Schacter, 2008; Kahana, 2012), and neuronal evidence highlights that the recall of a given observation is probabilistic (stochastic recall, see Shadlen and Shohamy, 2016). Experiments in economics find that humans are (partially) unaware of their memory distortions (Zimmermann, 2020; Gödker et al., 2022; Enke et al., 2023). Although my assumptions are consistent with this evidence, I do not attempt to model the encoding and retrieval process on a neural level but focus on the effect of memory on beliefs, as is relevant for macro-finance.

et al. (2022) list similarity as a law of human memory. I show that subjective beliefs under similarity-weighted memory are consistent with empirical findings. The agent expects high growth in good times (Nagel and Xu, 2022; Bordalo et al., 2023b), expectations overreact to news (Coibion and Gorodnichenko, 2015; Bordalo et al., 2020a), and the subjective volatility of fundamentals is lower in good times than in bad times (Lochstoer and Muir, 2022).

Incorporating the agent's beliefs about fundamentals under similarity-weighted memory into the asset pricing model explains empirically observed differences of subjectively expected and objectively realized returns. Empirically, as well as in my model, subjectively expected returns are procyclical (Amromin and Sharpe, 2014; Greenwood and Shleifer, 2014), not predictable by aggregate valuation ratios, and positively related to the subjectively expected volatility (Nagel and Xu, 2023). In contrast, realized returns are countercyclical (Shiller, 1981; Mehra and Prescott, 1985), predictable by aggregate valuation ratios (Campbell and Shiller, 1988), and do not vary with objective risk measures (Lettau and Ludvigson, 2010). Quantitatively, the model generates a realistically high objective risk premium and a low risk-free rate.

Beliefs under selective memory. I first show that selective memory systematically affects the beliefs of the agent—even in the very long term—and can explain deviations from rational expectations while retaining Bayesian learning. Throughout the analysis, I focus on learning from an infinite sample to identify the systematic effect of selective memory on beliefs and asset prices, and relax this assumption in simulations. Methodologically, I characterize the beliefs of the agent under selective memory as memory-weighted likelihood maximizers as in Fudenberg et al. (2023). The agent observes many draws from the fixed distribution of fundamentals. Without memory distortions, the agent recalls all observations and the histogram of recalled observations converges almost surely and uniformly to the true distribution. With selective memory distortions, the histogram of recalled observations reflects a memory-weighted version of the true distribution. Bayesian learning then implies that the agent's beliefs concentrate on distributions that maximize the likelihood of the recalled observations. For normal distributions, I show that the agent's posterior mean is higher (lower) than the true mean if the agent is more (less) likely to recall high than low observations; while the agent's posterior volatility is higher (lower) than the true volatility if the agent is more (less) likely to recall extreme observations. These results are general and hold without a structural assumption on selective memory.

Similarity-weighted memory. I use the characterization of the agent's beliefs to incorporate similarity-weighted memory into an asset pricing model. I consider a representative agent endowment economy with Epstein and Zin (1989)-preferences. Endowment growth is drawn from an i.i.d. two-state Markov chain with observable states as in Mehra and Prescott (1985), whereby one state captures normal times and the other state recessions. Conditional on the state, endowment growth is log-normally distributed. The agent learns the state-wise mean of log endowment growth from her recalled observations, which are distorted by similarity-weighted memory. Assets are claims on the aggregate endowment (Lucas, 1978).

Similarity-weighted memory explains empirically relevant patterns of beliefs: (i) the posterior mean varies procyclically and overreacts to new information; that is, an upward revision of the agent's posterior mean predicts a negative forecast error because the posterior mean is systematically too high after an upward revision (Coibion and Gorodnichenko, 2015); and (ii) the agent's subjective volatility of fundamentals varies countercyclically.

The intuition for procyclicality and overreaction of the posterior mean is as follows: If today's endowment growth is high (low), the agent overremembers past high (low) endowment growth rates. The agent's posterior mean is then high (low) after observing high (low) endowment growth today (procyclicality). Overreaction of the agent's posterior mean occurs for a similar reason: The agent revises her posterior mean up if and only if today's endowment growth exceeds yesterday's endowment growth. Conditional on an upward revision of the agent's expectation, today's endowment growth is more likely to be above than below the fundamental mean. Moreover, the agent's posterior mean exceeds the fundamental mean if and only if today's endowment growth exceeds the fundamental mean. Consequently, the agent's posterior mean is more likely above than below the fundamental mean after an upward revision, implying a predictably negative forecast error.

The intuition for the countercyclical variation of subjective volatility is more subtle. The economy has two states. With two states, the subjective volatility of log endowment growth depends on the perceived difference of the mean log endowment growth in each state, which is time-varying. Under similarity-weighted memory, today's endowment growth has a stronger effect on the recalled growth rates when the growth rates are more spread out. During recessions, endowment growth is more spread out than during normal times, such that today's log endowment growth affects the location of the recalled endowment growth from recessions more than from normal times. Put differently, the agent is oblivious of recessions during good times, but recalls them vividly during bad times. If today's log endowment growth is high (low), the difference of the posterior means is small (large), implying that the agent perceives the economy as less (more) volatile.

In equilibrium, the agent's subjective beliefs about fundamentals affect subjectively expected as well as objectively realized returns, as all assets are claims to the aggregate endowment. Consistent with survey evidence by Greenwood and Shleifer (2014), return expectations are procyclical. When today's log endowment growth is high, the agent becomes optimistic and expects high log endowment growth going forward. The expected return (and the risk-free rate) must then increase to induce investment in the risky asset. The subjective risk premium—the difference between the subjectively expected return and the real risk-free rate—depends on the agent's risk-aversion and on the perceived riskiness of the economy. The agent's risk-aversion is constant, but the perceived riskiness of the economy (subjective volatility) is time-varying, which leads to time-variation in the subjective risk premium. However, the variation in subjective volatility is small under similarity-weighted memory, such that the subjective risk premium is acyclical, not predictable by aggregate valuation ratios, and positively related to the agent's perception of risk, consistent with evidence in Nagel and Xu (2023).

Next, I examine objectively realized returns. First, the real risk-free rate varies procyclically. Intuitively, the risk-free rate must be high if the agent expects high endowment growth to induce savings in the risk-free asset. Second, the objective risk premium is predictably countercyclical. In equilibrium, objectively realized returns depend on changes in the agents' beliefs and are predictable if belief changes are predictable. An outside observer with access to the same data as the agent can recover the parameters of the endowment growth process and predict mean reversion of the agent's beliefs. If today's endowment growth is high, the agent becomes too optimistic about the fundamentals and pushes up the price of the risky asset too much. Next period's realization then disappoints on average, beliefs mean revert, and objectively realized returns are low after a high endowment growth (Bordalo et al., 2023b). Moreover, the agent updates beliefs as if the recalled experiences were all that ever occurred and perceives the new beliefs to be persistent, which leads to volatile objective returns. Objective returns are unrelated to changes in objective risk or risk-aversion as both are constant.

I then calibrate the model to analyze the quantitative implications of similarity-weighted memory. My simulations confirm the qualitative properties of beliefs and asset prices discussed above. In addition, the objective risk premium is realistically high if the agent learns from a realistic number of past observations (30-100 years of data),[2] and can even become negative if the agent is very optimistic (Greenwood and Hanson, 2013; Cassella and Gulen, 2018). The real risk-free rate is low and does not vary much, as is consistent with data.

Peak-end memory. I briefly show that my framework can be used to analyze further selective memory biases. I consider a *peak-end memory* distortion that captures the higher memorability of extreme observations as well as observations that are similar to today's realization (Kahneman, 2000). The experience effects literature highlights the persistent

[2]With finitely many observations, the agent is uncertain about her posterior beliefs, and this uncertainty depends on the (stochastic) number of recalled observations. As in Collin-Dufresne et al. (2016), parameter uncertainty generates an additional source of risk which affects the subjective risk premium. In my simulations, the subjective risk premium is only slightly higher with than without parameter uncertainty. Instead, the higher volatility of beliefs when learning from a limited sample leads high realized returns.

influence of extreme experiences on risk taking (Malmendier and Nagel, 2011), inflation expectations (Malmendier and Nagel, 2016), managerial decisions (Malmendier et al., 2011), and real estate purchases (Happel et al., 2022). Relatedly, Kensinger and Ford (2020) argue that emotional events are more likely to be stored in memory and are often retrieved more vividly (flashbulb memories; see Phelps, 2006). Additionally, the end of an experience is typically more memorable than the beginning or middle (recency; see Kahana, 2012; Barberis, 2018; Wachter and Kahana, 2023).

In addition to the implications of similarity-weighted memory, the peak-end memory distortion leads to a high subjective volatility because the agent overremembers extreme observations. Being risk-averse, the agent thus requires a comparably high subjective risk premium, which is in line with the empirical findings reviewed in Adam and Nagel (2023).

Related literature. This paper contributes to a growing literature that examines the importance of memory on belief formation and decision making. Empirical and experimental work shows that investors' information sets are systematically affected by selective memory (Zimmermann, 2020; Charles, 2021, 2022; Gödker et al., 2022; Goetzmann et al., 2022; Graeber et al., 2022; Burro et al., 2023; Enke et al., 2023; Jiang et al., 2023). In line with a large literature in psychology (Tulving and Schacter, 1990; Schacter, 2008; Kahana, 2012), a common finding is that the recall of past observations is affected by the similarity of the past observation and the current context.[3] Using a representative survey of individual investors, Jiang et al. (2023) identify similarity-weighted memory as a key mechanism in the formation of investor beliefs. A theoretical literature analyzes the effect of memory primarily on individual decision making (Gilboa and Schmeidler, 1995; Mullainathan, 2002; Azeredo da Silveira and Woodford, 2019; Bodoh-Creed, 2020; Nagel and Xu, 2022; Bordalo et al., 2023a; Fudenberg et al., 2023; Wachter and Kahana, 2023). I extend this literature by showing that similarity-weighted memory can jointly explain empirical facts about subjective

[3]The memory literature identified three main regularities of selective recall: Similarity (a higher likelihood of recalling observations that are similar to today's context), recency (a higher likelihood of recalling recent rather than past observations), and contiguity (a higher likelihood of recalling observations that co-occur temporally). Recency is closely related to extrapolative expectations (Nagel and Xu, 2022).

expectations, subjective risk perceptions, and both subjective and objective returns in the aggregate. Policy makers may need to take the interdependence between memory, beliefs, and asset prices into account.

My paper also contributes to the growing theoretical literature on subjective beliefs in asset pricing (for an overview, see Adam and Nagel, 2023). A common empirical finding is that individual investor expectations are procyclical: Investors expect asset prices to rise further after high returns and to continue to fall after low returns (Vissing-Jorgensen, 2004; Bacchetta et al., 2009; Amromin and Sharpe, 2014; Greenwood and Shleifer, 2014; Kuchler and Zafar, 2019; Da et al., 2021). Theoretical research modeled procyclical expectations by assuming over-extrapolation (Barberis et al., 2015; Adam et al., 2017; Barberis et al., 2018; Jin and Sui, 2022), diagnostic expectations (Bordalo et al., 2018, 2019), partial-equilibrium thinking (Bastianello and Fontanier, 2022), or overlapping generations that learn from personal experiences (Ehling et al., 2018; Malmendier et al., 2020). Li and Liu (2023) show theoretically that procyclical fundamental expectations, but not procyclical return expectations, lead to a volatile equity premium; and Nagel and Xu (2022) and Bordalo et al. (2023b) empirically find that procyclical fundamental expectations explain the predictably countercyclical and quantitatively high equity premium. However, models that assume fundamental extrapolation (Hirshleifer et al., 2015; Nagel and Xu, 2022) typically do not give rise to return extrapolation. I show that similarity-weighted memory can simultaneously explain procyclical fundamental expectations, return extrapolation, and time-variation in the subjective volatility.

Studies that incorporate memory into asset pricing have focused mainly on fading memory to account for the evidence that lifetime experiences shape macroeconomic expectations (Malmendier and Nagel, 2011, 2016; Happel et al., 2022; Malmendier and Wachter, 2022), and Nagel and Xu (2022) analyze asset prices an economy in which a representative agent learns with fading memory. Fading memory implies that the impact of a past experience on the agent's beliefs gradually decreases, and any past experience will eventually be forgotten. In contrast and consistent with the evidence on the long-term effect of experiences on economic decisions, past experiences have a long-lasting effect on the agent's beliefs and are never truly forgotten under selective memory.[4] In addition to the asset pricing results in Nagel and Xu (2022), I also obtain a time-varying subjective risk premium and return extrapolation.

Selective memory is used as motivation for diagnostic expectations (Bordalo et al., 2022), which have been used to explain credit cycles (Bordalo et al., 2018) and cross-sectional variation of stock returns (Bordalo et al., 2019). In this paper, I explicitly model the recall of past observations, obtain overreaction of expectations in an i.i.d. economy, and apply the model to analyze time-series properties of aggregate asset prices. Wachter and Kahana (2023) provide a psychologically motivated theory of associative recall, in which the agent's current context cues memories of prior associated contextual states. Their focus is on decision making, which is distinct from this paper.

My model also builds on previous work that analyzes the asset-pricing implications of learning (Timmermann, 1993; Lewellen and Shanken, 2002; Weitzman, 2007). Collin-Dufresne et al. (2016) analyze the asset pricing implications of parameter learning without memory distortions. In their model, a representative agent with Epstein and Zin (1989)-preferences prices the parameter uncertainty that emerges from gradual Bayesian learning. I abstract from parameter uncertainty in my main analysis. Using the numerical methods developed in Johnson (2007) and Collin-Dufresne et al. (2016), I show that parameter uncertainty emerging from selective memory leads to a realistically high risk premium.

I also incorporate results from the statistics and economics literature on misspecified learning (Berk, 1966; Esponda and Pouzo, 2016; Molavi, 2019; Heidhues et al., 2021; Molavi et al., 2023) into the asset pricing literature. Fudenberg et al. (2023) propose the concept

[4]Early studies of human memory (Ebbinghaus, 1885; Jost, 1897; Müller and Pilzecker, 1900; Carr, 1931) focused on the finding that past experiences seem to be forgotten as a function of time decay (power law of forgetting), but experimental findings challenged the notion that experiences are ever truly forgotten. Instead, if the context of the original experience is reinstated, seemingly forgotten memories are typically recalled (Kahana, 2012).

of posterior beliefs as maximizers of the memory-weighted likelihood that is central to my characterization of subjective long-term beliefs.

The paper proceeds as follows. In Section 2, I describe the framework for my analysis of subjective beliefs under selective memory and introduce the asset pricing model. In Section 3, I apply the model to analyze the effect of similarity-weighted memory on investor beliefs and asset prices. Section 4 briefly analyzes an extension of the model to the peak-end memory distortion, and Section 5 concludes. All proofs are in the appendix.

2 Beliefs and asset prices under a general selective memory distortion

In this section, I characterize the agent's long-term beliefs under selective memory and describe the asset pricing framework. Selective memory is my only departure from a rational expectations model, and all asset pricing effects are driven by the agent's subjective long-term beliefs. I specify the learning environment and model of selective memory in Section 2.1. To simplify the exposition, I focus on discrete distributions as in Fudenberg et al. (2023). An extension to continuous distributions is in Online Appendix OA.1. Proposition 1 in Section 2.2 is new and characterizes the agent's subjective beliefs for normal distributions, as is relevant for applications in finance. Similarity-weighted memory (Section 3) and the peak-end memory distortion (Section 4) are special cases of the treatment. I then describe the asset pricing framework in Section 2.3, where I use the canonical model by Martin (2013) that nests the standard consumption-based asset pricing model (Mehra and Prescott, 1985). Readers who are interested in the applications may prefer to go directly to the specific section.

2.1 Learning framework

In this subsection, I formalize the learning problem of an agent with selective memory.

Economy. I study a representative agent endowment economy in discrete time. In every period t, the agent observes the state of the world s_t and log endowment growth $g_t = \log C_t/C_{t-1}$. The assets in the economy are levered claims on the endowment stream. I assume that the state $s_t = s$ is drawn from a finite set $S \subseteq \mathbb{N}$ according to the fixed, i.i.d. and full support distribution $\Xi \in \Delta(S)$. The state $s_t = s$ induces a fixed and i.i.d. distribution $q_s^* \in \Delta(G)$ over the finite set of possible endowment growth realizations G, that is $q_s^*(g) = Pr\,(g_t = g|s_t = s)$.[5] I assume that q_s^* belongs to the family of parametric distributions, $q_s^* \in \{q_\theta : \theta \in \Theta\}$, $\forall s \in S$, with $\Theta \subseteq \mathbb{R}^k, k \in \mathbb{N}$, closed and convex.

Learning. The agent knows the distribution Ξ of states, but must learn the distribution of log endowment growth. To model uncertainty about the distribution of log endowment growth, I assume that the agent holds a prior belief b_0 over potential distributions $q \in \Delta(G)^{|S|}$, where $q_s(g)$ denotes the probability of $g_t = g$ when $s_t = s$, and q specifies one induced distribution q_s for each state $s \in S$. The support of the prior is $\mathcal{Q}$ and contains all q that the agent considers possible. I focus on the case in which the agent considers only parametric distributions $q_s \in \{q_\theta : \theta \in \Theta\}$, $\forall s \in S$, and impose two additional regularity conditions on the prior. First, the agent is correctly specified $q^* \in Q$ (Esponda and Pouzo, 2016; Fudenberg et al., 2023), which implies that the agent eventually learns the true distribution without memory distortions. Second, for all $q \in Q$ and all $s \in S$, it holds that $q_s^*(g) > 0$ implies $q_s(g) > 0$.

Memory. The agent observes an infinite history of log endowment growth and state realizations, $H_t = \{(g_\tau, s_\tau)\}_{\tau=-\infty}^{t}$, where I call the tuple (g_τ, s_τ) an *experience*. $t_s = \sum_\tau \mathbb{1}_{\{s_\tau = s\}}$ denotes the number of experiences with $s_\tau = s, \tau \leq t$. In any period t, the agent recalls a subset of past experiences. The agent always observes and recalls the current experience (g_t, s_t), but her memory of any past experience is distorted by the *memory function*

[5] The assumption of a finite set of possible endowment growth realizations is for simplicity and allows me to directly use the results from Fudenberg et al. (2023) here. Behaviorally, the restriction can be justified by assuming that the agent only observes and recalls a discrete approximation of endowment growth, potentially due to limited attention. The results extend to continuous distributions, but I defer the discussion of continuous distributions to Online Appendix OA.1.

$m_{(g_t,s_t)} : G \times S \mapsto [0,1]$. For $\tau < t$, the value of the memory function $m_{(g_t,s_t)}(g_\tau, s_\tau)$ specifies the probability with which the agent recalls past experience (g_τ, s_τ) given (g_t, s_t). The *recalled periods* r_t are a subset of $\{-\infty, ..., t\}$ and the *recalled history* $H_t^R \subseteq H_t$ is the collection of recalled experiences $\{(g_\tau, s_\tau)\}_{\tau \in r_t}$ with $|H_t^R|$ past experiences. Similarly, $r_{s,t}$ denotes the recalled periods in state s and the *recalled history of state s* $H_{s,t}^R \subseteq H_t^R$ is the collection of recalled experiences with $s_\tau = s$.

Beliefs. The agent forms Bayesian beliefs as if her recalled history H_t^R is all that occurred (naïvety).[6] Her posterior belief in period t is

$$b_t\big(A|H_t^R\big) = \frac{\int_{q\in A} \prod_{\tau\in r_t} q_{s_\tau}(g_\tau)\, db_0(q)}{\int_{q\in Q} \prod_{\tau\in r_t} q_{s_\tau}(g_\tau)\, db_0(q)} \;\forall A \subseteq \mathcal{Q}, \tag{1}$$

where A is a (sub-)set of probability distributions in the agent's prior support $\mathcal{Q}$.

2.2 Subjective long-term beliefs

I next characterize the agent's subjective long-term beliefs. Define the *memory-weighted likelihood maximizer* (Fudenberg et al., 2023) conditional on this period's experience (g_t, s_t) as

$$\mathrm{LM}(g_t, s_t) = \underset{q\in Q}{\mathrm{argmax}}\Big(\sum_{s\in S} \psi(s) \sum_{g\in G} m_{(g_t,s_t)}(g, s)\, q_s^*(g)\, \log q_s(g)\Big). \tag{2}$$

The memory-weighted likelihood maximizer is the element of the agent's prior support that maximizes the likelihood of the recalled history H_t^R. Fudenberg et al. (2023) show that the agent's beliefs after a sufficiently long realized history H_t are given by the memory-weighted likelihood maximizer. The agent observes an infinite history of experiences and

[6] First, the agent recomputes her beliefs each period based on all recalled information and does not sequentially update her belief in period $t-1$ based on the experience (g_t, s_t). d'Acremont et al. (2013) and Sial et al. (2023) present evidence that humans access their accumulated evidence when forming beliefs. Second, modelling partial naïvety requires assumptions on the agent's perception of her memory function. If the agent anticipates her memory selectivity, she will perfectly undo any memory bias and learn q^*. Alternatively, if she believes that her recalled experiences are representative for the experiences she does not recall, then her belief b_t is not affected by partial naïvety. An analysis of intermediate assumptions about the agent's perception of her memory selectivity is provided in Fudenberg et al. (2023).

the empirical frequency of log endowment growth conditional on the state converges almost surely to the true distribution, q_s^*. However, the agent selectively recalls past experiences, and the frequency of recalled experiences converges to a memory-weighted version of the true distribution. It is a property of Bayesian learning that distributions that do not maximize the likelihood of the recalled experiences have vanishing posterior probability,[7] which implies that the agent's beliefs concentrate on the memory-weighted likelihood maximizer.[8]

I next illustrate the effect of selective memory on the agent's subjective beliefs if log endowment growth is normally distributed conditional on the state, as is relevant for asset pricing applications. Assume that log endowment growth is drawn from a normal distribution, $q_s^* \in \{q_\theta = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(g-\mu)^2}{2\,\sigma^2}\right) | \theta = (\mu; \sigma^2) \in \Theta, \mu \in \mathbb{R}, \sigma^2 \in \mathbb{R}_{\geq 0}, g \in \mathbb{R}\} =: \Theta_{\mathcal{N}}$, where Θ is closed and convex.[9] I associate each q_s^* with a parameter vector $\theta_s = (\mu_s, \sigma_s^2)$ whenever no confusion arises. Assumption 1 holds for the remainder of this paper, and Proposition 1 shows how the agent's state-wise posterior belief depends on selective memory. The agent observes the state realizations and thus performs state-wise inference.

Assumption 1 The prior support is $\Theta_{\mathcal{N}}^{|S|}$, and $q^* \in \Theta_{\mathcal{N}}^{|S|}$.

Proposition 1 (Normal posterior under selective memory)**.** *For each state $s \in S$, the agent's belief $b_{s,t}$ is almost surely given by the unique normal distribution with $\hat{\theta}_{s,t} := (\hat{\mu}_{s,t}, \hat{\sigma}^2_{s,t})$, and*

$$\hat{\mu}_{s,t} = \mu_s + \underbrace{\mathbb{E}\Big[\frac{t_s}{|H^R_{s,t}|}\Big]}_{\textit{Forgetfulness}} \cdot \underbrace{Cov\Big[g, \mathbb{1}_{\{g \in H^R_{s,t}\}}\Big]}_{\textit{Selectivity}}, \textit{ and} \tag{3}$$

[7]See Berk (1966)'s concentration result, and the Bernstein-von-Mises theorem under model misspecification (Kleijn and Van Der Vaart, 2012).

[8]Without memory distortions, $m_{(g_t,s_t)}(g_\tau, s_\tau) = 1,\ \forall (g_\tau, s_\tau) \in G \times S$, the distribution of recalled experiences is identical to the distribution of log endowment growth, and the agent learns q^* because she is correctly specified. Similarly, $\text{LM}(g_t, s_t)$ does not depend on the "scale" of the memory function. If $m_{(g_t,s_t)}(\cdot) = c\ m'_{(g_t,s_t)}(\cdot), c > 0$, then both memory functions have the same memory-weighted likelihood maximizer. The agent learns q^* if her memory is only stochastic, but not selective. The framework nests rational expectations.

[9]I use g to refer to endowment growth as a random variable instead of as a specific realization g_τ. The closedness of Θ implies that $\mu \in [\underline{\mu}; \bar{\mu}]$ with $\underline{\mu} > -\infty$ and $\bar{\mu} < \infty$, and $\sigma^2 \in [0, \bar{\sigma}^2]$ with $\bar{\sigma}^2 < \infty$.

$$\hat{\sigma}^2_{s,t} = \sigma^2_s + \left(\hat{\mu}_{s,t} - \mu_s\right)^2 + \underbrace{\mathbb{E}\left[\frac{t_s}{|H^R_{s,t}|}\right]}_{\text{Forgetfulness}} \cdot \underbrace{Cov\left[(g - \hat{\mu}_{s,t})^2, \mathbb{1}_{\{g \in H^R_{s,t}\}}\right]}_{\text{Selectivity}}, \tag{4}$$

where the indicator random variable $\mathbb{1}_{\{g \in H^R_{s,t}\}}$ *equals one if the agent recalls endowment growth* $g_\tau = g$ *with* $s_\tau = s$*, and zero otherwise.*

Equation 3 shows that the agent's posterior mean of log endowment growth in state s depends on two elements: (i) the true fundamental mean μ_s, and (ii) an adjustment term that arises from selective memory. The "forgetfulness" term in the adjustment, $\mathbb{E}\left[\frac{t_s}{|H^R_{s,t}|}\right]$, is the expected size of the realized history relative to the recalled history of state s. If the agent recalls almost all past observations, $m_{(g_t,s_t)}(g_\tau, s_\tau) \approx 1$, the recalled history will be as long as the realized history, $|H^R_{s,t}| \approx t_s$. The recalled history will be "shorter" than the realized history if the agent, instead, barely recalls past observations.[10] The second term in the adjustment, $\text{Cov}\left[g, \mathbb{1}_{\{g \in H^R_{s,t}\}}\right]$, captures the selectivity. The posterior mean will be higher than the true mean if the agent is more likely to recall high log endowment growth rates from state s, as measured by the covariance between g and the propensity of recalling $g_\tau = g$.[11] On the contrary, if the agent is as likely to recall high as low log endowment growth rates, such that $\text{Cov}\left[g, \mathbb{1}_{\{g \in H^R_{s,t}\}}\right] \approx 0$, the agent almost surely learns μ_s.

The agent's posterior variance of log endowment growth in state s (Equation 4) is anchored at the true underlying variance σ^2_s and learned by an agent without memory distortions. The second term in Equation 4, $(\hat{\mu}_{s,t} - \mu_s)^2$, is the usual adjustment for the usage of a biased mean estimate. The last term in Equation 4, $\mathbb{E}\left[\frac{t_s}{|H^R_{s,t}|}\right] \cdot \text{Cov}\left[(g - \hat{\mu}_{s,t})^2, \mathbb{1}_{\{g \in H^R_{s,t}\}}\right]$, captures the direct effect of selective memory on the agent's posterior variance. Selective memory increases the posterior variance whenever the agent is more likely to recall more spread-out log endowment growth rates, while selectivity decreases the posterior variance if

[10] Using Jensen's Inequality, it is $\mathbb{E}\left(\frac{1}{X}\right) \geq \frac{1}{\mathbb{E}(X)}$ for a positive random variable X; and therefore $\mathbb{E}\left[\frac{t}{|H^R_t|}\right] \geq \frac{t}{\mathbb{E}\left[|H^R_t|\right]}$.

[11] In my formulation of selective memory, the agent recalls a past experience with probability $m_{(g_t,s_t)}(g_\tau, s_\tau)$ and does not recall the experience otherwise. Therefore, $\mathbb{1}_{\{g \in H^R_t\}}$ is a random variable with $\mathbb{E}\left(\mathbb{1}_{\{g \in H^R_t\}}\right) = m_{(g_t,s_t)}(g_\tau, s_\tau)$.

the agent tends to recall endowment growth rates that are close to the posterior mean.

Depending on the memory specification and on today's experience (g_t, s_t), the agent can be too optimistic $(\hat{\mu}_{s,t} > \mu_s)$ or too pessimistic $(\hat{\mu}_{s,t} < \mu_s)$. Likewise, the agent might perceive the economy as more risky $(\hat{\sigma}^2_{s,t} > \sigma^2_s)$ or less risky $(\hat{\sigma}^2_{s,t} < \sigma^2_s)$ than it truly is. In addition, selective memory can also lead to time-variation in the agent's beliefs because the propensity to recall a given experience may depend on the current experience (g_t, s_t). I apply Proposition 1 to similarity-weighted memory (Section 3) and the peak-end memory distortion (Section 4).

2.3 Asset pricing framework

In this subsection, I incorporate the agent's subjective beliefs that arise from selective memory into a standard consumption-based asset pricing model (Lucas, 1978; Mehra and Prescott, 1985). For analytical tractability, I follow the framework of Martin (2013).[12]

I assume that the representative agent has Epstein and Zin (1989)-preferences

$$U_t = \left\{ (1-\beta) C_t^{\frac{1-\gamma}{\eta}} + \beta \left(\tilde{\mathbb{E}}_t \left[U_{t+1}^{1-\gamma} \right] \right)^{\frac{1}{\eta}} \right\}^{\frac{\eta}{1-\gamma}}, \tag{5}$$

with discount factor β, risk-aversion γ, elasticity of intertemporal substitution (EIS) ψ, and composite parameter $\eta = \frac{1-\gamma}{1-1/\psi}$. In any period t, the agent maximizes expected lifetime utility under subjective expectations $\tilde{\mathbb{E}}_t(\cdot)$ formed under posterior belief b_t.

The agent is unaware of her memory distortions and treats her recalled experiences as if they were all that ever occurred. Although the agent's recalled information does not form a filtration, the agent, at any time t, holds an internally consistent set of beliefs and behaves as if the law of iterated expectations holds (Adam and Nagel, 2023), such that the economy is as in Martin (2013).

Consider an asset that pays a dividend stream $\{D_{t+k}\}_{k \geq 0}$ with $D_{t+k} = C^{\lambda}_{t+k}$ for some

[12] Appendix F gives a more detailed derivation of the results.

constant λ. If $\lambda = 0$, the asset is a riskless bond that pays 1 in each period; if $\lambda = 1$, the asset is the aggregate consumption claim; and $\lambda > 1$ is a levered claim (Campbell, 1986; Abel, 1999). Define the *log dividend-price ratio* as $dp_t = \log(1 + \frac{D_t}{P_t})$. The return on any asset is $R_{t+1} = \frac{D_{t+1}+P_{t+1}}{P_t} = \frac{D_{t+1}}{D_t} \frac{D_t}{P_t} \left(1 + \frac{P_{t+1}}{D_{t+1}}\right)$, and the *log subjective expected return* is $\tilde{er}_t = \log\left(\tilde{\mathbb{E}}_t \; R_{t+1}\right)$. Similarly, the *log risk-free rate* is the log (subjective) expected return on the riskless bond,[13] and the *subjective risk premium* on the λ-asset is the difference between the log subjective expected return and the log risk-free rate. Martin (2013) shows that[14]

$$r_t^f = -\log(\beta) - \mathcal{K}_t(-\gamma) + \left(1 - \frac{1}{\eta}\right) \mathcal{K}_t(1-\gamma), \tag{6}$$

$$dp_t = -\log(\beta) - \mathcal{K}_t(\lambda - \gamma) + \left(1 - \frac{1}{\eta}\right) \mathcal{K}_t(1-\gamma), \tag{7}$$

$$\tilde{er}_t = dp_t + \mathcal{K}_t(\lambda), \tag{8}$$

$$\tilde{rp}_t = \tilde{er}_t - r_t^f = \mathcal{K}_t(\lambda) + \mathcal{K}_t(-\gamma) - \mathcal{K}_t(\lambda - \gamma), \tag{9}$$

where $\mathcal{K}_t(k)$ is the cumulant-generating function under the agent's subjective beliefs in period t. The moment-generating function $\mathcal{M}_t(k)$ and the cumulant-generating function $\mathcal{K}_t(k)$ under the agent's subjective beliefs are defined as

$$\mathcal{M}_t(k) := \tilde{\mathbb{E}}_t\left(e^{k\, g_{t+1}}\right), \text{ and}$$

$$\mathcal{K}_t(k) := \log\left(\mathcal{M}_t(k)\right) = \log \tilde{\mathbb{E}}_t\left(e^{k\, g_{t+1}}\right), \text{ respectively.}$$

Both the moment-generating function and the cumulant-generating function provide expressions for the moments of log endowment growth under the agent's posterior belief b_t.

[13] Note that, in equilibrium, the risk-free rate and asset prices in period t are determined under the agent's subjective measure and thus objectively realized. The realized return and the realized risk-premium, instead, may deviate from subjective expectations because they depend on next period's dividend payment and price.

[14] The asset-pricing quantities for the power-utility case follow for $\psi = 1/\gamma$, which implies $\eta = 1$.

Expanding the cumulant-generating function $\mathcal{K}_t(k)$ as a power series yields

$$\mathcal{K}_t(k) = \sum_{n=1}^{\infty} c_n \frac{k^n}{n!},$$

with cumulants c_n. The first four cumulants are related to the first four moments of the agent's posterior belief: $c_1 \equiv \hat{\mu}_t$ is the posterior mean of the agent, $c_2 \equiv \hat{\sigma}_t^2$ is the agent's posterior variance, $\frac{c_3}{c_2^{3/2}}$ is the skewness, and $\frac{c_4}{c_2^2}$ is the excess kurtosis under the agent's posterior belief b_t.

I next discuss how the agent's subjective beliefs affect equilibrium asset prices. To gain intuition, consider power-utility preferences ($\eta = 1$) and a second-order approximation of the cumulant-generating function, $\mathcal{K}_t(k) \approx k\, c_1 + \frac{1}{2} k^2\, c_2 = k\, \hat{\mu}_t + \frac{1}{2} k^2\, \hat{\sigma}_t^2$, which yields

$$\begin{aligned}
r_t^f &= -\log(\beta) + \gamma\, \hat{\mu}_t - \frac{1}{2}\gamma^2\, \hat{\sigma}_t^2, \\
dp_t &= -\log(\beta) - (\lambda - \gamma)\, \hat{\mu}_t - \frac{1}{2}(\lambda - \gamma)^2 \hat{\sigma}_t^2, \\
\tilde{er}_t &= -\log(\beta) + \gamma\, \hat{\mu}_t + \lambda\, \gamma\, \hat{\sigma}_t^2 - \frac{1}{2}\gamma^2\, \hat{\sigma}^2, \\
\tilde{rp}_t &= \lambda\, \gamma\, \hat{\sigma}_t^2.
\end{aligned}$$

Both the risk-free rate and the subjective expected return are increasing in the posterior mean of the agent, $\hat{\mu}_t$, while the dividend-price ratio is decreasing in $\hat{\mu}_t$ if $\lambda > \gamma$. Intuitively, if the agent becomes more optimistic (higher $\hat{\mu}_t$), she consumes more today. The risk-free rate and subjective expected return must then increase to induce saving/investment. The dividend yield dp_t decreases in $\hat{\mu}_t$ because the price—which reflects the discounted sum of all future dividends—increases in $\hat{\mu}_t$ if leverage λ exceeds the agent's risk-aversion γ.[15] Note that the subjective risk premium is independent of $\hat{\mu}_t$ because the risk-free rate and the subjective expected return both depend positively on $\gamma\, \hat{\mu}_t$.

[15] Instead, if $\gamma > \lambda$—which is empirically more relevant and considered below—the agent discounts future high dividend-payments more heavily due to the low marginal utility of high consumption and the divdend yield is increasing in $\hat{\mu}_t$. The effect does not arise under Epstein and Zin (1989)-preferences, which decouple risk-aversion and EIS.

Moreover, the agent's posterior variance—the agent's subjectively perceived risk in the economy—affects asset prices. The risk-free rate is decreasing in $\hat{\sigma}_t^2$, because the risk-averse agent has a precautionary savings motive that is stronger the more risky the economy appears. The decreasing risk-free rate also leads to a decreasing dividend yield dp_t due to a discount-rate effect. In addition, the posterior variance has two opposite effects on the subjectively expected return: First, the decrease in the risk-free rate leads to a decrease of the subjectively expected return, as the overall level of returns in the economy decreases. Second, the risk-averse agent requires a positive subjective risk premium, $rp_t = \lambda\,\gamma\,\hat{\sigma}_t^2$, which increases in the posterior variance. The expected return increases in the posterior variance if the risk-premium effect dominates the risk-free rate effect ($\lambda > \frac{1}{2}\,\gamma$).

3 Similarity-weighted memory

In this section, I assume that the agent's memory is distorted by a similarity-weighted memory function and impose structure on log endowment growth. I focus on similarity with respect to the log endowment growth (Section 3.1). Using Proposition 1, I analyze the agent's long-term beliefs under similarity-weighted memory in Section 3.2, and discuss the asset pricing implications in Section 3.3. Finally, I use standard data to estimate the parameters of log endowment growth under my structural assumptions and simulate asset prices in Section 3.4.

The mechanism in this section works as follows: The economy is either in a normal state or in a recession. Let endowment growth always be high in normal times, while it can be high or low during recessions (recessions have higher fundamental uncertainty). The agent observes the contemporaneous endowment growth (current context) and recalls past experiences with similar endowment growth. Since endowment growth is always high in normal times, the agent can only recall high endowment growth from normal times and her belief about normal times does not react to the current context. On the contrary, whether the agent recalls high or low endowment growth from recessions depends on the current context. If contemporaneous endowment growth is high (low), the agent recalls more high (low) endowment growth experiences from recessions, and her posterior mean will be high (low). Therefore, the agent's posterior mean covaries positively with the current context, and recessions are perceived to be worse if contemporaneous endowment growth is low than if it is high.[16] The agent becomes oblivious of recessions during good times, but recalls them vividly during bad times.

In line with this motivating example, Enke et al. (2023) conducted a series of laboratory experiments and document that similarity-weighted memory causes an overreaction of experimental market prices. In their experiment, subjects observe news about different companies, some of which are shown in a memorable context. The authors find that subjects asymmetrically recall past news if it is cued by the current context, which leads to an overreaction of expectations and of prices in a parimutuel betting market. Using a representative survey of retail investors, Jiang et al. (2023) present additional evidence of similarity-weighted memory as a key mechanism of belief formation in financial markets. The authors elicit investor memories of past returns and find support for similarity-weighted recall in that investors recall more positive past returns if today's stock market return is high. The recalled memories are highly correlated with expectations and have a higher explanatory power for investor beliefs than actual experiences.

3.1 Structural assumptions

The assumptions in Section 2.1 continue to hold, but I impose additional structure on the economy. Let $S = \{1, 2\}$, and $s_t = s$ follows a two-state observable Markov chain with constant transition matrix Π. The elements of Π are $\pi_{ij} = \Pr[s_t = j | s_{t-1} = i]$, and I restrict

[16] The example here is an extension of the thirsty traveler example in Bordalo et al. (2020b). Consider a thirsty traveler who recalls water prices from memory. At the airport, water prices are always high, thus the traveler can only recall high water prices at the airport experiences regardless of the current context. In contrast, water prices downtown are sometimes low (at the corner store) and sometimes high (at a luxury hotel), such that similarity-weighted memory systematically affects retrieved water prices downtown.

$\pi_{11} = \pi_{21} =: \pi_1$ and $\pi_{12} = \pi_{22} = 1 - \pi_1 =: \pi_2$ to ensure that the process is i.i.d. Conditional on state $s_t = s$, log endowment growth is normally distributed

$$g_t = \mu_{s_t} + \sigma_{s_t}\, \epsilon_t, \quad \epsilon_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1). \tag{10}$$

The mean μ_s and variance σ_s^2 are state-dependent. Let $\mu_1 > \mu_2$ and $\sigma_1^2 < \sigma_2^2$, such that state 1 corresponds to normal times, while state 2 captures recessions.

Markov-switching models have been widely used to model aggregate endowment dynamics due to their flexibility and tractability (Mehra and Prescott, 1985; Rietz, 1988; Barro, 2006; Johannes et al., 2016). I use the two-state structure to analyze how similarity-weighted memory affects the agent's perception of recessions compared to normal times. All results that relate to the agent's posterior mean hold if endowment growth is log-normally distributed, but the perceived riskiness of the economy is then constant (see Online Appendix OA.2.1).

As before, the agent relies on an infinite history of past endowment growth and state realizations, $H_t = \{(g_\tau, s_\tau)\}_{\tau=-\infty}^{t}$, to learn state-dependent parameters such that Proposition 1 holds state-wise. The recalled history H_t^R is distorted by a similarity-weighted memory function (see also Kahana, 2012; Jiang et al., 2023)

$$m^{\text{sim}}_{(g_t, s_t)}\,(g_\tau, s_\tau) = \exp\left[-\frac{(g_\tau - g_t)^2}{2\,\kappa}\right], \tag{11}$$

where $\kappa > 0$ captures the *scrutiny* with which the agent examines her memory database, and a high scrutiny implies that the agent recalls almost all past observations. An extension in which similarity also depends on the state is in Online Appendix OA.2.2.

3.2 Long-term beliefs

I next analyze the agent's long-term beliefs under similarity-weighted memory. First, I highlight central properties of the agent's subjective beliefs, and then analyze the rationality of the agent's forecast. Finally, I discuss the predictability of belief revisions.

Subjective beliefs. Proposition 2 characterizes the state-wise beliefs of an agent with a similarity-weighted memory distortion.

Proposition 2 (State-wise subjective beliefs)**.** *Under similarity-weighted memory as in Equation 11, almost surely,*

$$\hat{\mu}_{s,t} = \frac{\kappa}{\kappa + \sigma_s^2}\,\mu_s + \frac{\sigma_s^2}{\kappa + \sigma_s^2}\,g_t = (1-\alpha_s)\,\mu_s + \alpha_s\,g_t, \text{ and} \tag{12}$$

$$\hat{\sigma}_{s,t}^2 = (1-\alpha_s)\,\sigma_s^2, \tag{13}$$

where $\alpha_s := \frac{\sigma_s^2}{\kappa+\sigma_s^2} \in (0,1)$ *measures the sensitivity of the agent's belief to this period's log endowment growth* g_t*.*

Equation 12 shows that the state-dependent posterior mean of the agent is a convex combination of the true state-dependent mean μ_s and this period's log endowment growth g_t. If today's endowment growth is high, the agent is more likely to recall past experiences with a high endowment growth than with a low endowment growth due to similarity. The agent will therefore be more (less) optimistic if this period's endowment growth is high (low). Similarity-weighted memory distortions provide a microfoundation for extrapolative beliefs in that the agent's posterior is formed as if the agent overweights contemporaneous endowment growth when forming beliefs (for an overview on extrapolative beliefs, see Barberis, 2018).

The sensitivity of the agent's state-wise posterior mean to contemporaneous endowment growth g_t depends on the state-dependent variance σ_s^2 and the scrutiny κ, as summarized in α_s. If $\sigma_s^2 \to 0$, the true distribution of log endowment growth in state s is concentrated at μ_s. All observations that the agent can recall are very close to μ_s and the agent's posterior mean must be $\hat{\mu}_{s,t} \approx \mu_s$. If $\sigma_s^2 \to \infty$, the distribution of endowment growth in state s becomes flat, close to a uniform distribution, and the agent observes all endowment growth rates equally often. The agent's recalled experiences are then determined by the agent's memory function, which is symmetric around g_t and her posterior mean will be $\hat{\mu}_{s,t} \approx g_t$.

The reverse intuition holds for scrutiny parameter κ. If scrutiny is high ($\kappa \to \infty$), similarity becomes irrelevant because the agent always consults all memories and the agent's recalled experiences are determined by the distribution of log endowment growth in state s. If scrutiny is low ($\kappa \to 0$), similarity is very important and past experiences with log endowment growth that differs from the contemporaneous log endowment growth will not be recalled. The posterior mean of the agent equals g_t.

The agent's state-dependent posterior variance (Equation 13) is independent of the contemporaneous endowment growth and smaller than the fundamental state-dependent variance σ_s^2 since $\alpha_s \in (0,1)$. The agent's similarity-weighted memory distortion symmetrically overweights observations that are close to the contemporaneous log endowment growth, and the agent tends to forget experiences that are less similar to the contemporaneous log endowment growth. Therefore, the scale of the posterior distribution is smaller than the scale of the fundamental distribution.[17]

I focus on the case in which the agent knows the state-dependent variances, σ_1^2 and σ_2^2, but learns about the mean endowment growth μ_1 and μ_2.[18] Moreover, I now discuss properties of the agent's unconditional time-t belief about log endowment growth, which determines asset prices. Note that the unconditional distribution of log endowment growth is a mixture of the state-wise distributions, which is generally not a normal distribution even though log endowment growth is normally distributed conditional on the state. I thus characterize the unconditional (perceived and actual) distribution of log endowment growth using the cumulant-generating function (Section 2.3):

$$\mathcal{K}_t(k) = \log\left[\mathcal{M}_t(k)\right] = \log\left[\pi_1\, e^{k\,\hat{\mu}_{1,t} + \frac{1}{2}\,k^2\,\sigma_1^2} + \pi_2\, e^{k\,\hat{\mu}_{2,t} + \frac{1}{2}\,k^2\,\sigma_2^2}\right]. \tag{14}$$

The same expression holds under rational expectations with $\hat{\mu}_{1,t} = \mu_1$ and $\hat{\mu}_{2,t} = \mu_2$, and I

[17]In terms of Proposition 1, the covariance between recalling an experience and that experience being distant from the posterior mean is negative under similarity-weighted memory, but constant over time.

[18]Qualitatively, all results hold when the agent simultaneously learns about the state-wise variances. As the state-dependent posterior variances are constant, one only needs to replace σ_s^2 by $\hat{\sigma}_s^2$.

find the n'th (non-central) moment by taking the n'th derivative of the moment-generating function with respect to k and evaluating the derivative at $k = 0$. Although Martin (2013) shows that higher-order moments have a non-trivial effect on asset prices, I focus on the subjective mean and variance of log endowment growth, as these two moments have the most pronounced effect on asset prices. The simulation in Section 3.4 incorporates higher-order moments, and Figure D.1 in Appendix D plots the dependence of the first four moments on g_t.

The unconditional expected log endowment growth under the agent's posterior belief is

$$\tilde{\mathbb{E}}_{\mathrm{t}}\left(g_{t+1}\right) = \pi_1\,\hat{\mu}_{1,t} + \pi_2\,\hat{\mu}_{2,t}. \tag{15}$$

Equation 15 shows that agent's expected log endowment growth is the probability-weighted average of the state-wise posterior means. The expected log endowment growth is thus increasing in this period's endowment growth (procyclical). The sensitivity of the expected log endowment growth to the contemporaneous endowment growth depends on a weighted average of the state-wise variances and on the scrutiny κ. Proposition 3 characterizes the agent's state-wise and unconditional posterior mean.

Proposition 3 (Subjective mean)**.** *The average state-dependent posterior mean conditional on the current state is*

$$\mathbb{E}\left(\hat{\mu}_{1,t+1} | s_{t+1} = 1\right) = \mu_1 \tag{16}$$

$$\mathbb{E}\left(\hat{\mu}_{1,t+1} | s_{t+1} = 2\right) = \mu_1 + \alpha_1\,\left(\mu_2 - \mu_1\right), \tag{17}$$

and the average state-dependent posterior mean is

$$\mathbb{E}\left(\hat{\mu}_{1,t+1}\right) = \mu_1 + \alpha_1\,\pi_2\,\left(\mu_2 - \mu_1\right). \tag{18}$$

Equations 16—18 hold for $\hat{\mu}_{2,t+1}$ with the respective change of indices. Moreover, the average *unconditional posterior mean of log endowment growth is*

$$\mathbb{E}\left(\pi_1\,\hat{\mu}_{1,t+1}+\pi_2\,\hat{\mu}_{2,t+1}\right)=\pi_1\,\mu_1+\pi_2\,\mu_2+\pi_1\,\pi_2\,\left(\mu_2-\mu_1\right)\,\left(\alpha_1-\alpha_2\right). \tag{19}$$

The posterior mean of state 1 is unbiased if state 1 occurs, but is biased toward the mean of state 2 if state 2 occurs. The reasoning is as follows: If today's state $s_t = s$, then log endowment growth is drawn from a distribution centered at μ_s, such that the realized endowment growth g_t will, on average, be μ_s in state s. On average, the posterior mean of state 1 is thus unbiased if $s_t = 1$, but pulled to μ_2 if $s_t = 2$. As a result, the average state-dependent posterior mean, which is a combination of the average state-dependent posterior mean conditional on each state, is biased towards the mean of the other state.

The average unconditional posterior mean given in Equation 19 is biased upward, since $\mu_1 > \mu_2$ and $\alpha_1 < \alpha_2$ by assumption. In normal times, $s_t = 1$, endowment growth tends to be high and the posterior mean of state 2 is biased upwards. Similarly, the posterior mean of state 1 is biased downwards in a recession, $s_t = 2$, but the downward bias of the posterior mean of state 1 is smaller than the upward bias of the posterior mean of state 2 because $\alpha_1 < \alpha_2$. The net effect is an upward bias of the unconditional posterior mean, such that the agent is, on average, too optimistic about endowment growth.[19]

The posterior variance of log endowment growth under the agent's beliefs is

$$\mathrm{Var}_t\left(g_{t+1}\right)=\pi_1\,\sigma_1^2+\pi_2\,\sigma_2^2+\pi_1\,\pi_2\,\left(\hat{\mu}_{1,t}-\hat{\mu}_{2,t}\right)^2. \tag{20}$$

Equation 20 highlights that the perceived riskiness of the economy, $\mathrm{Var}_t(g_{t+1})$, depends not only on the state-wise variances (σ_1^2 and σ_2^2), but also on the squared distance between the state-wise posterior means $(\hat{\mu}_{1,t}-\hat{\mu}_{2,t})^2$. Intuitively, the agent perceives the economy as more risky if recessions are severe compared to normal times ($\hat{\mu}_{2,t} << \hat{\mu}_{1,t}$). Proposition 4

[19] It follows that the posterior mean under similarity-weighted memory is unbiased if endowment growth is log-normally distributed.

characterizes the average posterior variance.

Proposition 4 (Subjective unconditional variance)**.** *The average posterior variance of the agent is*

$$\mathbb{E}\left[\mathrm{Var}_t\left(g_{t+1}\right)\right] = \left(\pi_1\,\sigma_1^2 + \pi_2\,\sigma_2^2\right)\left[1 + \pi_1\,\pi_2\,\left(\alpha_1 - \alpha_2\right)^2\right]$$
$$+ \left(\mu_1 - \mu_2\right)^2\,\pi_1\,\pi_2\,\left[\pi_1\,\pi_2\left(\alpha_1 - \alpha_2\right)^2 + \left[1 - \left(\alpha_1\,\pi_2 + \alpha_2\,\pi_1\right)\right]^2\right], \tag{21}$$

which is larger than the true variance $\mathrm{Var}\left(g_{t+1}\right)$ *if*

$$\frac{\left(\alpha_1 - \alpha_2\right)^2\,\left(\pi_1\,\sigma_1^2 + \pi_2\,\sigma_2^2\right)}{2\,\left(\pi_2\,\alpha_1 + \pi_1\,\alpha_2\right) - \left(\pi_2\,\alpha_1^2 + \pi_1\,\alpha_2^2\right)} \geq \left(\mu_1 - \mu_2\right)^2, \tag{22}$$

and bounded by

$$0 \leq \mathbb{E}\left[\mathrm{Var}_t\left(g_{t+1}\right)\right] \leq 1.25\ \mathrm{Var}(g_{t+1}).$$

Condition 22 allows me to characterize situations in which the average perceived riskiness of the economy exceeds the fundamental riskiness. The left-hand side of condition 22 is always positive, so that Condition 22 holds for $(\mu_1 - \mu_2) \to 0$. Similarity-weighted memory systematically increases the perceived variance if the mean in both states is approximately equal because the sensitivity of the posterior means to the contemporaneous endowment growth differs. Additionally, the left-hand side of Condition 22 increases in the difference of the state-dependent variances σ_1^2 and σ_2^2. The sensitivity of the posterior mean of state s depends on the state-dependent variance σ_s^2. In situations in which $\sigma_2^2 >> \sigma_1^2$, the posterior mean of state 2 reacts more strongly to this period's endowment growth g_t than the posterior mean of state 1, such that the squared difference of the posterior means is, on average, higher if $\sigma_2^2 >> \sigma_1^2$ than if $\sigma_2^2 \approx \sigma_1^2$.

Moreover, I have $\alpha_2 > \alpha_1$, such that the posterior mean of state 2 is more sensitive to the contemporaneous endowment growth than the posterior mean of state 1. Define

$g^* = \frac{(1-\alpha_1)\,\mu_1 + (1-\alpha_2)\,\mu_2}{\alpha_2 - \alpha_1}$ as the log endowment growth for which $\hat{\mu}_{1,t} = \hat{\mu}_{2,t}$. The difference between the posterior means $\hat{\mu}_{1,t} - \hat{\mu}_{2,t}$ decreases in the log endowment growth for $g_t < g^*$, equals zero for $g_t = g^*$, and increases in g_t whenever $g_t > g^*$. Therefore, the unconditional posterior variance of the agent also decreases in g_t for $g_t < g^*$, and increases in g_t for $g_t > g^*$. The perceived riskiness of the economy is a convex function of the contemporaneous log endowment growth.

Forecast rationality. I now compare the expectations of an agent under similarity-weighted memory to the expectations of an agent without memory distortions.[20] Intuitively, realizations of a stochastic process should move one-for-one with a rational forecast, such that forecast rationality implies $a_{MZ} = 0$ and $\beta_{MZ} = 1$ in the following Mincer and Zarnowitz (1969)-regression:

$$g_{t+h} = a_{MZ} + \beta_{MZ}\,\tilde{\mathbb{E}}_{\mathrm{t}}\left(g_{t+h}\right) + u_{t+h}. \tag{23}$$

Another property of rational forecasts is that they should neither overreact nor underreact to new information. Coibion and Gorodnichenko (2015) propose the following regression to test over- or underreaction of forecasts:

$$g_{t+h} - \tilde{\mathbb{E}}_{\mathrm{t}}\left(g_{t+h}\right) = a_{CG} + \beta_{CG}\left[\tilde{\mathbb{E}}_{\mathrm{t}}\left(g_{t+h}\right) - \mathbb{E}_{t-1}\left(g_{t+h}\right)\right] + u_{t+h}. \tag{24}$$

Forecast rationality implies $a_{CG} = 0$ and $\beta_{CG} = 0$ in Equation 24, because the forecast revision $\tilde{\mathbb{E}}_{\mathrm{t}}\left(g_{t+h}\right) - \mathbb{E}_{t-1}\left(g_{t+h}\right)$ is known to the agent at time t and should thus not predict forecast errors. Otherwise, the agent would adjust the forecast. If $\beta_{CG} < 0$, the forecaster overreacts since the revision is too strong on average, while $\beta_{CG} > 0$ captures an underreaction of the forecast. Proposition 5 shows that I can reject forecast rationality in Mincer and Zarnowitz (1969)-regressions and find overreaction in Coibion and Gorodnichenko (2015)-regressions for an agent with similarity-weighted memory.

[20] Note that I can always obtain the rational benchmark for $\kappa \to \infty$.

Proposition 5 (Properties of subjective forecasts)**.** *Under similarity-weighted memory, I can reject rationality of the agent's forecast as measured by Mincer and Zarnowitz (1969)-regressions in Equation 23, since*

$$\beta_{MZ} = 0 < 1, \qquad \textit{and} \qquad a_{MZ} = \pi_1\,\mu_1 + \pi_2\,\mu_2 \neq 0. \tag{25}$$

The long-term beliefs of an agent with similarity-weighted memory overreact as measured by Coibion and Gorodnichenko (2015)-regressions in Equation 24, since

$$\beta_{CG} = -\frac{1}{2} < 0, \qquad \textit{and} \qquad a_{GC} = \pi_1\,\pi_2\,(\mu_2 - \mu_1)\,(\alpha_1 - \alpha_2)\,. \tag{26}$$

The results summarized in Proposition 5 show that the agent's forecast is uninformative for the realization of log endowment growth ($\beta_{MZ} = 0$). The best forecast of future realizations of log endowment growth is the long-term mean $\pi_1\mu_1 + \pi_2\mu_2$, because log endowment growth is i.i.d. The agent's expectation, however, covaries with the current realization of endowment endowment growth g_t due to similarity-weighted memory, but the current realization of an i.i.d. process is not predictive for future realizations, yielding $\beta_{MZ} = 0$. Similarity-weighted memory also leads to an overreaction of the agent's forecast ($\beta_{GC} < 0$). For simplicity, focus on $\hat{\mu}_{1,t} = (1 - \alpha_1)\,\mu_1 + \alpha_1\,g_t$. The agent revises the posterior mean of state 1 upward if and only if tomorrow's log endowment growth exceeds the contemporaneous log endowment growth, $g_{t+1} > g_t$. Thus, conditional on upward revision, it is $\Pr(g_{t+1} \geq \mu_1 | g_{t+1} > g_t) > 0.5$, since g_{t+1} must exceed g_t. In addition, we also know that $\hat{\mu}_{1,t+1}$ exceeds the true mean μ_1 if and only if $g_{t+1} > \mu_1$. Consequently, the agent's posterior mean is more likely above than below the fundamental mean after an upward revision, implying a predictably negative forecast error.

Belief predictability. As a last step in the analysis of the agent's belief, I examine the predictability of the agent's belief revisions. The agent's subjective beliefs drive asset prices, such that the predictability of beliefs implies predictability of objective returns.

An econometrician with access to the same realized and infinite history H_t as the agent will perfectly uncover the true parameters of the data-generating process, and can forecast the agent's next-period beliefs. The expected subjective moment-generating function is

$$\begin{aligned}\tilde{\mathcal{M}}(m) := \mathbb{E}\left[\mathcal{M}_{t+1}(m)\right] =& \pi_1\,\mathbb{E}\left[\,e^{m\,\alpha_1\,g_{t+1}}\right]\;e^{m\,(1-\alpha_1)\,\mu_1+\frac{1}{2}\,m^2\,\sigma_1^2}+\\ & \pi_2\,\mathbb{E}\left[\,e^{m\,\alpha_2\,g_{t+1}}\right]\;e^{m\,(1-\alpha_2)\,\mu_2+\frac{1}{2}\,m^2\,\sigma_2^2}\\ =& \pi_1\,e^{\mathcal{K}^*(m\,\alpha_1)+m\,(1-\alpha_1)\,\mu_1+\frac{1}{2}\,m^2\,\sigma_1^2}\\ & +\pi_2\,e^{\mathcal{K}^*(m\,\alpha_2)+m\,(1-\alpha_2)\,\mu_2+\frac{1}{2}\,m^2\,\sigma_2^2},\end{aligned} \tag{27}$$

where $\mathcal{K}^*(k) = \log\left(\pi_1\,e^{k\,\mu_1+\frac{1}{2}\,k^2\,\sigma_1^2}+\pi_2\,e^{k\,\mu_2+\frac{1}{2}\,k^2\,\sigma_2^2}\right)$ denotes the true cumulant-generating function of endowment growth. Equation 27 shows that the expected belief of an agent with similarity-weighted memory is constant over time as a result of the i.i.d.-structure of the economy. A constant expectation of the agent's posterior beliefs implies that belief revisions are predictable and mean-reverting. If this period's log endowment growth is very high, such that the relevant moments of the agent's posterior beliefs are inflated, then I expect to observe a downward revision of the agent's beliefs in the next period.

Summary. I find that the agent's posterior belief under similarity-weighted memory is time-varying, although the agent has access to infinite data and forms beliefs using infinite data. Consistent with empirical evidence, I find that (i) the agent's posterior mean varies procyclically and overreacts to new information; and that (ii) the agent's subjective volatility varies countercyclically. The beliefs are predictably mean-reverting.

3.3 Asset pricing implications

In this section, I examine the equilibrium asset pricing implications of the agent's subjective long-term beliefs under similarity-weighted memory, which are time-varying with contemporaneous log endowment growth. I first analyze subjectively expected asset prices and then discuss realized asset prices.

Subjective asset prices. Proposition 6 characterizes the equilibrium asset prices under similarity-weighted memory using the subjective cumulant-generating function in Equation 14 and the results from Section 2.3. I focus on power utility for simplicity and give the results for Epstein and Zin (1989)-preferences in Appendix B.5.

Proposition 6 (Asset prices under similarity-weighted memory)**.** *Focus on power utility ($\psi = 1/\gamma$). Under similarity-weighted memory as in Equation 11 and an i.i.d. two-state Markov-switching process for log endowment growth, it is*

$$r_t^f = -\log(\beta) - \log\left(\pi_1\, e^{-\gamma\,\hat{\mu}_{1,t}+\frac{1}{2}\gamma^2\,\sigma_1^2} + \pi_2\, e^{-\gamma\,\hat{\mu}_{2,t}+\frac{1}{2}\,\gamma^2\,\sigma_2^2}\right), \tag{28}$$

$$dp_t = -\log(\beta) - \log\left(\pi_1\, e^{(\lambda-\gamma)\,\hat{\mu}_{1,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_1^2} + \pi_2\, e^{(\lambda-\gamma)\,\hat{\mu}_{2,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_2^2}\right), \tag{29}$$

$$\tilde{er}_t = dp_t + \log\left(\pi_1\, e^{\lambda\,\hat{\mu}_{1,t}+\frac{1}{2}\,\lambda^2\,\sigma_1^2} + \pi_2\, e^{\lambda\,\hat{\mu}_{2,t}+\frac{1}{2}\,\lambda^2\,\sigma_2^2}\right), \tag{30}$$

$$\begin{aligned}\tilde{rp}_t = {} & \log\left(\pi_1\, e^{-\gamma\,\hat{\mu}_{1,t}+\frac{1}{2}\gamma^2\,\sigma_1^2} + \pi_2\, e^{-\gamma\,\hat{\mu}_{2,t}+\frac{1}{2}\,\gamma^2\,\sigma_2^2}\right) + \log\left(\pi_1\, e^{\lambda\,\hat{\mu}_{1,t}+\frac{1}{2}\,\lambda^2\,\sigma_1^2} + \pi_2\, e^{\lambda\,\hat{\mu}_{2,t}+\frac{1}{2}\,\lambda^2\,\sigma_2^2}\right) \\ & - \log\left(\pi_1\, e^{(\lambda-\gamma)\,\hat{\mu}_{1,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_1^2} + \pi_2\, e^{(\lambda-\gamma)\,\hat{\mu}_{2,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_2^2}\right).\end{aligned} \tag{31}$$

First, the risk-free rate r_t^f, which is determined by the agent's subjective belief in equilibrium, increases in the contemporaneous log endowment growth due to the procyclicality of the agent's expectation. The posterior state-wise means $\hat{\mu}_{1,t}$ and $\hat{\mu}_{2,t}$ both increase in contemporaneous log endowment growth g_t, so that the cumulant-generating function at $-\gamma$, $\mathcal{K}_t(-\gamma) = \log\left(\pi_1\, e^{-\gamma\,\hat{\mu}_{1,t}+\frac{1}{2}\gamma^2\,\sigma_1^2} + \pi_2\, e^{-\gamma\,\hat{\mu}_{2,t}+\frac{1}{2}\,\gamma^2\,\sigma_2^2}\right)$, decreases in g_t. However, $\mathcal{K}_t(-\gamma)$ enters the expression for the risk-free rate negatively, and the risk-free rate increases in g_t. Intuitively, if contemporaneous log endowment growth is high, the agent selectively recalls past experiences with a high log endowment growth due to similarity. Thus, the agent becomes too optimistic and expects a high endowment growth going forward. The risk-free rate must then be high to make saving in a risk-free asset attractive, as is consistent with evidence (Adam and Nagel, 2023).

Second, the dividend-price ratio of the λ-asset dp_t decreases in the contemporaneous log endowment growth if $\lambda > \gamma$. As discussed in Section 2.3, the price of the asset increases in the agent's posterior mean if the leverage factor λ exceeds the curvature of the utility function as determined by the risk-aversion parameter γ.[21] A decreasing dividend-price ratio is consistent with evidence showing that the price-dividend ratio (the reciprocal of dp_t) is positively correlated with subjective expectations of future growth (De La O and Myers, 2021; Bordalo et al., 2023b).

Third, the subjectively expected return, which depends on the dividend-price ratio dp_t and on the subjectively expected payout growth $\mathcal{K}_t(\lambda) = \log\left(\pi_1\, e^{\lambda\, \hat{\mu}_{1,t} + \frac{1}{2}\lambda^2 \sigma_1^2} + \pi_2\, e^{\lambda\, \hat{\mu}_{2,t} + \frac{1}{2}\lambda^2 \sigma_2^2}\right)$, increases in the contemporaneous log endowment growth. The subjectively expected payout growth is increasing in the agent's posterior mean and thus in the contemporaneous log endowment growth. For $\gamma > \lambda$, the dividend-price ratio increases in contemporaneous log endowment growth, such that both components of the subjectively expected return are increasing in contemporaneous log endowment growth. For $\gamma < \lambda$, instead, the dividend-price ratio decreases in contemporaneous log endowment growth. However, $\lambda > \gamma > 0$ implies that the expected payout growth is more sensitive to the agent's state-wise posterior means $\hat{\mu}_{1,t}$ and $\hat{\mu}_{2,t}$ than the dividend-price ratio, such that the increase of the expected payout dominates the decreasing dp_t. Intuitively, the agent is more optimistic after observing a high contemporaneous log endowment growth, and the risky asset must deliver a high expected return to induce investment.[22] Consistent with survey evidence (Amromin and Sharpe, 2014; Greenwood and Shleifer, 2014), the expected return of an agent with similarity-weighted memory is procyclical.

Fourth, the subjective risk premium rp_t depends on the convexity of the cumulant-generating function over the interval $[-\gamma, \lambda]$, as highlighted by Martin (2013). Under similarity-weighted memory, however, the convexity of the cumulant-generating function

[21]In contrast, the price of the asset decreases in the agent's posterior mean if $\gamma > \lambda$ under power utility preferences, but not under Epstein and Zin (1989)-preferences.

[22]Figure D.2 in Appendix D shows that an increase in contemporaneous log endowment growth rotates the subjective cumulant-generating function. As the cumulant-generating function is convex (Martin, 2013), the effect of an increase in the contemporaneous log endowment growth on the subjective expected return is not necessarily monotone, and Figure 1 below shows that the expected return is a convex function of g_t. The statements in the main text hold for a second-order approximation of the subjective cumulant-generating function.

changes over time with g_t. Consider the second-order approximation of the cumulant-generating function under the agent's subjective beliefs discussed in Section 2.3, $rp_t \approx \lambda\,\gamma\,\mathrm{Var}_t\left(g_{t+1}\right) = \lambda\,\gamma\,\left(\pi_1\,\sigma_1^2 + \pi_2\,\sigma_2^2 + \pi_1\,\pi_2\,\left(\hat{\mu}_{1,t} - \hat{\mu}_{2,t}\right)^2\right)$.[23] The agent knows the state-wise variances σ_1^2 and σ_2^2, but the wedge between the posterior mean of states 1 and 2 is time-varying with the contemporaneous log endowment growth. The agent's posterior variance is a convex function of the contemporaneous log endowment growth. The subjective risk premium of the agent is, up to a second-order approximation of the subjective cumulant-generating function, proportional to the unconditional posterior variance, and thus a convex function of the contemporaneous log endowment growth. The agent requires a high risk premium in exceptionally good and bad times, while the risk premium is moderate in intermediate regions of log endowment growth.

Figure 1 displays the risk-free rate, dividend-price ratio, expected return, and subjective risk premium. In contrast to the discussion so far, I consider a specification of the Epstein and Zin (1989)-preferences with $\psi \neq 1/\gamma$. The qualitative properties of asset prices highlighted for power utility continue to hold.Under similarity-weighted memory, the risk-free rate and the price-dividend ratio increase in the contemporaneous log endowment growth g_t, while the subjectively expected return is a convex function of g_t due to the effect of higher-order moments (see Footnote 22). Similarly, the subjective risk premium of the agent is a convex function of the contemporaneous log endowment growth due to the convexity of the posterior variance. In contrast, the risk-free rate, price-dividend ratio, subjective expected return, and subjective risk premium are constant in the i.i.d. economy under rational expectations, as shown by the dashed lines.

Objective asset prices. As noted earlier, the agent's beliefs are predictable by an outside observer with access to the same data as the agent. The objectively realized return and the objectively realized risk premium therefore deviate from their subjective counterpart. Using the Campbell-Shiller decomposition as in Campbell (1991), the objective risk premium

[23]Note that the subjective cumulant-generating function $\mathcal{K}_t(k)$ depends on the state-wise posterior means, $\hat{\mu}_{1,t}$ and $\hat{\mu}_{2,t}$. Therefore, the state-wise posterior means enter the expression for the risk premium.

Figure 1: Qualitative asset pricing implications of similarity-weighted memory

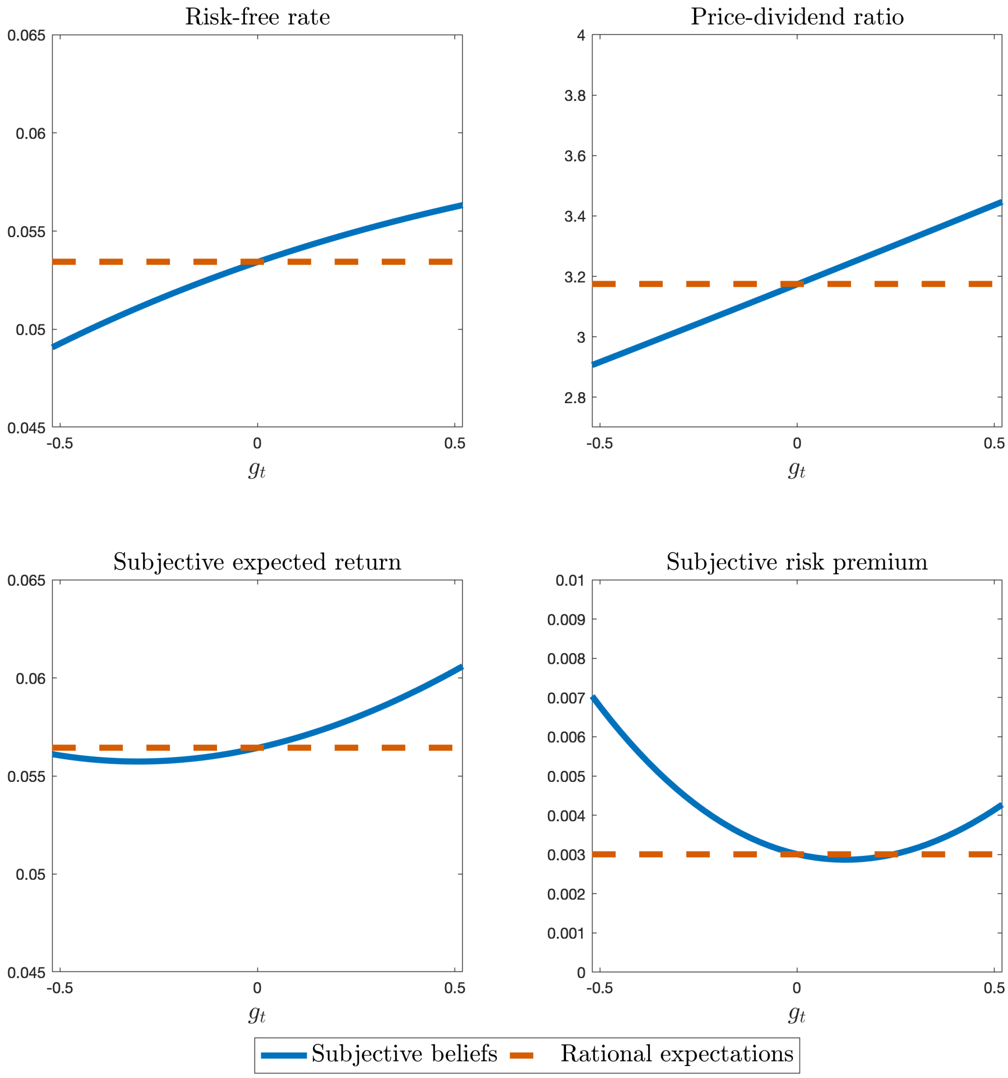

Figure 1 shows the risk-free rate, price-dividend ratio, subjectively expected return and subjectively expected risk premium for different values of contemporaneous log endowment growth g_t under rational expectations (dashed line) and under similarity-weighted memory (solid line). The parameters of the log endowment growth process are as in Table 1, and the preference parameters are $\beta = 0.95, \gamma = 10, \psi = 1.5, \kappa = 0.01$.

is (see Appendix F)

$$\mathbb{E}\left(r_{t+1}\right) - r_t^f = \lambda\, \mathbb{E}(g_{t+1}) + \mathcal{K}_t(-\gamma) - \mathcal{K}_t(\lambda - \gamma) - \frac{\bar{p}}{1-\bar{p}} \left(\mathbb{E}\left(dp_{t+1}\right) - dp_t\right), \tag{32}$$

with

$$\mathbb{E}(dp_{t+1}) = -\log(\beta) - \mathbb{E}\left[\mathcal{K}_{t+1}(\lambda-\gamma)\right] + \left(1-\frac{1}{\eta}\right)\mathbb{E}\left[\mathcal{K}_{t+1}(1-\gamma)\right],$$

and $\bar{p} = \frac{1}{1+\exp(\overline{d-p})}$ with a historical average dividend-price ratio $\overline{d-p}$ of 4% to 5%, implying $\bar{p} \approx 0.95$ (Campbell, 2017). The objective risk premium deviates from the subjective risk premium due to two effects: First, the econometrician's expected endowment growth $\mathbb{E}(g_{t+1}) = \pi_1\,\mu_1 + \pi_2\,\mu_2$ generally deviates from the subjectively expected endowment growth $\tilde{\mathbb{E}}_{\mathrm{t}}(g_{t+1}) = \pi_1\,\hat{\mu}_{1,t} + \pi_2\,\hat{\mu}_{2,t}$. Second, the econometrician expects a revision of the agent's beliefs to their long-term mean, which affects the dividend-price ratio of the economy. For $\bar{p} \approx 0.95$, differences in the expected dividend-price ratio under the objective and subjective measure are multiplied by $\frac{\bar{p}}{1-\bar{p}} \approx 19$, leading to large fluctuations in the objective risk premium.

To gain intuition, consider the case of log-normal endowment growth ($\mu = \mu_1 = \mu_2$ and $\sigma = \sigma_1 = \sigma_2$), since closed-form solutions exist in this case.[24] Under log-normality, it is

$$\mathbb{E}\left(r_{t+1}\right) - r_t^f = rp_t + \left(\frac{1}{1-\bar{p}}\,\lambda - \frac{\bar{p}}{1-\bar{p}}\,\frac{1}{\psi}\right)(\mu - \hat{\mu}_t) - \frac{1}{2}\,\lambda^2\,\sigma^2.$$

Similar to Nagel and Xu (2022), the objective risk premium has three components: The first term is the subjective risk premium, which is the risk compensation required by the representative agent in equilibrium. The second term is the time-varying belief wedge $\mu - \mu_t$, whose effect depends on the leverage of the asset λ and the inverse of the EIS ψ^{-1}. If the agent becomes too optimistic, $\mu_t > \mu$, the econometrician expects a mean reversion of the agent's beliefs towards μ and therefore a negative surprise in the next period. When the agent is overly optimistic, the return next period tends to be low. The last term is an adjustment that arises due to the difference between the agent's subjective beliefs and the

[24] I demonstrate how to approximate the expected subjective cumulant-generating function for the two-state Markov-switching process in Appendix F.

econometrician's objective beliefs.

Figure 2: Objective and subjective risk premium

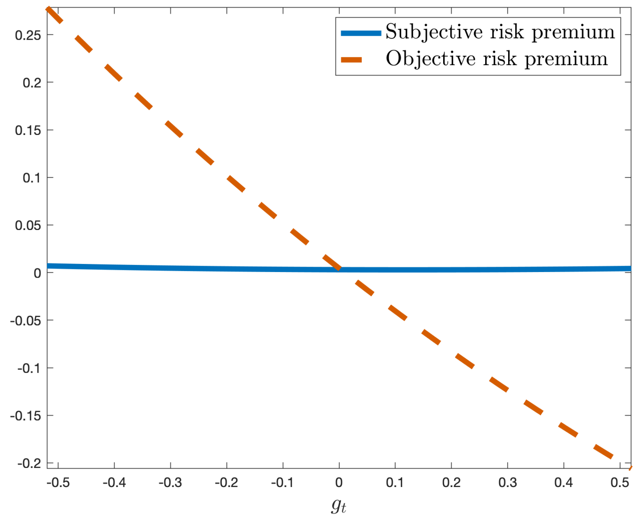

Figure 2 shows the objective (orange dashed line) and subjective (blue solid line) risk premium for different realizations of the contemporaneous log endowment growth g_t. The parameters of the log endowment growth process are as in Table 1, and the preference parameters are $\beta = 0.95, \gamma = 10, \psi = 1.5, \kappa = 0.01$.

Figure 2 shows the subjective and objective risk premium for the i.i.d. Markov-switching process, and shows that the intuition from the log-normal case discussed above extends to this case: When the contemporaneous log endowment growth is high, the agent becomes very optimistic and the realized risk premium in the next period will, on average, be low. Thus, the objective risk premium is predictably countercyclical. Moreover, Equation 32 shows that the price-dividend ratio (the reciprocal of dp_t) negatively predicts the realized risk premium. Figure 2 shows that the objective risk premium is low if g_t is high, which corresponds to a high price-dividend ratio as in the top right panel in Figure 1. The objective risk premium is thus predictably countercyclical using aggregate valuation ratios such as the price-dividend ratio (Campbell and Shiller, 1988). Additionally, Figure 2 indicates that the subjective risk premium varies significantly less than the objective risk premium. The subjective risk premium appears to be a straight line in Figure 2 because the objective risk premium is two orders of magnitude more volatile than the subjective risk premium (Nagel and Xu, 2023).

Summary. Similarity-weighted memory explains salient empirical differences of subjectively expected and objectively realized asset prices, especially patterns of cyclicality, predictability, and sensitivity to risk measures. First, subjectively expected returns are procyclical, while objectively realized returns are countercyclical. Second, the subjective risk premium is acyclical and not predictable by aggregate valuation ratios, while the objective risk premium is predictable by aggregate valuation ratios. Third, variation in the subjectively perceived riskiness of the economy leads to variation in the subjective risk premium, while variation of the objective risk premium is unrelated to the constant objective risk or risk-aversion.

3.4 Calibration

I next use simulations to analyze the model quantitatively. First, I describe the parameters of the log endowment growth process and the agent's preferences. I then simulate asset prices for two cases. In the first case, I assume that the agent has access to an infinte history H_t. In the second case, I assume that the agent has access to 30 years of data before "entering" the market, such that the agent learns from a limited sample.

Parameters. There are two types of parameters: endowment growth parameters and preference/belief parameters, and I summarize the parameters in Table 1. I estimate the parameters of the i.i.d. Markov-switching process for log endowment growth in Equation 10 using the methods in Johannes et al. (2016), and obtain estimates that are close to their results. I measure endowment growth using the quarterly growth of services and non-durable consumption from the Bureau of Economic Analysis from Q1 1947 to Q1 2023 and estimate the parameters with a Markov-Chain-Monte-Carlo (MCMC) method (for details, see Appendix G.1). The parameters reported in Table 1 are the *annualized* Bayesian maximum a-posterior parameters among $10,000$ parameter combinations. Furthermore, I estimate the leverage parameter λ by regressing quarterly aggregate dividends obtained from CRSP on endowment growth and find $\lambda \approx 3.29$, which is close to $\lambda = 3$ as used in Collin-Dufresne et al. (2016) and Nagel and Xu (2022). I use $\lambda = 3$ for comparability.

Table 1: Calibration parameters

Parameter	Symbol	Value	Source
Endowment growth process			
Mean growth in			
State 1	μ_1	1.96%	Estimated
State 2	μ_2	0.64%	Estimated
Growth volatility in			
State 1	σ_1	0.90%	Estimated
State 2	σ_2	4.76%	Estimated
Probability of state 1	π_1	0.86	Estimated
Leverage	λ	3.00	Nagel and Xu (2022)
Preferences and memory			
Risk aversion	γ	10	Jin and Sui (2022)
EIS	ψ	1.5	Nagel and Xu (2022)
Time discount factor	β	0.9967	Nagel and Xu (2022)
Memory scrutiny	κ	0.01	-

Table 1 reports the parameters used in the simulation. The parameters of the endowment growth process are estimated using the methods in Johannes et al. (2016) and annualized by multiplying means by four and standard deviations by 2. For the preference parameters, most values are the same as in Jin and Sui (2022) and Nagel and Xu (2022). Memory scrutiny κ is set to be of the same magnitude as the volatility of endowment growth.

For the preference parameters, I mostly follow previous papers and set γ, the relative risk-aversion coefficient, to ten (Jin and Sui, 2022).[25] Next, I set ψ, the elasticity of intertemporal substitution (EIS), to 1.5. There is a range of different values for the EIS in the asset pricing literature, but the majority of papers uses values above one (Beeler and Campbell, 2012).[26] Finally, I set β, the discount factor, to 0.9967 as in Nagel and Xu (2022), and κ, the scrutiny parameter to 0.01, so that memory scrutiny is of the same order of magnitude as the volatility of endowment growth. Assuming $\kappa = 0.01$ implies that $\alpha_1 = 0.002$ and $\alpha_2 = 0.053$, such that the agent's posterior mean is not too sensitive to g_t.[27]

[25]The long-run risk literature often assigns value of up to ten to the relative risk-aversion coefficient, and Mehra and Prescott (1985) argue that values up to and around ten are reasonable. Lowering the relative risk-aversion to $\gamma = 4$ (Nagel and Xu, 2022) yields a subjective risk premium of approximately 0.5 in the case without parameter uncertainty shown in Table 2, but otherwise does not affect the results qualitatively.

[26]Lowering ψ to one yields a higher risk-free rate and subjectively expected return, but leaves the subjective risk premium in the case without parameter uncertainty unchanged. Setting $\psi > 1$ simplifies the numerical calibrations of the case with parameter uncertainty.

[27]For $\kappa = 0.1\,(0.001, 0.0001)$, it is $\alpha_1 = 0.0002\,(0.0198, 0.1684)$ and $\alpha_2 = 0.0055\,(0.3577, 0.8478)$. The parameter κ does not affect the qualitative implications, but has quantitative implications. If the agent's

Simulation. I simulate the model at a quarterly frequency. Table 2 reports the annualized average moments from $10,000$ simulations of the model, each for 304 quarters. The left panel shows that the model can qualitatively match salient facts about asset prices.

Figure 3 shows the posterior beliefs from one realization of the endowment growth process. In combination with Table 2, Figure 3 shows that the agent's posterior mean of the good state $\hat{\mu}_{1,t}$, which is the state that occurs most often, is unbiased, while the posterior mean of the recession state is biased upward and fluctuates more strongly than $\hat{\mu}_{1,t}$. Furthermore, the posterior variance of the agent in the right panel of Figure 3 tends to spike during recessions (the vertical lines) and is on average slightly higher during recessions than during normal times. Subjective volatility is negatively correlated with endowment growth, such that the agent perceives a more risky economy in bad times than in good times. The bottom row in Table 3 shows the coefficient of a Coibion and Gorodnichenko (2015)-regression as in Proposition 5 and indicates a marginal overreaction of the agent's beliefs in that $b_{CG} < 0$.[28]

The left panel of Table 2 shows that the risk-free rate and the subjectively expected return are procyclical (positive correlation with endowment growth). Furthermore, the subjective risk premium is almost constant (average standard deviation of 0.005%), does not vary across normal times (average of 1.199%) or recessions (average of 1.204%), and is not predictable by the dividend-price ratio (Table 3). Consistent with empirical evidence, the subjective risk premium is acyclical with respect to aggregate valuation ratios (Adam and Nagel, 2023). However, because the subjective volatility changes over time, the subjective risk premium is correlated with the contemporaneous log endowment growth under similarity-weighted memory, and Nagel and Xu (2023) find evidence of a positive risk-return trade-off in the subjective risk premium.

beliefs react too strongly to g_t, an equilibrium might not exist as the agent might be too optimistic and does not invest in the risky asset.

[28]Using conventional t-Statistics, I cannot reject the null that $b_{CG} = 0$ in the sample. In unreported results, I confirm that increasing the sample size leads to $b_{CG} \to -0.5$ with high t-Statistics. Moreover, Mincer-Zarnowitz regressions yield $b_{MZ} \approx 0$, and I cannot reject the null of $b_{MZ} = 0$ using conventional t-Statistics.

Table 2: Simulation results with baseline parameters

Symbol	No parameter uncertainty			Parameter uncertainty		
	Total	Normal	Recession	Total	Normal	Recession
Endowment growth						
$\overline{g_t}$	1.774	1.961	0.626	1.777	1.960	0.652
$\text{Std}\,(g_t)$	1.958	0.901	4.725	1.958	0.900	4.719
Beliefs						
$\overline{\mu_t}$	1.783	1.960	0.700	1.782	1.959	0.697
$\text{Std}\,(\hat{\mu}_t)$	0.018	0.004	0.103	0.056	0.027	0.353
$\overline{\hat{\sigma}_t}$	1.966	1.965	1.968	1.985	1.984	1.990
$\text{corr}(\hat{\sigma}_t, g_t)$	−0.970	−0.999	−0.969	−0.319	−0.045	−0.001
Subjective asset prices						
$\overline{er_t}$	3.385	3.386	3.381	3.288	3.256	3.480
$\text{Std}\,(er_t)$	0.009	0.001	0.021	0.902	2.213	3.474
$\text{corr}\,(er_t, g_t)$	0.999	1.000	0.999	0.419	0.010	0.004
$\overline{r_t^f}$	2.186	2.187	2.178	2.174	2.180	2.136
$\text{Std}(r_t^f)$	0.027	0.012	0.065	0.686	0.848	0.902
$\text{corr}(r_t^f, g_t)$	0.999	1.000	0.999	0.497	0.069	0.326
$\overline{rp_t}$	1.199	1.199	1.204	1.114	1.076	1.344
$\text{Std}(rp_t)$	0.005	0.002	0.002	1.217	2.410	3.660
$\text{corr}(rp_t, g_t)$	−0.990	−1.000	−0.990	−0.381	−0.005	−0.057
Objective asset prices						
$\overline{rp_t}$	0.998	0.776	2.360	3.209	1.768	12.054
$\text{Std}(rp_t)$	2.237	1.027	5.397	26.864	42.188	69.590
$\text{corr}(rp_t, g_t)$	−1.000	−1.000	−1.000	−0.348	−0.112	−0.399

Table 2 reports the model moments obtained from $10,000$ simulations of the model for 304 quarters. The left panel (No parameter uncertainty) reports results under the assumption that the agent has an infinite sample of observations, while the right panel (Parameter uncertainty) reports results when the agent learns from a finite sample, such that the Bayesian posterior has a strictly positive variance around the parameter values. I use a burn-in period of 120 quarterly observations in the parameter uncertainty simulations. For each of the $10,000$ economies, I draw 10 realizations of the agent's memory and average them over the realizations. Returns and expectations are annualized as follows: the means are multiplied by four and the standard deviations are multiplied by two. For the risk-free rate, I multiply the quarterly mean and the standard deviation by four.

In contrast to the subjective risk premium, the objective risk premium varies more strongly (standard deviation of 2.237%), but is lower than the subjectively expected risk

Figure 3: Posterior long-term beliefs

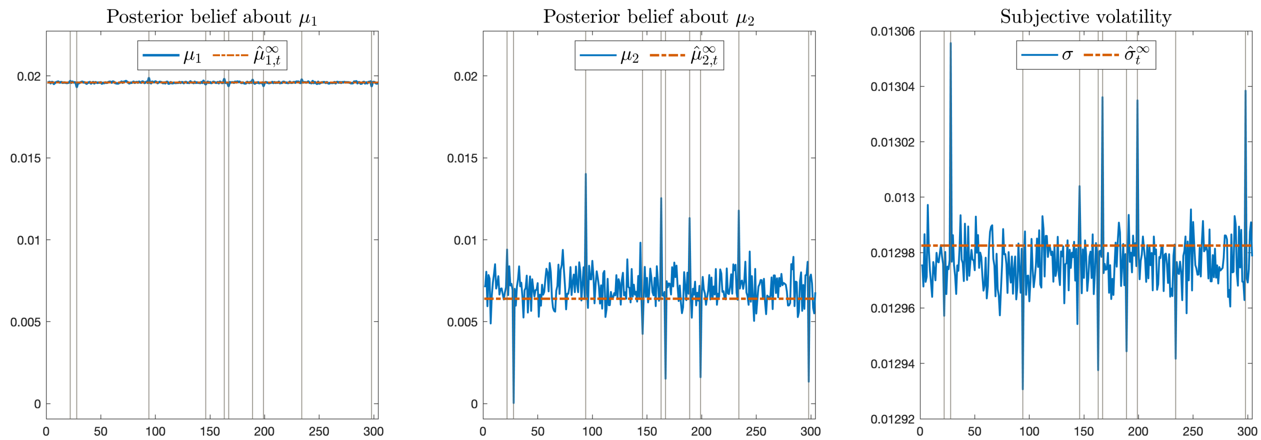

Figure 3 shows the annualized posterior beliefs of the agent for one realization of log endowment growth. The left (middle) panel shows the posterior mean of log endowment growth in the good state (bad state) and the right panel shows the subjective volatility. The dash-dotted line in each panel plots the respective quantity under full-information rational expectations, and the vertical lines mark realized bad states. The parameters are as in Table 1, except that I use $p_1 = 0.96$.

Table 3: Predictability and Coibion and Gorodnichenko (2015)-regressions

	No parameter uncertainty			Parameter uncertainty		
	RP_{Subj}	RP_{Obj}	$\hat{b}_{CG}$	RP_{Subj}	RP_{Obj}	$\hat{b}_{CG}$
dp_t	0.002	1.119		0.002	3.372	
$\left(\tilde{\mathbb{E}}_t - \mathbb{E}_{t-1}\right) g_{t+1}$			-0.026			-0.394

Table 3 reports the mean estimates from regressions for $10,000$ simulations of the model for 304 quarters plus a 120 quarters burn-in period in the parameter-uncertainty case. The first row shows the mean coefficients when regressing subjectively expected and objectively obtained risk premia on the log dividend-price ratio, as in Nagel and Xu (2023). The price-dividend ratio is rescaled to unit standard deviation. The second row shows the mean estimate from Coibion and Gorodnichenko (2015)-regressions of the forecast error on the forecast revision. The left (right) panel shows the results obtained without (with) parameter uncertainty.

premium with an average of 0.998%. The countercyclicality of the objective risk premium is more pronounced, since the objective risk premium is lower in normal times (0.007%) than in recessions (2.360%). Table 3 shows that the objective risk premium is predictably countercyclical using the dividend-price ratio ($\hat{b}_{obj} = 1.119$). The high volatility and predictability of the objective risk premium is consistent with empirical findings in Nagel and Xu (2023) and Campbell and Shiller (1988).

Table E.1 in Appendix E shows the average subjective and objective risk premium for each quantile of log endowment growth. The subjective risk premium is almost constant across quantiles, but the objective risk premium is high in the lowest quantile of endowment growth (6.42%), and becomes negative in the highest quantile (−3.91%), consistent with empirical results that find negative excess returns in times of high sentiment (Greenwood and Hanson, 2013; Cassella and Gulen, 2018).

Asset prices in an economy with similarity-weighted memory distortions are qualitatively consistent with stylized facts, but the objectively realized risk premium is comparably low. The results in Section 3.3 show that the objective risk premium depends on the predictability of the agent's subjective beliefs. Under the assumption that the agent already observed an infinitely long history of past log endowment growth, the agent's beliefs solely depend on next period's log endowment growth, which is very smooth (Mehra and Prescott, 1985).

Therefore, I next simulate asset prices when the agent learns from finitely many observations, which yields Bayesian parameter uncertainty. Collin-Dufresne et al. (2016) show that the parameter uncertainty emerging from Bayesian learning endogenously generates long-run risk and thus a high risk premium. I show that parameter uncertainty under similarity-weighted memory yields a realistically high risk premium and matches the aforementioned qualitative features of asset prices.

The agent's beliefs and asset prices must be simulated when the agent learns from finitely many observations (see Appendix G.2).[29] As before, I simulate the model $10,000$ times for 304 quarters. For each quarter and each simulation, I draw the agent's recalled observations ten times according to the similarity-weighted memory function (Equation 11). Throughout the simulations, I focus on the case in which the agent's prior is uninformative and generate 120 quarterly log endowment growth realizations as burn-in period, such that the agent has access to 30 years of data before "entering" the market.

Figure 4 plots one realization of the agent's beliefs and shows that the agent's beliefs

[29]We can obtain closed-form solutions for $\psi = 1$ and log-normal endowment growth $\mu_1 = \mu_2$ and $\sigma_1 = \sigma_2$, which I derive in Appendix G.2.

Figure 4: Posterior beliefs with parameter uncertainty

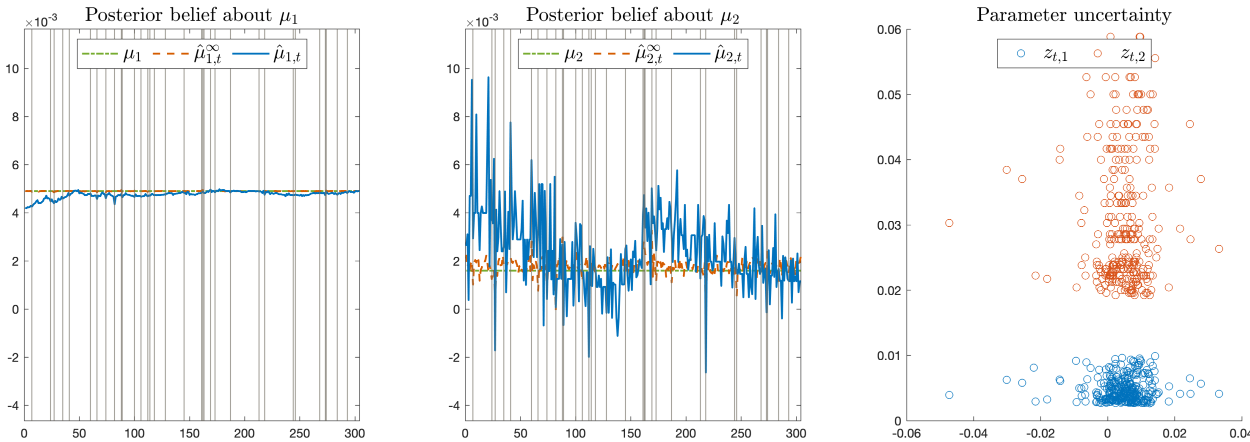

Figure 4 shows one time-series realization of the agent's posterior beliefs when allowing for parameter uncertainty. The left panel shows the beliefs about μ_1, and the middle panel shows the beliefs about μ_2. In both panels, the solid blue line shows the posterior with a limited number of observations, while the orange dotted line shows the beliefs in the $t \to \infty$-limit. The horizontal green dashed line plots the fundamentally true mean μ_s, and the vertical grey lines mark recession periods. The right panel plots the parameter uncertainty $z_{s,t}$ against the endowment growth g_t. I focus on an uninformative prior $\nu_s \to 0$. The simulation parameters are as in Table 1.

when learning from a limited sample are more volatile than in the $t \to \infty$-case, but the averages are close. Similarly, the properties of the subjective volatility discussed above hold under parameter uncertainty with a lower correlation of subjective volatility and log endowment growth. The bottom row of Table 3 shows that the agent's beliefs overreact more when learning from a limited sample (right panel), with a Coibion and Gorodnichenko (2015)-regression coefficient of $\hat{b}_{CG} = -0.419$.

The invariance of the qualitative features of beliefs when going from $t \to \infty$ to a limited-t implies that the qualitative features of asset prices discussed above hold under parameter uncertainty, but the quantitative fit is improved. Empirically, Jin and Sui (2022) report an average realized risk premium of 3.90%, while the average subjective risk premium is around 1.90%.[30] In the simulations, I obtain an average objective risk premium of 3.21%, which is considerably higher during recessions (12.05%) than during normal times (1.77%). The

[30] I measure the subjective risk premium as log expected return minus the log risk-free rate using the data provided by Nagel and Xu (2022), which combines several surveys. The data are available from the Review of Financial Studies.

subjective risk premium is lower than the objective risk premium at 1.11%, and does not vary much across normal times (1.08%) and recessions (1.34%). Similarly, the risk-free rate is within the range of empirical estimates with 2.17% (Jin and Sui, 2022; Nagel and Xu, 2022) and has a low volatility (Campbell and Cochrane, 1999).

Departing from rational expectations by considering subjective beliefs that arise from similarity-weighted memory explains the mismatch of subjective and objective asset prices regarding their cyclicality, predictability, and sensitivity to risk measures. In addition, subjective beliefs under similarity-weighted memory generate a realistically high risk premium and low risk-free rate.

4 Peak-end memory

I next demonstrate the flexibility of the general model introduced in Section 2 by analyzing another memory distortion. From now on, I assume that the agent is more likely to recall extreme observations as well as observations that are similar to today's observation, leading to a peak-end memory distortion (Kahneman, 2000). Psychologically, the peak-end memory distortion analyzed here is consistent with the literature on experience effects, which highlights the long-term influence of extreme experiences on risk taking (Malmendier and Nagel, 2011), inflation expectations (Malmendier and Nagel, 2016), managerial decisions (Malmendier et al., 2011), and real estate purchases (Happel et al., 2022). Emotional events are more likely to be stored in memory (Kensinger and Ford, 2020) and retrieved more vividly (flashbulb memories, see Phelps, 2006). Additionally, the end of an experience is generally more memorable than the beginning and middle (recency effect, Kahana, 2012; Barberis, 2018; Wachter and Kahana, 2023). I next outline the setup and discuss the implications of the peak-end memory distortion on the agent's beliefs and asset prices afterwards.

4.1 Framework

Consider log-normal endowment growth

$$g_t = \mu + \sigma\,\epsilon_t, \qquad \epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1)\,. \tag{33}$$

The agent knows that endowment growth is i.i.d. log-normally distributed, but learns the mean μ and volatility σ from her recalled history H_t^R. The agent's recalled history is distorted by the memory function

$$m^{\text{PE}}(g_\tau, g_t) := \underbrace{\exp\left[-e^{-\frac{(g_\tau-\mu)^2}{2\,\sigma^2}}\right]}_{\text{Extreme experience bias}} \cdot \underbrace{\exp\left[\frac{(g_\tau - g_t)^2}{2\,\kappa}\right]}_{\text{Similarity}}. \tag{34}$$

The extreme experience bias models the higher likelihood of recalling extreme observations. The inner exponential, $e^{-\frac{(g_\tau-\mu)^2}{2\,\sigma^2}}$, is proportional to the probability density function (pdf) of a normal distribution, with higher values close to the true mean μ and lower values in the tails of the distribution. The outer exponential around the pdf of the normal distribution implies that the center of the distribution—values around μ—are underweighted, while values in the tails of the distribution are overweighted.[31] The similarity term models the higher likelihood of recalling experiences at the end of the realized history H_t.[32] Figure D.3 plots the memory function.

4.2 Subjective long-term beliefs and asset prices

I next analyze the agent's subjective long-term beliefs under the peak-end memory function and highlight the implications for asset prices. Closed-form solutions for the agent's long-

[31] A more formal motivation for the functional form comes from the cumulative distribution function of the Gumbel distribution, which is the limiting distribution of the maximum of sequences of independent normal variables, see Appendix C.1.

[32] An alternative formulation could explicitly overweight experiences as a function of the passage of time, $t-\tau$. However, the proof of Proposition 1 in Section 2.2 uses the almost sure convergence of the empirical distribution of past experiences to the memory-weighted distribution under the law of large numbers, which does not hold if the agent deterministically forgets past experiences.

term belief do not exist under the memory function in Equation 34, such that I mostly discuss the impact of the extreme experience bias. The effect of the similarity term on the agent's beliefs and asset prices is as in Section 3.

Subjective beliefs. The agent is more likely to recall extreme experiences and experiences that are similar to the time-varying end of the realized history H_t. The extreme experience bias does not systematically affect the posterior mean of the agent since both tails of the distribution of log endowment growth are overweighted symmetrically, but an extreme experience bias leads to an increased posterior variance.[33] Appendix C.3 shows that a pure extreme experience bias yields $\hat{\mu}_t = \mu$ and $\hat{\sigma}_t^2 \approx 1.41 \cdot \sigma^2 > \sigma^2$.

Figure 5 shows the posterior mean and variance of the agent under the peak-end memory distortion given in Equation 34. I provide a numerical approximation of the agent's posterior mean in Appendix C.4. The agent's posterior mean is an increasing function of the contemporaneous log endowment growth g_t, since the posterior mean under the peak-end memory distortion varies only due to the similarity-term. The right panel of Figure 5 shows the posterior variance of the agent. Under peak-end memory, the posterior variance is higher than the true variance of the process due to the overweighting of extreme observations. Moreover, the posterior volatility is concave in contemporaneous log endowment growth with a maximum at $g_t = \mu$ due to the combination of extreme experience bias and similarity-weighting: If $g_t = \mu$, the memory function is identical to a pure overweighting of extreme experiences, and the posterior variance of the agent is $\hat{\sigma}_t^2 \approx 1.41 \cdot \sigma^2$. If $g_t \neq \mu$, the agent recalls more endowment growth rates that are close to g_t, shifting the mass of recalled experiences closer to one tail of the distribution, which reduces the posterior variance. The reduction is stronger the more the mass is shifted towards one tail of the distribution.

Asset pricing implications. I next analyze the effect of the peak-end memory distor-

[33] In line with Proposition 1, extreme experience bias does not induce a covariance between the propensity of recalling an experience and the value of log endowment growth, $\mathrm{Cov}\left[g, \mathbb{1}_{\{g \in H_t^R\}}\right] = 0$, but an extreme experience bias induces a positive covariance between the propensity of recalling an experience and the tails of the distribution, $\mathrm{Cov}\left[(g - \hat{\mu}_t)^2, \mathbb{1}_{\{g \in H_t^R\}}\right] > 0$.

Figure 5: Posterior beliefs under peak-end memory

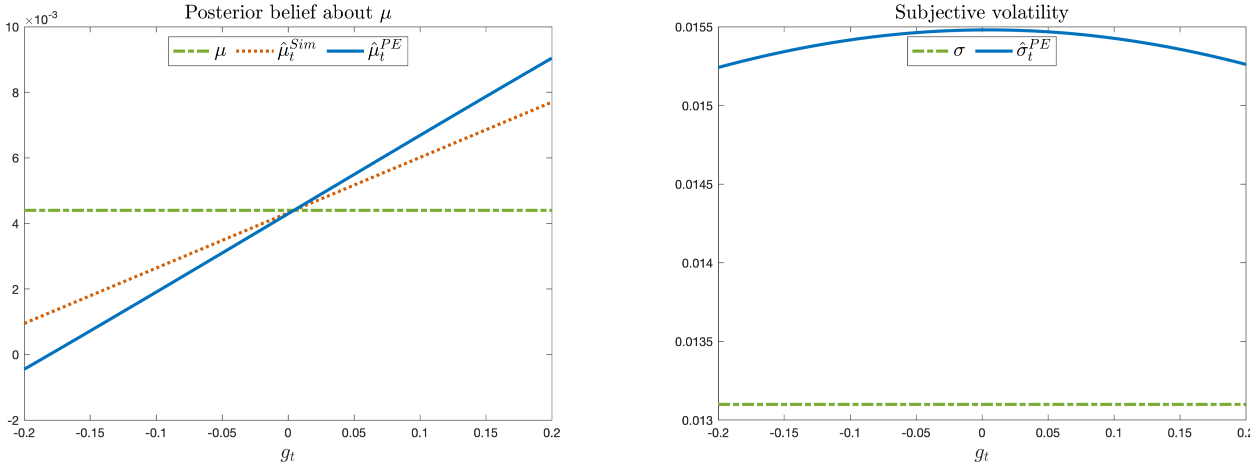

Figure 5 shows the posterior mean and variance of an agent with peak-end memory distortions as in Equation 34 for varying levels of contemporaneous endowment growth g_t. The parameters are $\mu = 0.44\%$, $\sigma = 1.31\%$, and $\kappa = 0.01$.

tion on asset prices using the framework in Section 2.3. The cumulant-generating function of log endowment growth under the agent's time-t belief is given by

$$\mathcal{K}_t^{PE}(k) = \log \tilde{\mathbb{E}}_t \left(e^{k\, g_{t+1}}\right) = k\, \hat{\mu}_t + \frac{1}{2}\, k^2\, \hat{\sigma}_t^2, \tag{35}$$

and I insert numerical estimates of $\hat{\mu}_t$ and $\hat{\sigma}_t^2$ to obtain subjective asset prices.

As in Section 3.4, I simulate the model $10,000$ times for 304 quarters and report average moments in Table 4. The parameters of the log endowment growth process are as in Nagel and Xu (2022) with a quarterly mean endowment growth of $\mu = 0.44\%$ (annualized: 1.76%) and a quarterly volatility of $\sigma = 1.31\%$ (annualized: 2.62%). All other parameters are as in Table 1.

The simulation results in Table 4 highlight that the agent's posterior mean is an unbiased estimate of the true mean, with an average posterior mean of 1.760%, and is relatively stable with an average standard deviation of 0.063%. Time-variation in the agent's posterior mean is entirely driven by the similarity component of the memory function in Equation 34, such that the agent's posterior mean is perfectly correlated with this period's endowment growth

Table 4: Asset prices under the peak-end memory distortion

Symbol	Mean	Std.	Corr. with g_t
Endowment growth and subjective beliefs			
g_t	1.759	2.620	1.000
$\hat{\mu}_t$	1.760	0.063	1.000
$\hat{\sigma}_t$	3.096	< 0.001	< 0.001
Subjective asset prices			
er_t	4.604	0.042	1.000
r_t^f	1.729	0.084	1.000
rp_t	2.875	< 0.001	< 0.001
Objective asset prices			
rp_t	2.573	19.063	-1.000

Table 4 reports the model moments obtained from $10,000$ simulations of the model for 304 quarters. I annualize the quantities as follows: Means are multiplied by four and the standard deviations are multiplied by two. For the risk-free rate, I multiply the quarterly mean and the standard deviation by four.

g_t, leading to procyclical beliefs. Table E.2 in Appendix E shows the results of a Coibion and Gorodnichenko (2015)-regression and highlights that the agent's beliefs overreact to new information as in Section 3.4. The higher likelihood of recalling extreme experiences leads to a higher posterior ($\hat{\sigma}_t = 3.096\%$) than fundamental volatility ($\sigma = 2.620\%$). The posterior volatility is very stable (standard deviation $< 0.001\%$) and not correlated with contemporaneous log endowment growth g_t.

The qualitative asset pricing implications of the peak-end memory distortion are as under similarity-weighted memory discussed in Section 3, because the time-variation in the agent's subjective long-term beliefs is due to the similarity component. The extreme experience bias inherent in the peak-end rule, however, affects the subjective risk premium. The agent is more likely to recall extreme past endowment growth observations, such that the agent perceives the economy as very risky. Being risk-averse, the agent must be compensated for holding a risky asset, and the extreme experience bias thus leads to a high subjective risk premium. Moreover, since the posterior volatility of the agent is very stable (standard deviation < 0.001), the subjective risk-premium is almost constant and not predictable using aggregate valuation ratios (see Table E.2).

The analysis of the peak-end memory distortion in this Section suggests that selective memory can not only explain qualitative features of beliefs and asset prices as in Section 3, but can parsimoniously generate a realistically high subjective and objective risk premium. A large literature on experience effects as summarized in Malmendier and Wachter (2022) and on the memorability of emotional events (Kensinger and Ford, 2020) argues that extreme experience leave "scars" that have a long-lasting effect on beliefs. The model here shows that the same memory mechanism—the higher likelihood of recalling extreme experiences—can explain why financial markets are perceived as risky and thus generate a high risk premium.

5 Conclusion

In this paper, I explore the implications of selective memory for asset prices and show that similarity-weighted selective memory simultaneously accounts for important facts about belief formation, survey data, and asset prices. With i.i.d. fundamentals and constant risk-aversion, my model explains empirically observed discrepancies between subjectively expected and objectively realized returns using a simple mechanism: The agent selectively recalls past observations of the fundamental that are similar to its contemporaneous realization. A good realization of the fundamental causes the agent to become too optimistic and to expect a high future growth with a low volatility, leading to a high expected return. The agent's optimism implies that asset prices today rise too much, and returns following a good realization of the fundamental are predictably low, although the fundamental risk in the economy and the agent's risk aversion are constant.

My framework can be used to analyze further selective memory distortions, and I demonstrate this flexibility by considering a peak-end memory distortion under which the agent additionally overremembers extreme past observations, consistent with the experience effects literature. Under the peak-end memory distortion, the subjective risk premium is high because the agent perceives the economy as risky. The framework could also be applied to active learning problems in a portfolio choice context (Gödker et al., 2022; Fudenberg et al., 2023), to agents who access false memories, or to heterogeneous agents.

This paper complements empirical evidence from finance (Bordalo et al., 2023b; Jiang et al., 2023; Nagel and Xu, 2023), experiments (Burro et al., 2023; Enke et al., 2023), and cognitive psychology (Kahana, 2012) by showing that selective—and especially similarity-weighted—memory theoretically explains conceptually puzzling patterns of subjective and objective asset prices without resorting to time-varying risk-aversion, long-term risk, or disaster risk. The emerging pattern suggests that selective memory systematically affects aggregate economic outcomes, such as asset prices. Understanding the role of memory in the formation of subjective beliefs could help researchers, practitioners, and policy makers to make sense of aggregate asset prices.

References

Abel, A. B. (1999). Risk premia and term premia in general equilibrium. *Journal of Monetary Economics*, *43*(1), 3–33.

Adam, K., Marcet, A., & Beutel, J. (2017). Stock price booms and expected capital gains. *American Economic Review*, *107*(8), 2352–2408.

Adam, K., & Nagel, S. (2023). Expectations Data in Asset Pricing. In R. Bachmann, G. Topa, & W. van der Klaauw (Eds.), *Handbook of Economic Expectations* (pp. 477–506). Academic Press.

Amromin, G., & Sharpe, S. A. (2014). From the horse's mouth: Economic conditions and investor expectations of risk and return. *Management Science*, *60*(4), 845–866.

Azeredo da Silveira, R., & Woodford, M. (2019). Noisy memory and over-reaction to news. *AEA Papers and Proceedings*, *109*, 557–561.

Bacchetta, P., Mertens, E., & Van Wincoop, E. (2009). Predictability in financial markets: What do survey expectations tell us? *Journal of International Money and Finance*, *28*(3), 406–426.

Barberis, N. (2018). Psychology-based models of asset prices and trading volume. In D. B. Bernheim, S. DellaVigna, & D. Laibson (Eds.), *Handbook of Behavioral Economics* (Vol. 1, pp. 79–175). North Holland, Amsterdam.

Barberis, N., Greenwood, R., Jin, L., & Shleifer, A. (2015). X-CAPM: An extrapolative capital asset pricing model. *Journal of Financial Economics*, *115*(1), 1–24.

Barberis, N., Greenwood, R., Jin, L., & Shleifer, A. (2018). Extrapolation and bubbles. *Journal of Financial Economics*, *129*(2), 203–227.

Barro, R. J. (2006). Rare disasters and asset markets in the twentieth century. *Quarterly Journal of Economics*, *121*(3), 823–866.

Bastianello, F., & Fontanier, P. (2022). Expectations and learning from prices. Working paper available at SRRN.

Beeler, J., & Campbell, J. Y. (2012). The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment. *Critical Finance Review*, *1*(1), 141–182.

Berk, R. H. (1966). Limiting behavior of posterior distributions when the model is incorrect. *Annals of Mathematical Statistics*, *37*(1), 51–58.

Bodoh-Creed, A. L. (2020). Mood, memory, and the evaluation of asset prices. *Review of Finance*, *24*(1), 227–262.

Bordalo, P., Conlon, J. J., Gennaioli, N., Kwon, S. Y., & Shleifer, A. (2023a). Memory and probability. *Quarterly Journal of Economics*, *138*(1), 265–311.

Bordalo, P., Gennaioli, N., La Porta, R., & Shleifer, A. (2019). Diagnostic expectations and stock returns. *Journal of Finance*, *74*(6), 2839–2874.

Bordalo, P., Gennaioli, N., La Porta, R., & Shleifer, A. (2023b). Belief overreaction and stock market puzzles. *Journal of Political Economy*, (Forthcoming).

Bordalo, P., Gennaioli, N., Ma, Y., & Shleifer, A. (2020a). Overreaction in macroeconomic expectations. *American Economic Review*, *110*(9), 2748–2782.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2018). Diagnostic expectations and credit cycles. *Journal of Finance*, *73*(1), 199–227.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2020b). Memory, attention, and choice. *Quarterly Journal of Economics*, *135*(3), 1399–1442.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2022). Overreaction and diagnostic expectations in macroeconomics. *Journal of Economic Perspectives*, *36*(3), 223–244.

Burro, G., Castagnetti, A., Cillo, A., & Crespi, P. G. (2023). A Walk Down Memory Lane: How Memories Influence Stock Investment and Information Processing. Working paper.

Campbell, J. Y. (1986). Bond and stock returns in a simple exchange model. *Quarterly Journal of Economics*, *101*(4), 785–803.

Campbell, J. Y. (1991). A Variance Decomposition for Stock Returns. *Economic Journal*, *101*(405), 157–179.

Campbell, J. Y. (2017). *Financial Decisions and Markets: A Course in Asset Pricing.* Princeton University Press.

Campbell, J. Y., & Cochrane, J. H. (1999). By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, *107*(2), 205–251.

Campbell, J. Y., & Shiller, R. (1988). The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors. *Review of Financial Studies*, *1*(3), 195–228.

Carr, H. A. (1931). The Laws of Association. *Psychological Review*, *38*(3), 212.

Cassella, S., & Gulen, H. (2018). Extrapolation bias and the predictability of stock returns by price-scaled variables. *Review of Financial Studies*, *31*(11), 4345–4397.

Charles, C. (2021). Memory and trading. Working paper available at SRRN.

Charles, C. (2022). Memory Moves Markets. Working paper available at SRRN.

Coibion, O., & Gorodnichenko, Y. (2015). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review*, *105*(8), 2644–2678.

Collin-Dufresne, P., Johannes, M., & Lochstoer, L. A. (2016). Parameter learning in general equilibrium: The asset pricing implications. *American Economic Review*, *106*(3), 664–98.

Cruciani, F., Berardi, A., Cabib, S., & Conversi, D. (2011). Positive and negative emotional arousal increases duration of memory traces: Common and independent mechanisms. *Frontiers in Behavioral Neuroscience*, *5*, 86.

d'Acremont, M., Schultz, W., & Bossaerts, P. (2013). The human brain encodes event frequencies while forming subjective beliefs. *Journal of Neuroscience*, *33*(26), 10887–10897.

Da, Z., Huang, X., & Jin, L. J. (2021). Extrapolative beliefs in the cross-section: What can we learn from the crowds? *Journal of Financial Economics*, *140*(1), 175–196.

De La O, R., & Myers, S. (2021). Subjective cash flow and discount rate expectations. *Journal of Finance*, *76*(3), 1339–1387.

Ebbinghaus, H. (1885). *Ueber das Gedaechtnis*. Duncker & Humblot.

Ehling, P., Graniero, A., & Heyerdahl-Larsen, C. (2018). Asset prices and portfolio choice with learning from experience. *Review of Economic Studies*, *85*(3), 1752–1780.

Enke, B., Schwerter, F., & Zimmermann, F. (2023). Associative Memory, Beliefs and Market Interactions. Working paper.

Epstein, L. G., & Zin, S. E. (1989). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica*, *57*(4), 937–969.

Esponda, I., & Pouzo, D. (2016). Berk–Nash equilibrium: A framework for modeling agents with misspecified models. *Econometrica*, *84*(3), 1093–1130.

Fudenberg, D., Lanzani, G., & Strack, P. (2023). Selective Memory Equilibrium. Working paper available at SRRN.

Gilboa, I., & Schmeidler, D. (1995). Case-based decision theory. *Quarterly Journal of Economics*, *110*(3), 605–639.

Gödker, K., Jiao, P., & Smeets, P. (2022). Investor memory. Working paper available at SRRN.

Goetzmann, W. N., Watanabe, A., & Watanabe, M. (2022). Evidence on Retrieved Context: How History Matters. NBER Working Paper.

Graeber, T., Zimmermann, F., & Roth, C. (2022). Stories, statistics, and memory. Working paper available at SRRN.

Greenwood, R., & Hanson, S. G. (2013). Issuer quality and corporate bond returns. *Review of Financial Studies*, *26*(6), 1483–1525.

Greenwood, R., & Shleifer, A. (2014). Expectations of returns and expected returns. *Review of Financial Studies*, *27*(3), 714–746.

Hansen, L. P., Heaton, J. C., & Li, N. (2008). Consumption strikes back? Measuring long-run risk. *Journal of Political Economy*, *116*(2), 260–302.

Happel, J., Karabulut, Y., Schäfer, L., & Tuzel, S. (2022). Shattered Housing. Working paper available at SRRN.

Heidhues, P., Kőszegi, B., & Strack, P. (2021). Convergence in models of misspecified learning. *Theoretical Economics*, *16*(1), 73–99.

Hirshleifer, D., Li, J., & Yu, J. (2015). Asset pricing in production economies with extrapolative expectations. *Journal of Monetary Economics*, *76*, 87–106.

Jiang, Z., Liu, H., Peng, C., & Yan, H. (2023). Investor Memory and Biased Beliefs: Evidence from the Field. Working paper.

Jin, L. J., & Sui, P. (2022). Asset pricing with return extrapolation. *Journal of Financial Economics*, *145*(2), 273–295.

Johannes, M., Lochstoer, L. A., & Mou, Y. (2016). Learning about consumption dynamics. *Journal of Finance*, *71*(2), 551–600.

Johnson, T. C. (2007). Optimal learning and new technology bubbles. *Journal of Monetary Economics*, *54*(8), 2486–2511.

Jost, A. (1897). *Die Assoziationsfestigkeit in ihrer Abhängigkeit von der Verteilung der Wiederholungen*. L. Voss.

Kahana, M. J. (2012). *Foundations of Human Memory*. Oxford University Press.

Kahana, M. J., Diamond, N. B., & Aka, A. (2022). Laws of Human Memory. Working paper.

Kahneman, D. (2000). Evaluation by moments: Past and future. *Choices, Values, and Frames*, 693–708.

Kensinger, E. A., & Ford, J. H. (2020). Retrieval of emotional events from memory. *Annual Review of Psychology*, *71*, 251–272.

Kleijn, B., & Van Der Vaart, A. (2012). The Bernstein-Von-Mises theorem under misspecification. *Electronic Journal of Statistics*, *6*.

Kuchler, T., & Zafar, B. (2019). Personal experiences and expectations about aggregate outcomes. *Journal of Finance*, *74*(5), 2491–2542.

Lettau, M., & Ludvigson, S. C. (2010). Measuring and modeling variation in the risk-return trade-off. *Handbook of Financial Econometrics*, 617–690.

Lewellen, J., & Shanken, J. (2002). Learning, asset-pricing tests, and market efficiency. *Journal of Finance*, *57*(3), 1113–1145.

Li, K., & Liu, J. (2023). Extrapolative asset pricing. *Journal of Economic Theory*, *210*, 105651.

Lochstoer, L. A., & Muir, T. (2022). Volatility expectations and returns. *Journal of Finance*, *77*(2), 1055–1096.

Lucas, R. E. (1978). Asset prices in an exchange economy. *Econometrica*, 1429–1445.

Malmendier, U., & Nagel, S. (2011). Depression babies: Do macroeconomic experiences affect risk taking? *Quarterly Journal of Economics*, *126*(1), 373–416.

Malmendier, U., & Nagel, S. (2016). Learning from inflation experiences. *Quarterly Journal of Economics*, *131*(1), 53–87.

Malmendier, U., Pouzo, D., & Vanasco, V. (2020). Investor experiences and financial market dynamics. *Journal of Financial Economics*, *136*(3), 597–622.

Malmendier, U., Tate, G., & Yan, J. (2011). Overconfidence and early-life experiences: The effect of managerial traits on corporate financial policies. *Journal of Finance*, *66*(5), 1687–1733.

Malmendier, U., & Wachter, J. A. (2022). Memory of past experiences and economic decisions. Working paper available at SRRN.

Martin, I. (2013). Consumption-based asset pricing with higher cumulants. *Review of Economic Studies*, *80*(2), 745–773.

Mehra, R., & Prescott, E. (1985). The equity premium: A puzzle. *Journal of Monetary Economics*, *15*(2), 145–161.

Mincer, J. A., & Zarnowitz, V. (1969). The Evaluation of Economic Forecasts. In J. A. Mincer (Ed.), *Economic Forecasts and Expectations* (pp. 3–46). National Bureau of Economic Research.

Molavi, P. (2019). Macroeconomics with learning and misspecification: A general theory and applications. Working paper.

Molavi, P., Tahbaz-Salehi, A., & Vedolin, A. (2023). Model Complexity, Expectations, and Asset Prices. *Review of Economic Studies*, (Forthcoming).

Mullainathan, S. (2002). A memory-based model of bounded rationality. *Quarterly Journal of Economics*, *117*(3), 735–774.

Müller, G. E., & Pilzecker, A. (1900). *Experimentelle Beiträge zur Lehre vom Gedächtniss*. Barth.

Nagel, S., & Xu, Z. (2022). Asset Pricing with Fading Memory. *Review of Financial Studies*, (Forthcoming).

Nagel, S., & Xu, Z. (2023). Dynamics of subjective risk premia. *Journal of Financial Economics*, *150*(2).

Phelps, E. A. (2006). Emotion and cognition: Insights from studies of the human amygdala. *Annual Review of Psychology*, *57*, 27–53.

Rietz, T. A. (1988). The equity risk premium: A solution. *Journal of Monetary Economics*, *22*(1), 117–131.

Schacter, D. L. (2008). *Searching for memory: The brain, the mind, and the past*. Basic books.

Shadlen, M. N., & Shohamy, D. (2016). Decision making and sequential sampling from memory. *Neuron*, *90*(5), 927–939.

Shiller, R. (1981). Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends? *American Economic Review*, *71*(3), 421–36.

Shorack, G. R., & Wellner, J. A. (1986). *Empirical processes with applications to statistics*. New York: Wiley.

Sial, A. Y., Sydnor, J. R., & Taubinsky, D. (2023). Biased Memory and Perceptions of Self-Control. NBER Working Paper.

Timmermann, A. G. (1993). How learning in financial markets generates excess volatility and predictability in stock prices. *Quarterly Journal of Economics*, *108*(4), 1135–1145.

Tulving, E., & Schacter, D. L. (1990). Priming and human memory systems. *Science*, *247*(4940), 301–306.

Vissing-Jorgensen, A. (2004). Perspectives on behavioral finance: Does" irrationality" disappear with wealth? Evidence from expectations and actions. *NBER Macroeconomics Annual 2003*, *18*, 139–194.

Wachter, J. A., & Kahana, M. J. (2023). A retrieved-context theory of financial decisions. *Quarterly Journal of Economics*, (Forthcoming).

Weitzman, M. L. (2007). Subjective expectations and asset-return puzzles. *American Economic Review*, *97*(4), 1102–1130.

Zimmermann, F. (2020). The dynamics of motivated beliefs. *American Economic Review*, *110*(2), 337–61.

A Proofs for Section 2

A.1 Proposition 1

I explicitly solve for the parameters of the maximizers of the memory-weighted likelihood. For simplicity, I focus on the case where S is a singleton, such that I do not condition on the state. The result holds state-wise, as the agent makes state-wise inference.

First, I ensure that the memory-weighted true probability distribution integrates to one by defining the integration constant

$$M = \sum_{s \in S} \psi(s) \int_{-\infty}^{\infty} m_{(g_s,t_t)}(g,s)\, q_s^*(g)\, dg.$$

With the transformed memory function $\tilde{m}_{(g_s,t_t)}(g,s) = \frac{1}{M}\, m_{(g_s,t_t)}(g,s)$, we can solve the dual problem

$$\begin{aligned}
LM(g_t,s_t) &= \underset{q \in Q}{\operatorname{argmin}} \left(-M \sum_{s \in S} \psi(s) \int_{-\infty}^{\infty} \tilde{m}_{(g_s,t_t)}(g,s)\, q_s^*(g)\, \log q_s\left(g\right)\, dg \right) \\
&= \underset{q \in Q}{\operatorname{argmin}} \left(-M \sum_{s \in S} \psi(s) \int_{-\infty}^{\infty} \tilde{m}_{(g_s,t_t)}(g,s)\, q_s^*(g)\, \log \left[\frac{1}{\sqrt{2\,\pi\,\sigma^2}} e^{-\frac{(g-\mu)^2}{2\,\sigma^2}} \right] dg \right) \\
&= \underset{q \in Q}{\operatorname{argmin}} \left(-M \sum_{s \in S} \psi(s) \int_{-\infty}^{\infty} \tilde{m}_{(g_s,t_t)}(g,s)\, q_s^*(g) \left[-\frac{\log(2\pi)}{2} - \frac{\log(\sigma^2)}{2} - \frac{(g-\mu)^2}{2\sigma^2} \right] dg \right) \\
&= \underset{\theta \in \Theta}{\operatorname{argmin}} \left(M\, \frac{\log(2\pi)}{2} + M\, \frac{\log(\sigma^2)}{2} + M\, \frac{\tilde{\sigma}_m^2 + (\tilde{\mu}_m - \mu)^2}{2\sigma^2} \right),
\end{aligned}$$

where the last line follows by noting that the integral of the memory-weighted density over the real line is one due to the rescaling, and where I defined the mean and variance of endowment growth under the memory-weighted density by $\tilde{\mu}_m$ and $\tilde{\sigma}_m^2$, respectively. Evaluating the first-order conditions, the distribution that maximizes the memory-weighted likelihood has parameter $\theta_{LM} = (\mu_{LM}, \sigma_{LM}^2)$ given by

$$\mu_{LM} = \tilde{\mu}_m,$$

$$\sigma_{LM}^2 = \tilde{\sigma^2}_m.$$

The parameter space Θ is closed and convex (see definition of $\Theta_{\mathcal{N}}$ in Section 2.1), such that these parameters are unique.

A consistent and unbiased estimator of $\tilde{\mu}_m$, the mean of the memory-weighted true probability distribution, is the sample mean $\hat{\mu}_t$ when drawing from the memory-weighted true probability distribution, $\hat{\mu}_t = \frac{1}{|H_t^R|}\sum_{\tau=-\infty}^{t} g_\tau \mathbb{1}_{\{g_\tau \in H_t^R\}}$. Now, the sample mean almost surely equals it's expected value if the sample is infinite, such that for $\tau_0 \to -\infty$

$$\begin{aligned}
\hat{\mu}_t =& \mathbb{E}\Big[\frac{1}{|H_t^R|}\sum_{\tau=\tau_0}^{t} g_\tau \mathbb{1}_{\{g_\tau \in H_t^R\}}\Big] \\
=& \mathbb{E}\Big[\frac{1}{|H_t^R|}\Big]\,\mathbb{E}\Big[\sum_{\tau=\tau_0}^{t} g_\tau \mathbb{1}_{\{g_\tau \in H_t^R\}}\Big] + \underbrace{\mathrm{Cov}\Big[\frac{1}{|H_t^R|}, \sum_{\tau=\tau_0}^{t} g_\tau \mathbb{1}_{\{g_\tau \in H_t^R\}}\Big]}_{0 \text{ for } \tau_0 \to -\infty} \\
=& \mathbb{E}\Big[\frac{1}{|H_t^R|}\Big]\sum_{\tau=\tau_0}^{t} \mu \cdot m_{(g_t,s_t)}(g_\tau, s_\tau) + \mathbb{E}\Big[\frac{1}{|H_t^R|}\Big]\sum_{\tau=\tau_0}^{t} \mathrm{Cov}\Big[g, \mathbb{1}_{\{g \in H_t^R\}}\Big] \\
=& \mu \cdot \mathbb{E}\Big[\frac{1}{\sum_{\tau=\tau_0}^{t} \mathbb{1}_{\{g_\tau \in H_t^R\}}}\Big] \cdot \sum_{\tau=\tau_0}^{t} \mathbb{E}\left(\mathbb{1}_{\{g_\tau \in H_t^R\}}\right) + \mathbb{E}\Big[\frac{1}{|H_t^R|}\Big]\sum_{\tau=\tau_0}^{t} \mathrm{Cov}\Big[g, \mathbb{1}_{\{g \in H_t^R\}}\Big] \\
=& \mu \cdot \mathbb{E}\Big[\frac{1}{\sum_{\tau=\tau_0}^{t} \mathbb{1}_{\{g_\tau \in H_t^R\}}}\Big] \cdot \mathbb{E}\left(\sum_{\tau=\tau_0}^{t} \mathbb{1}_{\{g_\tau \in H_t^R\}}\right) + \mathbb{E}\Big[\frac{1}{|H_t^R|}\Big]\sum_{\tau=\tau_0}^{t} \mathrm{Cov}\Big[g, \mathbb{1}_{\{g \in H_t^R\}}\Big] \\
=& \mu + \mathbb{E}\Big[\frac{t}{|H_t^R|}\Big] \cdot \mathrm{Cov}\Big[g, \mathbb{1}_{\{g \in H_t^R\}}\Big].
\end{aligned} \tag{A.1}$$

Note that the last step follows since $|H_t^R| = \infty$ deterministically for $\tau_0 \to -\infty$.

Similarly, an estimator of the variance of the memory-weighted true probability distribution is the sample variance, $\hat{\sigma}_t^2 = \frac{1}{|H_t^R|}\sum_{\tau=1}^{t}(g_\tau \mathbb{1}_{\{g_\tau \in H_t^R\}} - \hat{\mu}_t)^2$, which almost surely equals it's expected value for $\tau_0 \to -\infty$

$$\hat{\sigma}_t^2 = \mathbb{E}\Big[\frac{1}{|H_t^R|}\sum_{\tau=\tau_0}^{t} \mathbb{1}_{\{g_\tau \in H_t^R\}}(g_\tau - \hat{\mu}_t)^2\Big]$$

$$=\mathbb{E}\Big[\frac{1}{|H_t^R|}\Big]\ \mathbb{E}\Big[\sum_{\tau=\tau_0}^{t}\mathbb{1}_{\{g_\tau\in H_t^R\}}\big(g_\tau-\hat{\mu}_t\big)^2\Big]$$

$$=\mathbb{E}\Big[\frac{1}{|H_t^R|}\Big]\ \sum_{\tau=\tau_0}^{t}\mathbb{E}\big[\mathbb{1}_{\{g_\tau\in H_t^R\}}\big]\ \mathbb{E}\big[(g_\tau-\hat{\mu}_t)^2\big]+\mathrm{Cov}\big(\mathbb{1}_{\{g\in H_t^R\}},(g-\hat{\mu}_t)^2\big). \tag{A.2}$$

We now solve for

$$\mathbb{E}\big[(g_\tau-\hat{\mu}_t)^2\big]=\underbrace{\mathbb{E}\big[(g_\tau-\mu)^2\big]}_{=\sigma^2}-2\ \underbrace{\mathbb{E}\big[(g_\tau-\mu)(\hat{\mu}_t-\mu)\big]}_{(a)}+\underbrace{\mathbb{E}\big[(\hat{\mu}_t-\mu)^2\big]}_{(b)},$$

where I added and subtracted the true mean μ. We now evaluate (a) and (b) in turn:

$$\begin{aligned}(a)\ \ \mathbb{E}\big[(g_\tau-\mu)(\hat{\mu}_t-\mu)\big]&=\underbrace{\mathbb{E}\big[(g_\tau-\mu)\big]}_{=0}\ \mathbb{E}\big[(\hat{\mu}_t-\mu)\big]+\mathrm{Cov}\big((g_\tau-\mu),(\hat{\mu}_t-\mu)\big)\\&=\mathrm{Cov}\Big(g_\tau,\frac{1}{|H_t^R|}\sum_{j=\tau_0}^{t}g_j\mathbb{1}_{\{g_j\in H_t^R\}}\Big)\\&=\underbrace{\frac{1}{|H_t^R|}}_{=\frac{1}{\infty}\text{ for }\tau_0\to-\infty}\ \underbrace{\mathrm{Cov}\big(g_\tau,g_\tau\mathbb{1}_{\{g_\tau\in H_t^R\}}\big)}_{<\sigma^2<\infty}\\&=0.\end{aligned}$$

From the second to the third line, I used the assumptions that (1) g_τ is i.i.d. and (2) that the memory of $g_j, j\neq\tau$ is independent of g_τ. Next, note that for $\tau_0\to-\infty$, we have that $\hat{\mu}_t=\mathbb{E}(\hat{\mu}_t)$ almost surely, such that we find

$$\begin{aligned}(b)\ \ \mathbb{E}\big[(\hat{\mu}_t-\mu)^2\big]&=\mathbb{E}\big[(\mathbb{E}(\hat{\mu}_t)-\mu)^2\big]\\&=\mathbb{E}\left[\Big(\mathbb{E}\big[\frac{t}{|H_t^R|}\big]\cdot\mathrm{Cov}\big[g,\mathbb{1}_{\{g\in H_t^R\}}\big]\Big)^2\right]\\&=\Big(\mathbb{E}\big[\frac{t}{|H_t^R|}\big]\cdot\mathrm{Cov}\big[g,\mathbb{1}_{\{g\in H_t^R\}}\big]\Big)^2\\&=(\hat{\mu}_t-\mu)^2.\end{aligned}$$

Overall, we thus have

$$\mathbb{E}\big[(g_\tau - \hat{\mu}_t)^2\big] = \sigma^2 + (\hat{\mu}_t - \mu)^2\,,$$

and inserting into Equation A.2 yields the claim.

B Proofs for Section 3

B.1 Proposition 2

I derive the posterior belief of an agent with similarity-weighted memory as given by Equation 11. Note that the state $s \in \{1, 2\}$ follows an observable Markov chain, such that the agent performs state-wise inference. The memory-weighted probability distribution is given by

$$
\begin{aligned}
m(g, g_t) \cdot q(g|s_\tau = s) &= \frac{1}{z_t} \exp\left[-\frac{(g-g_t)^2}{2\,\kappa}\right] \exp\left[-\frac{(g-\mu_s)^2}{2\,\sigma_s^2}\right] \\
&= \frac{1}{z} \exp\left[-\frac{g^2 - 2\,\frac{\kappa\,\mu_s + \sigma_s^2\,g_t}{\kappa+\sigma_s^2}\,g + \frac{\kappa\,\mu_s^2+\sigma_s^2\,g_t^2}{\kappa+\sigma_s^2}}{2\frac{\kappa\,\sigma_s^2}{\kappa+\sigma_s^2}}\right]
\end{aligned}
$$

with integration constant z. The exponential term is Gaussian with

$$
\begin{aligned}
\hat{\mu}_{s,t} &= \frac{\kappa\,\mu_s + \sigma_s^2\,g_t}{\kappa+\sigma_s^2} = \frac{\kappa}{\kappa+\sigma_s^2}\,\mu_s + \frac{\sigma_s^2}{\kappa+\sigma_s^2}\,g_t = (1-\alpha_s)\mu_s + \alpha_s\,g_t, \text{ and} \\
\hat{\sigma}_{s,t}^2 &= \frac{\kappa\,\sigma_s^2}{\kappa+\sigma_s^2} = (1-\alpha_s)\,\sigma_s^2,
\end{aligned}
$$

with $\alpha_s := \frac{\sigma_s^2}{\kappa+\sigma_s^2}$. The prior support of the agent contains all normal distributions. Therefore, the unique maximizer of the memory-weighted likelihood given in Equation (2) is the normal distribution $\mathcal{N}\left(\hat{\mu}_{s,t}, \hat{\sigma}_{s,t}\right)$, see Proposition 1. An alternative, but slightly longer, proof that explicitly uses Proposition 1 is available.

B.2 Proposition 3

It is

$$
\begin{aligned}
\mathbb{E}\left[\hat{\mu}_{s,t+1}|s_{t+1} = s\right] &= \mathbb{E}\left[(1-\alpha_{s_{t+1}})\,\mu_{s_{t+1}} + \alpha_{s_{t+1}}\,g_{t+1}|s_{t+1} = s\right] \\
&= (1-\alpha_{s_{t+1}})\,\mu_s + \alpha_{s_{t+1}}\,\mathbb{E}[g_{t+1}|s_{t+1} = s] \\
&= \mu_s,
\end{aligned}
$$

for $s \in \{1,2\}$. Denote the other state by s_- ($s_- = 2$ if $s = 1$) and

$$\begin{aligned}
\mathbb{E}\left[\hat{\mu}_{s_-,t+1}|s_{t+1}=s\right] &= \mathbb{E}\left[\left(1-\alpha_{s_-}\right)\mu_{s_-} + \alpha_{s_-}\, g_{t+1}|s_{t+1}=s\right] \\
&= \left(1-\alpha_{s_-}\right)\mu_{s_-} + \alpha_{s_-}\,\mathbb{E}[g_{t+1}|s_{t+1}=s] \\
&= \mu_{s_-} + \alpha_{s_-}\,\left(\mu_s - \mu_{s_-}\right).
\end{aligned}$$

Combining both expressions, it is

$$\mathbb{E}\left[\hat{\mu}_{s,t+1}\right] = \pi_s\,\mathbb{E}\left[\hat{\mu}_{s,t+1}|s_{t+1}=s\right] + \left(1-\pi_s\right)\mathbb{E}\left[\hat{\mu}_{s,t+1}|s_{t+1}=s_-\right],$$

and inserting yields the claim in Proposition 4.

B.3 Proposition 4

On average, the posterior variance of endowment growth g_{t+1} is

$$\mathbb{E}\left[\mathrm{Var}_t\left(g_{t+1}\right)\right] = \pi_1\,\sigma_1^2 + \pi_2\,\sigma_2^2 + \pi_1\,\pi_2\,\mathbb{E}\left[\left(\hat{\mu}_{1,t}-\hat{\mu}_{2,t}\right)^2\right].$$

In addition, it is

$$\begin{aligned}
\mathbb{E}\left(\hat{\mu}_{1,t}-\hat{\mu}_{2,t}\right) &= \left(\mu_1-\mu_2\right)\left[1-\left(\alpha_1\,\pi_2+\alpha_2\,\pi_1\right)\right] \\
\mathrm{Var}\left(\hat{\mu}_{1,t}-\hat{\mu}_{2,t}\right) &= \mathrm{Var}\left[(1-\alpha_1)\mu_1-(1-\alpha_2)\mu_2+(\alpha_1-\alpha_2)g_t\right] = \left(\alpha_1-\alpha_2\right)^2\,\mathrm{Var}(g_t).
\end{aligned}$$

By the i.i.d. process and the definition of variance, we can rewrite

$$\begin{aligned}
\mathbb{E}\left[\mathrm{Var}_t\left(g_{t+1}\right)\right] =& \pi_1\,\sigma_1^2 + \pi_2\,\sigma_2^2 + \pi_1\,\pi_2\,\left[\left(\alpha_1-\alpha_2\right)^2\,\mathrm{Var}(g_{t+1}) + \left(\mu_1-\mu_2\right)^2\left[1-\left(\alpha_1\,\pi_2+\alpha_2\,\pi_1\right)\right]^2\right] \\
=& \left(\pi_1\,\sigma_1^2+\pi_2\,\sigma_2^2\right)\left[1+\pi_1\,\pi_2\,\left(\alpha_1-\alpha_2\right)^2\right] \\
&+ \left(\mu_1-\mu_2\right)^2\,\pi_1\,\pi_2\,\left[\pi_1\,\pi_2\left(\alpha_1-\alpha_2\right)^2 + \left[1-\left(\alpha_1\,\pi_2+\alpha_2\,\pi_1\right)\right]^2\right].
\end{aligned}$$

The average perceived riskiness of the agent is larger than the true fundamental riskiness if

$$
\begin{aligned}
\mathbb{E}\left[\operatorname{Var}_t\left(g_{t+1}\right)\right] &\geq \operatorname{Var}\left(g_{t+1}\right) \\
\iff \mathbb{E}\left[\left(\hat{\mu}_{1,t}-\hat{\mu}_{2,t}\right)^2\right] &\geq \left(\mu_1-\mu_2\right)^2 \\
\iff \left(\alpha_1-\alpha_2\right)^2 \operatorname{Var}(g_{t+1}) &\geq \left(\mu_1-\mu_2\right)^2 \left(1-\left[1-\left(\alpha_1\,\pi_2+\alpha_2\,\pi_1\right)\right]^2\right) \\
\iff \left(\alpha_1-\alpha_2\right)^2 \left(\pi_1\,\sigma_1^2+\pi_2\,\sigma_2^2\right) &\geq \left(\mu_1-\mu_2\right)^2 \left[2\,\left(\pi_2\,\alpha_1+\pi_1\,\alpha_2\right)-\pi_2\,\alpha_1^2-\pi_1\,\alpha_2^2\right] \\
\iff \frac{\left(\alpha_1-\alpha_2\right)^2\,\left(\pi_1\,\sigma_1^2+\pi_2\,\sigma_2^2\right)}{2\,\left(\pi_2\,\alpha_1+\pi_1\,\alpha_2\right)-\left(\pi_2\,\alpha_1^2+\pi_1\,\alpha_2^2\right)} &\geq \left(\mu_1-\mu_2\right)^2,
\end{aligned}
$$

where dividing by $(2\,(\pi_2\,\alpha_1+\pi_1\,\alpha_2)-(\pi_2\,\alpha_1^2+\pi_1\,\alpha_2^2))$ does not change the inequality since $2\,(\pi_2\,\alpha_1+\pi_1\,\alpha_2)>\pi_2\,\alpha_1^2+\pi_1\,\alpha_2^2$.

The upper bound on the expected subjective variance is found as

$$
\begin{aligned}
\frac{\mathbb{E}\left[\operatorname{Var}_t\left(g_{t+1}\right)\right]}{\operatorname{Var}(g_{t+1})} =& \frac{\operatorname{Var}(g_{t+1})+\pi_1\,\pi_2\,(\alpha_1-\alpha_2)^2\,\operatorname{Var}(g_{t+1})}{\operatorname{Var}(g_{t+1})} \\
&+\frac{\pi_1\,\pi_2\,\left(\mu_1-\mu_2\right)^2\,\left[-2\,\left(\alpha_1\,\pi_2+\alpha_2\,\pi_1\right)+\left(\alpha_1\,\pi_2+\alpha_2\,\pi_1\right)^2\right]}{\operatorname{Var}(g_{t+1})} \\
=& 1+\underbrace{\pi_1\,\pi_2}_{\leq 0.25}\underbrace{\left(\alpha_1-\alpha_2\right)^2}_{\leq 1}+\underbrace{\left(\alpha_1\,\pi_2+\alpha_2\,\pi_1-2\right)}_{\leq -1}\underbrace{\frac{\pi_1\,\pi_2\,\left(\mu_1-\mu_2\right)^2\,\left(\alpha_1\,\pi_2+\alpha_2\,\pi_1\right)}{\operatorname{Var}(g_{t+1})}}_{>0} \\
\leq& 1.25.
\end{aligned}
$$

B.4 Proposition 5

Consider first the Mincer and Zarnowitz (1969)-regressions, and write

$$
\tilde{\mathbb{E}}_t(g_{t+h})=\pi_1\,\left(\left(1-\alpha_1\right)\mu_1+\alpha_1\,g_t\right)+\pi_2\,\left(\left(1-\alpha_2\right)\mu_2+\alpha_2\,g_t\right)=\tilde{\mu}+\left(\pi_1\,\alpha_1+\pi_2\,\alpha_2\right)\,g_t,
$$

where $\tilde{\mu} = \pi_1\,(1-\alpha_1)\,\mu_1 + \pi_2\,(1-\alpha_2)\,\mu_2$ is the fixed component of the agent's forecast. It is

$$\beta_{MZ} = \frac{\mathrm{Cov}(g_{t+h}, \tilde{\mathbb{E}}_t(g_{t+h}))}{\mathrm{Var}\left(\tilde{\mathbb{E}}_t(g_{t+h})\right)} = \frac{\mathrm{Cov}(g_{t+h}, (\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\,g_t)}{\mathrm{Var}\left((\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\,g_t\right)} = 0,$$

where the last step follows form the i.i.d. structure of endowment growth. Using $\beta_{MZ} = 0$, we find $a_{MZ} = \pi_1\,\mu_1 + \pi_2\,\mu_2$.

Similarly, for the Coibion and Gorodnichenko (2015)-regression, note that

$$\tilde{\mathbb{E}}_t(g_{t+h}) - \mathbb{E}_{t-1}(g_{t+h}) = (\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\,(g_t - g_{t-1}),$$

and denote $\mathbb{E}(g) = \mu = \pi_1\,\mu_1 + \pi_2\,\mu_2$ to obtain

$$\begin{aligned}
\beta_{GC} &= \frac{\mathrm{Cov}\left[g_{t+h} - \tilde{\mathbb{E}}_t(g_{t+h}), \tilde{\mathbb{E}}_t(g_{t+h}) - \mathbb{E}_{t-1}(g_{t+h})\right]}{\mathrm{Var}\left[\tilde{\mathbb{E}}_t(g_{t+h}) - \mathbb{E}_{t-1}(g_{t+h})\right]} \\
&= \frac{(\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\,\mathbb{E}\left[(g_{t+h} - (\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\,g_t - \mu + (\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\,\mu)\cdot(g_t - g_{t-1})\right]}{2\,(\pi_1\,\alpha_1 + \pi_2\,\alpha_2)^2\,\mathrm{Var}(g)} \\
&= \frac{\mathbb{E}\left[g_{t+h}\,g_t - g_{t+h}\,g_{t-1} - (\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\,g_t^2 + (\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\,g_t\,g_{t-1}\right]}{2\,(\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\,\mathrm{Var}(g)} \\
&= \frac{-\,(\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\,\left[\mathbb{E}(g_t^2) - \mu^2\right]}{2\,(\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\,\mathrm{Var}(g)} = -\frac{1}{2}.
\end{aligned}$$

Finally, using Equation 19 and noting that $\mathbb{E}\left(\tilde{\mathbb{E}}_t(g_{t+h}) - \mathbb{E}_{t-1}(g_{t+h})\right) = 0$, it is $a_{GC} = \pi_1\,\pi_2\,(\mu_2 - \mu_1)\,(\alpha_1 - \alpha_2)$.

B.5 Proposition 6

The results follow from inserting the cumulant-generating function given in Equation 14 into Result 1 in Martin (2013), as shown in Section 2.3. Under Epstein and Zin (1989)-preferences, it is

$$
\begin{aligned}
dp_t = & -\log(\beta) - \log\left(\pi_1\, e^{(\lambda-\gamma)\,\hat{\mu}_{1,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_1^2} + \pi_2\, e^{(\lambda-\gamma)\,\hat{\mu}_{2,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_2^2}\right) \\
& + \left(1-\frac{1}{\eta}\right)\,\log\left(\pi_1\, e^{(1-\gamma)\,\hat{\mu}_{1,t}+\frac{1}{2}\,(1-\gamma)^2\,\sigma_1^2} + \pi_2\, e^{(1-\gamma)\,\hat{\mu}_{2,t}+\frac{1}{2}\,(1-\gamma)^2\,\sigma_2^2}\right), \qquad \text{(B.1)} \\
r_t^f = & -\log(\beta) - \log\left(\pi_1\, e^{-\gamma\,\hat{\mu}_{1,t}+\frac{1}{2}\gamma^2\,\sigma_1^2} + \pi_2\, e^{-\gamma\,\hat{\mu}_{2,t}+\frac{1}{2}\,\gamma^2\,\sigma_2^2}\right) \\
& + \left(1-\frac{1}{\eta}\right)\,\log\left(\pi_1\, e^{(1-\gamma)\,\hat{\mu}_{1,t}+\frac{1}{2}\,(1-\gamma)^2\,\sigma_1^2} + \pi_2\, e^{(1-\gamma)\,\hat{\mu}_{2,t}+\frac{1}{2}\,(1-\gamma)^2\,\sigma_2^2}\right), \qquad \text{(B.2)} \\
er_t = & \, dp_t + \log\left(\pi_1\, e^{\lambda\,\hat{\mu}_{1,t}+\frac{1}{2}\,\lambda^2\,\sigma_1^2} + \pi_2\, e^{\lambda\,\hat{\mu}_{2,t}+\frac{1}{2}\,\lambda^2\,\sigma_2^2}\right), \qquad \text{(B.3)} \\
rp_t = & \log\left(\pi_1\, e^{-\gamma\,\hat{\mu}_{1,t}+\frac{1}{2}\gamma^2\,\sigma_1^2} + \pi_2\, e^{-\gamma\,\hat{\mu}_{2,t}+\frac{1}{2}\,\gamma^2\,\sigma_2^2}\right) + \log\left(\pi_1\, e^{\lambda\,\hat{\mu}_{1,t}+\frac{1}{2}\,\lambda^2\,\sigma_1^2} + \pi_2\, e^{\lambda\,\hat{\mu}_{2,t}+\frac{1}{2}\,\lambda^2\,\sigma_2^2}\right) \\
& - \log\left(\pi_1\, e^{(\lambda-\gamma)\,\hat{\mu}_{1,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_1^2} + \pi_2\, e^{(\lambda-\gamma)\,\hat{\mu}_{2,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_2^2}\right). \qquad \text{(B.4)}
\end{aligned}
$$

Expressions for power utility are found by setting $\psi = \frac{1}{\gamma}$, implying $\eta = 1$.

C Proofs for Section 4: Peak-end rule

In this appendix, I derive numerical approximations for the agent's beliefs under the peak-end rule. Recall from Section 4, the memory function under the peak-end rule is defined as

$$
\begin{aligned}
m^{\mathrm{PE}}\left(g_\tau, g_t\right) &= \exp\left[-e^{-\frac{(g_\tau-\mu)^2}{2\,\sigma^2}}\right]\,\exp\left[\frac{(g_\tau-g_t)^2}{2\,\kappa}\right] \\
&= m^P(g_\tau)\cdot m(g_\tau, g_t), \qquad \text{(C.1)}
\end{aligned}
$$

where the first component $m^P(g_\tau) := \exp\left[-e^{-\frac{(g_\tau-\mu)^2}{2\,\sigma^2}}\right]$ overweights extreme experiences and the second component $m(g_\tau, g_t)$ is the similarity-weighted memory function analyzed in Section 3 and Appendix B. Therefore, I focus on the effect of the extreme-experience bias on the agent's posterior beliefs in this appendix.

C.1 Motivation of extreme experience formulation using Extreme Value Theory

Extreme experience bias posits that humans are more likely to recall extreme events Cruciani et al. (2011). The agent observes the realized history of i.i.d. normally distributed random variables $H_t = (g_k, g_{k+1}, ..., g_{t-1}, g_t)$, with $k \to -\infty$. Let $g^m_{k,t} = \max H_t$ be the maximum in a sequence of observations of length k. The distribution of $g^m_{k,t}$ converges to a Gumbel-distribution (or Type-I generalized extreme value distribution) for large k.

Proof. Let $X_\tau = \frac{g_\tau - \mu}{\sigma}$ have i.i.d. standard Normal distribution $X_\tau \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, with CDF $\Phi(x)$ and PDF $\phi(x)$ Define $X^*_n = \max_{1\le\tau\le n} X_\tau$. We search for sequences $\{a_n\}, \{b_n\}$ and a limiting CDF $G(z)$ for $\frac{X^*_n - a_n}{b_n}$ to apply the Fisher–Tippett–Gnedenko theorem.

The CDF of X^*_n is

$$\mathbb{P}(X^*_n \le x) = \mathbb{P}\left(X_1 \le x, X_2 \le x, ..., X_n \le x\right) = \prod_{j=1}^{n} P(X_j \le n) = \Phi^n(x).$$

As X_τ is unbounded, we have $\Phi(x) < 1 \; \forall x$, and $\Phi^n(x) \to 0$ for $n \to \infty$. The maximum $X^*_n \xrightarrow{P} \infty$. In order to achieve a non-degenerate limit, we must standardize X^*_n using (increasing) sequences a_n and b_n.

For $x > 0$, we can use the symmetry of the normal distribution to get

$$\begin{aligned}\Phi(-x) &= \int_x^\infty \phi(z)\, dz \\ &\le \int_x^\infty \frac{z}{x}\phi(z)\, dz = \frac{1}{x\sqrt{2\pi}} \int_x^\infty z\, e^{\frac{z^2}{2}}\, dz = \frac{1}{x}\phi(x) = \frac{1}{x}\phi(x).\end{aligned}$$

We can tighten the bound using Gordon's Inequality as

$$1 \le \frac{\phi(x)}{x\,\Phi(-x)} \le 1 + \frac{1}{x^2}.$$

Now, let $a_n = -\Phi^{-1}\left(\frac{1}{n}\right)$ be the $\left(1 - \frac{1}{n}\right)$'th quantile and set $b_n = \frac{1}{a_n}$. Using the Taylor rule, we find

$$\log \Phi(-a_n - b_n z) = \log \Phi(a_n) + b_n\, z\, \frac{\phi(-a_n)}{\Phi(-a_n)} + o(b_n\, z)$$
$$= \log \frac{1}{n} - z + o(b_n\, z).$$

We can then find

$$\Pr\left(X_1 \leq a_n + b_n z\right) = \Phi(a_n + b_n z) = 1 - \Phi(-a_n - b_n z) \approx 1 - \frac{1}{n} e^z,$$

and

$$\Pr\left(X_n^* \leq a_n + b_n z\right) \approx \left[1 - \frac{1}{n} e^z\right]^n$$
$$\approx \exp\left(-e^{-z}\right) := G(z),$$

with $G(z)$ being the CDF of the Gumbel-distribution. The last approximation follows from $\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = \exp(x)$. For $\{g_\tau\} \overset{i.i.d.}{\sim} \mathcal{N}(\mu, \sigma)$, then we need to change $a_n = \mu - \sigma\Phi^{-1}(1/n)$ and $b_n = -\sigma\Phi^{-1}(1/n)$ to find the Gumbel-distribution as the limit of the standardized maximum. $\square$ The CDF of the Gumbel-distribution for the maximum is

$$G(z; \mu, \sigma^2) = \exp\left(-e^{-\frac{z-\mu}{2\,\sigma^2}}\right),$$

and I obtain $m^P(g_\tau)$ by squaring the distance in the double exponential.

C.2 Memory-weighted probability distribution

Under the assumptions of Proposition 1, the agent's posterior will concentrate on a memory-weighted version of the true probability distribution. I here show that such a distribution exists under extreme-experience bias by finding an integration constant A that implies $\int_{-\infty}^{\infty} m^P(g)\, q^*(g)\, dg = 1$. Let us consider a generalized version of the extreme-experience bias with $\tilde{m}^P(g_\tau) = \exp\left[-e^{-\frac{(g_\tau - a)^2}{2\,b}}\right]$, and

$$\begin{aligned}
\int_{-\infty}^{\infty} \tilde{m}^P(g)\, q^*(g)\, dg &= \int_{\mathbb{R}} e^{-\frac{(g-\mu)^2}{2\,\sigma^2}}\ e^{-e^{-\frac{(g-a)^2}{2\,b}}}\ dg \\
&= \int_{\mathbb{R}} e^{-\frac{(g-\mu)^2}{2\,\sigma^2}} \ \sum_{k=0}^{\infty} \frac{(-1)^k}{k!}\, e^{-k\,\frac{(g-a)^2}{2\,b}}\, dg \\
&= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{\mathbb{R}} e^{-\frac{g^2\left(\sigma^2 k + b\right) - 2g\left(\sigma^2\, k\, a + \mu\, b\right) + \left(\sigma^2\, k\, a^2 + \mu^2\, b\right)}{2\sigma^2\, b}}\, dg \\
&= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \sqrt{\frac{2\,\pi\,\sigma^2\, b}{\sigma^2\, k + b}}\, e^{-\frac{k}{2}\,\frac{(\mu-a)^2}{\sigma^2\, k + b}} = A^{-1},
\end{aligned}$$

where I used the series expansion of the exponential function in the first line. The integration constant A exists and is a well-defined function of the parameters.

C.3 Numerical approximation of the subjective moments under extreme experience bias

Restrict attention to the extreme-experience bias $m^P(g_\tau)$ with $a = \mu$ and $b = \sigma^2$.[34] Under this assumption, the agent is more likely to recall experiences that are further away from the mean of the underlying growth-rate distribution while acknowledging the scale of the underlying distribution σ^2. Behaviorally, the specification implies a memory-formulation evaluates extremess relative to the true underlying process. If the growth-rates are generated from a more volatile process, an observation needs to be larger (in absolute terms) to be considered extreme. Similarly, a growth-rate that is close to μ is considered "normal" and thus less likely to be recalled under extreme-experience biased memory.

Using this formulation, the agent's posterior expectation of the growth-rate is

$$\hat{\mu}_t = A \int_{\mathbb{R}} g\, e^{-\frac{(g-\mu)^2}{2\sigma^2}}\ e^{-e^{-\frac{(g-\mu)^2}{2\,\sigma^2}}}\ dg$$

[34] Similar results exists for more general versions with either $a \neq \mu$ or $b \neq \sigma^2$ and are available upon request.

$$
\begin{aligned}
&= A \int_{\mathbb{R}} (x+\mu)\, e^{-\frac{x^2}{2\sigma^2}}\, e^{-e^{-\frac{x^2}{2\sigma^2}}}\, dx \\
&= A \int_{\mathbb{R}} x\, e^{-\frac{x^2}{2\sigma^2}}\, e^{-e^{-\frac{x^2}{2\sigma^2}}}\, dx + \mu\, A \int_{\mathbb{R}} e^{-\frac{x^2}{2\sigma^2}}\, e^{-e^{-\frac{x^2}{2\sigma^2}}}\, dx \\
&= A \int_{\mathbb{R}} x\, e^{-\frac{x^2}{2\sigma^2}}\, e^{-e^{-\frac{x^2}{2\sigma^2}}}\, dx + \mu,
\end{aligned}
$$

where I used a change of variables and the last line follows from the definition of A (to see this, you can reverse the substitution $x = \Delta c - \mu$). Next, let us define $y = e^{-\frac{x^2}{2\sigma^2}}$ to find

$$
\begin{aligned}
\int_{\mathbb{R}} x\, e^{-\frac{x^2}{2\sigma^2}}\, e^{-e^{-\frac{x^2}{2\sigma^2}}}\, dx &= \int_{-\infty}^{0} x\, e^{-\frac{x^2}{2\sigma^2}}\, e^{-e^{-\frac{x^2}{2\sigma^2}}}\, dx + \int_{0}^{\infty} x\, e^{-\frac{x^2}{2\sigma^2}}\, e^{-e^{-\frac{x^2}{2\sigma^2}}}\, dx \\
&= \int_0^1 -\sigma^2\, e^{-y}\, dy + \int_1^0 -\sigma^2\, e^{-y}\, dy \\
&= -\sigma^2 \left(\int_0^1 e^{-y}\, dy - \int_0^1 e^{-y}\, dy \right) = 0,
\end{aligned}
$$

which implies that $\hat{\mu}_t = \mu$ under extreme experience bias. If the agent symmetrically overweights the tails of the underlying distribution (which is also symmetric), she will learn the correct mean growth rate.

Next, I approximate the perceived variance of the agent. Define $u = \frac{(g-\mu)}{\sqrt{2\,\sigma^2}}$. It is

$$
\begin{aligned}
\hat{\sigma}_t^2 &= A \int_{\mathbb{R}} (g-\mu)^2\; e^{-\frac{(g-\mu)^2}{2\,\sigma^2}}\; e^{-e^{-\frac{(g-\mu)^2}{2\,\sigma^2}}}\; dg \\
&= A\, 2\sqrt{2}\,\sigma^3 \int_{\mathbb{R}} u^2\, e^{-u^2}\; e^{-e^{-u^2}}\; du.
\end{aligned}
$$

In general, no closed form solution exists for the integral, but we can approximate it using various substitutions. First, use $y = e^{-u^2}$:

$$
\begin{aligned}
\int_{\mathbb{R}} u^2\, e^{-u^2}\, e^{-e^{-u^2}}\, du &= 2 \int_0^{\infty} u^2\, e^{-u^2}\, e^{-e^{-u^2}}\, du \\
&= \int_0^1 \sqrt{-\ln(y)}\; e^{-y}\, dy \\
&= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_0^1 \sqrt{-\ln(y)}\; y^k\, dy.
\end{aligned}
$$

The inner integral can be solved by using $v = -\ln(y)$ as

$$\begin{aligned}
\int_0^1 \sqrt{-\ln(y)}\, y^k \, dy &= \int_\infty^0 \sqrt{v}\, e^{-k\,v} \left(\frac{dy}{dv}\right) dv \\
&= \int_0^\infty \sqrt{v} e^{-(k+1)v} \, dv \\
&= \frac{\Gamma(1.5)}{(k+1)^{3/2}} \\
&= \frac{\sqrt{\pi}}{2\,(k+1)^{3/2}},
\end{aligned}$$

where the last lines follows by the properties of the Gamma-function for half-integers.

Putting terms together, it is

$$\begin{aligned}
\hat{\sigma}_t^2 &= A\, 2\sqrt{2}\, \sigma^3 \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{\sqrt{\pi}}{2\,(k+1)^{3/2}} \\
&= A \sqrt{2\,\pi}\, \sigma^3 \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{1}{(k+1)^{3/2}} \\
&= \sigma^2 \frac{\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{1}{(k+1)^{3/2}}}{\sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \frac{1}{\sqrt{m+1}}} \\
&\approx \sigma^2 \cdot 1.4108 > \sigma^2.
\end{aligned}$$

As expected, extreme experience-biased memory leads to a higher fundamental variance. Intuitively, the agent's memory overweights observations that are further away from the mean of the underlying distribution. Therefore, the growth-rate process seems riskier than it actually is under the agent's filtration.

C.4 Numerical approximation of the subjective moments under peak-end rule

Let us consider the memory function defined in Equation 34. First, I show that the memory-weighted probability distribution exists by showing that the integration constant $\mathcal{V}$ exists:

$$\begin{aligned}
\int_{-\infty}^{\infty} m^{PE}(g, g_t)\, q^*(g)\, dg &= \int_{\mathbb{R}} e^{-\frac{(g-\mu)^2}{2\,\sigma^2}}\; e^{-\frac{(g-g_t)^2}{2\,\kappa}}\; e^{-e^{-\frac{(g-\mu)^2}{2\,\sigma^2}}}\, dg \\
&= \int_{\mathbb{R}} e^{-\left[\frac{(g-\mu)^2}{2\,\sigma^2}+\frac{(g-g_t)^2}{2\,\kappa}\right]} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!}\, e^{-k\,\frac{(g-\mu)^2}{2\,\sigma^2}}\, dg \\
&= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{\mathbb{R}} e^{-\left[(1+k)\,\frac{(g-\mu)^2}{2\,\sigma^2}+\frac{(g-g_t)^2}{2\,\kappa}\right]}\, dg \\
&= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{\mathbb{R}} e^{-\frac{g^2\left((1+k)\,\kappa+\sigma^2\right)-2g\left((1+k)\,\kappa\,\mu+\sigma^2\,g_t\right)+\left((1+k)\,\kappa\,\mu^2+\sigma^2\,g_t^2\right)}{2\sigma^2\,\kappa}}\, dg \\
&= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \sqrt{\frac{2\,\pi\,\sigma^2\,\kappa}{\sigma^2+(1+k)\,\kappa}}\; e^{-\frac{(1+k)}{2}\,\frac{(\mu-g_t)^2}{(1+k)\,\kappa+\sigma^2}} = \mathcal{V}^{-1},
\end{aligned}$$

where I used the series expansion of the exponential function in the first line. The integration constant $\mathcal{V}$ exists and is a well-defined function of the parameters.

Next, we approximate the agent's posterior mean under the peak-end rule memory function. Define the mean and variance under similarity-weighted memory (see Proposition 2) as

$$\begin{aligned}
\hat{\mu}_t^S &= \frac{\kappa\,\mu+\sigma^2\,g_t}{\kappa+\sigma^2} = (1-\alpha)\,\mu+\alpha\,g_t \\
(\hat{\sigma}_t^S)^2 &= \frac{\kappa\sigma^2}{\kappa+\sigma^2},
\end{aligned}$$

with $\alpha = \frac{\sigma^2}{\kappa+\sigma^2}$. We can then rewrite the peak-end memory function as

$$m^{PE}(g_\tau, g_t) = e^{-\frac{(g_\tau-\mu)^2}{2\,\sigma^2}}\; e^{-\frac{(g_\tau-g_t)^2}{2\,\kappa}}\; e^{-e^{-\frac{(g_\tau-\mu)^2}{2\,\sigma^2}}}$$

$$= e^{\frac{(\mu - g_t)^2}{2}} \, e^{\frac{(g_T - \hat{\mu}_t^S)^2}{2\,(\hat{\sigma}_t^S)^2}} \, e^{-e^{-\frac{(g_T - \mu)^2}{2\,\sigma^2}}}.$$

The agent's posterior mean under the peak-end rule memory function can then be obtained as

$$\begin{aligned}
\hat{\mu}_t &= \mathcal{V}\, e^{\frac{(\mu - g_t)^2}{2}} \int_{\mathbb{R}} g\, e^{\frac{(g - \hat{\mu}_t^S)^2}{2\,(\hat{\sigma}_t^S)^2}} \, e^{-e^{-\frac{(g-\mu)^2}{2\,\sigma^2}}} \, dg \\
&= \mathcal{V}\, e^{\frac{(\mu - g_t)^2}{2}} \int_{\mathbb{R}} \left(x + \mu + \hat{\mu}_t^S\right) e^{-\frac{(x+\mu)^2}{2\,(\hat{\sigma}_t^S)^2}} \, e^{-e^{-\frac{(x+\hat{\mu}_t^S)^2}{2\,\sigma^2}}} \, dx \\
&= \mathcal{V}\, e^{\frac{(\mu - g_t)^2}{2}} \int_{\mathbb{R}} x\, e^{-\frac{(x+\mu)^2}{2\,(\hat{\sigma}_t^S)^2}} \, e^{-e^{-\frac{(x+\hat{\mu}_t^S)^2}{2\,\sigma^2}}} \, dx + \mu + \hat{\mu}_t^S.
\end{aligned}$$

We cannot simplify the first integral using the same steps as in Appendix C.3, because the function $f(x) = x\, e^{-\frac{(x+\mu)^2}{2\,(\hat{\sigma}_t^S)^2}} \, e^{-e^{-\frac{(x+\hat{\mu}_t^S)^2}{2\,\sigma^2}}}$ is, in general, not symmetric. Therefore, I approximate the integral using the series expansion of the exponential function as

$$\begin{aligned}
\int_{\mathbb{R}} x\, e^{-\frac{(x+\mu)^2}{2\,(\hat{\sigma}_t^S)^2}} \, e^{-e^{-\frac{(x+\hat{\mu}_t^S)^2}{2\,\sigma^2}}} &= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{\mathbb{R}} x\, e^{-\frac{(x+\mu)^2}{2\,(\hat{\sigma}_t^S)^2}} \, e^{-k\,\frac{(x+\hat{\mu}_t^S)^2}{2\,\sigma^2}} \, dx \\
&= -\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{\sqrt{2\,\pi}\,\sigma\,\hat{\sigma}_t^S}{(\sigma^2 + k\,(\hat{\sigma}_t^S)^2)^{3/2}} \left(\mu\,\sigma^2 + k\,\hat{\mu}_t^S\,(\hat{\sigma}_t^S)^2\right) e^{-\frac{k}{2}\cdot\frac{(\mu - \hat{\mu}_t^S)^2}{\sigma^2 + k\,(\hat{\sigma}_t^S)^2}}.
\end{aligned}$$

Thus, putting terms together, it is

$$\begin{aligned}
\hat{\mu}_t &= \mu + \hat{\mu}_t^S - \mathcal{V}\, e^{\frac{(\mu - g_t)^2}{2}} \left(\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{\sqrt{2\,\pi}\,\sigma\,\hat{\sigma}_t^S}{(\sigma^2 + k\,(\hat{\sigma}_t^S)^2)^{3/2}} \left(\mu\,\sigma^2 + k\,\hat{\mu}_t^S\,(\hat{\sigma}_t^S)^2\right) e^{-\frac{k}{2}\cdot\frac{(\mu - \hat{\mu}_t^S)^2}{\sigma^2 + k\,(\hat{\sigma}_t^S)^2}} \right) \\
&= \mu + \hat{\mu}_t^S - e^{\frac{(\mu - g_t)^2}{2}} \frac{\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{\kappa + \sigma^2}{((1+k)\,\kappa + \sigma^2)^{3/2}} \left(\mu + k\,\frac{\kappa}{\kappa + \sigma^2}\hat{\mu}_t^S\right) e^{-\frac{k}{2}\cdot\frac{\alpha^2\,(\mu - g_t)^2}{\sigma^2 + k\,(\hat{\sigma}_t^S)^2}}}{\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \sqrt{\frac{1}{\sigma^2 + (1+k)\,\kappa}}\, e^{-\frac{(1+k)}{2}\,\frac{(\mu - g_t)^2}{(1+k)\,\kappa + \sigma^2}}}.
\end{aligned}$$

D Additional figures

Figure D.1: First four posterior moments of endowment growth under similarity-weighted memory

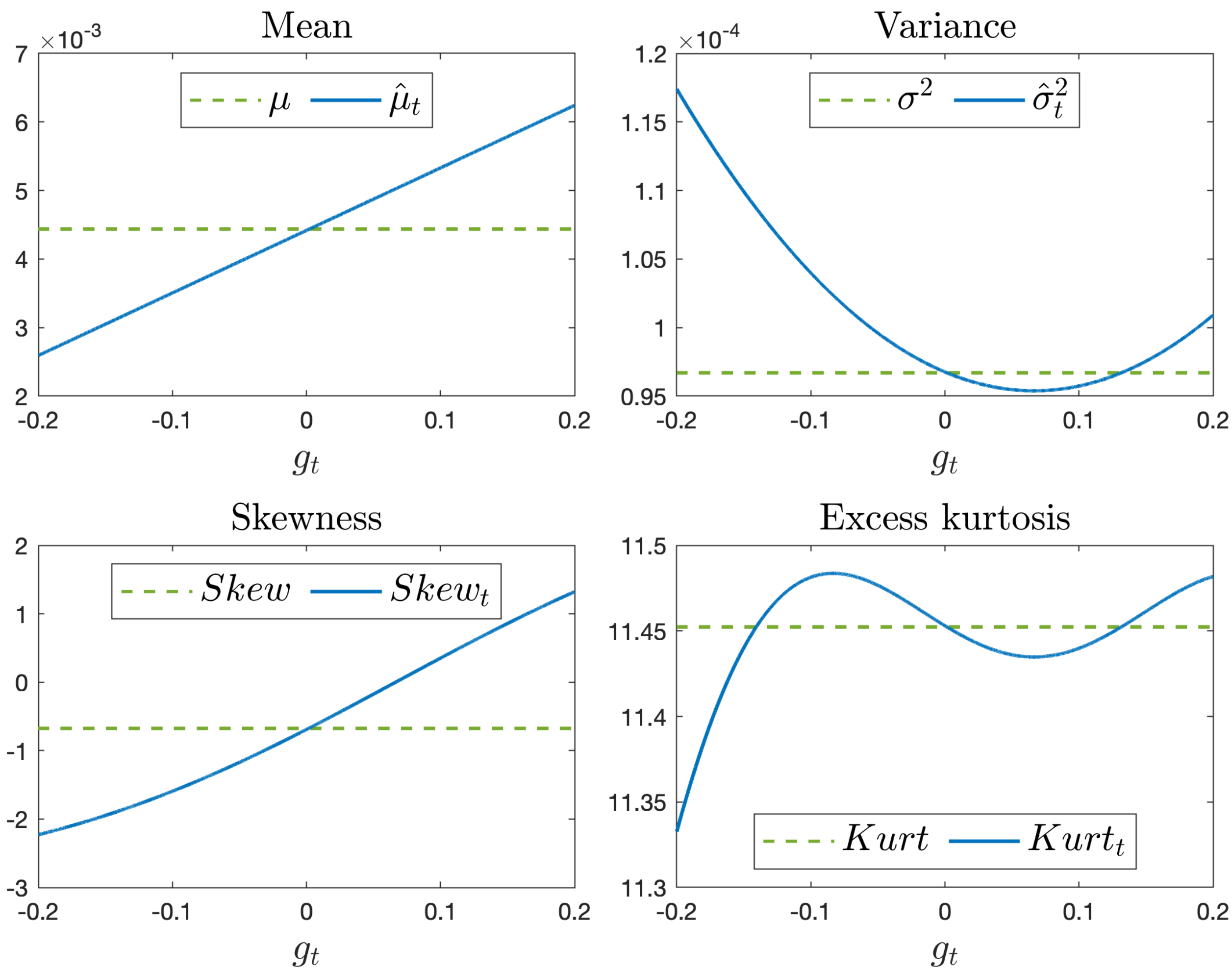

Figure D.1 shows the first four moments of endowment growth under the agent's subjective beliefs derived from similarity-weighted memory (solid blue line) and the true underlying values (green dashed line). Endowment growth is distributed as in Equation 10, and the similarity-weighted memory function is as in Equation 11. The parameters used to generate Figure D.1 are as in Table 1.

Figure D.2: Cumulant-generating function for different realizations of endowment growth

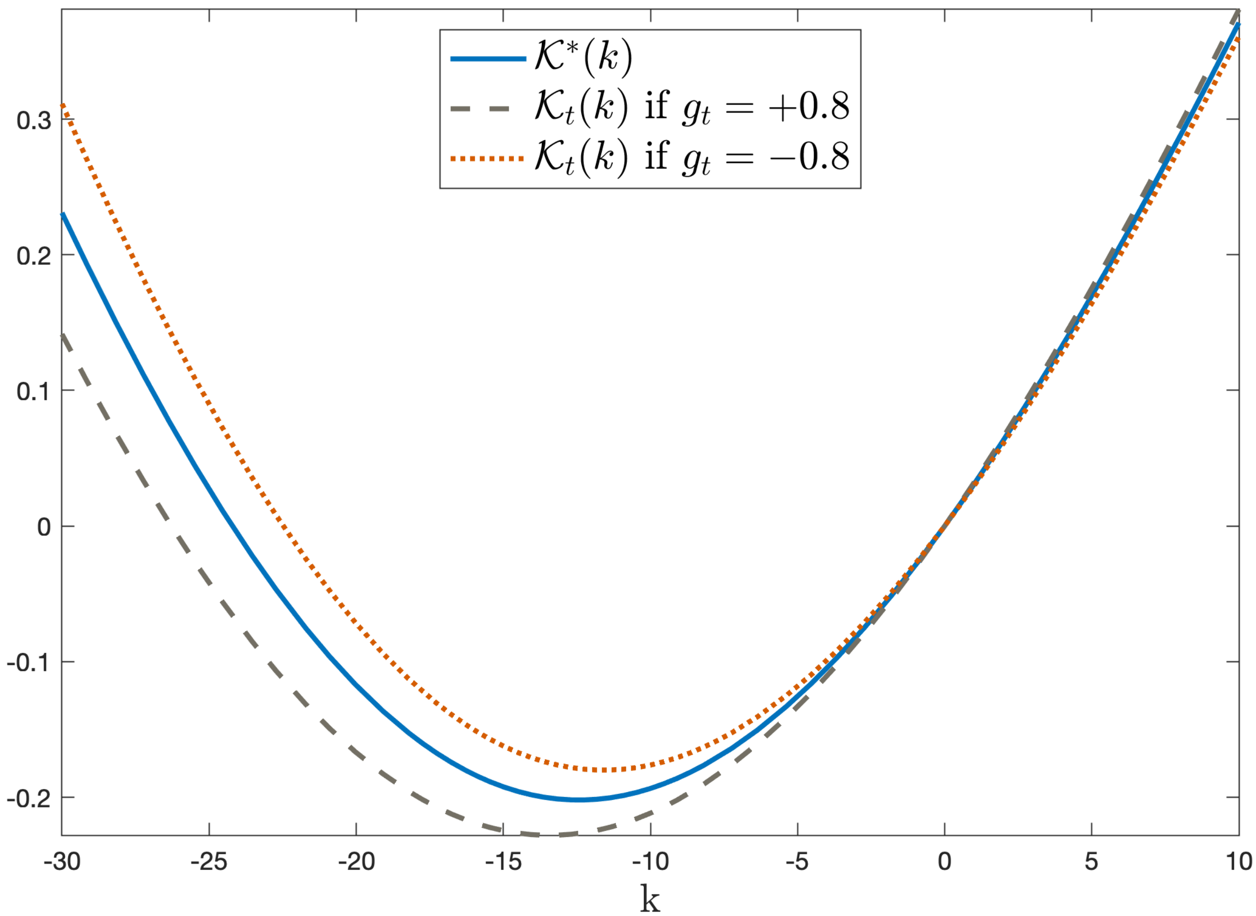

Figure D.2 shows the true cumulant-generating function of endowment growth $\mathcal{K}^*(k)$ (blue solid line) and the agent's subjective cumulant-generating function for a highly positive (grey dashed line) and negative (orange dotted line) current endowment growth g_t. The cumulant-generating functions given in Equation 14. The parameters are $\mu_1 = 0.05$, $\mu_2 = -0.05$, $\sigma_1 = 0.01$, $\sigma_2 = 0.02$, $\pi_1 = 0.8$, and $\kappa = 0.1$. Note that for any cumulant-generating function $\mathcal{K}(0)$, it is $\mathcal{K}(0) = 0$, such that changes of the current endowment growth lead to a rotation of the cumulant-generating function.

Figure D.3: Graphical construction of the memory-weighted true probability distribution

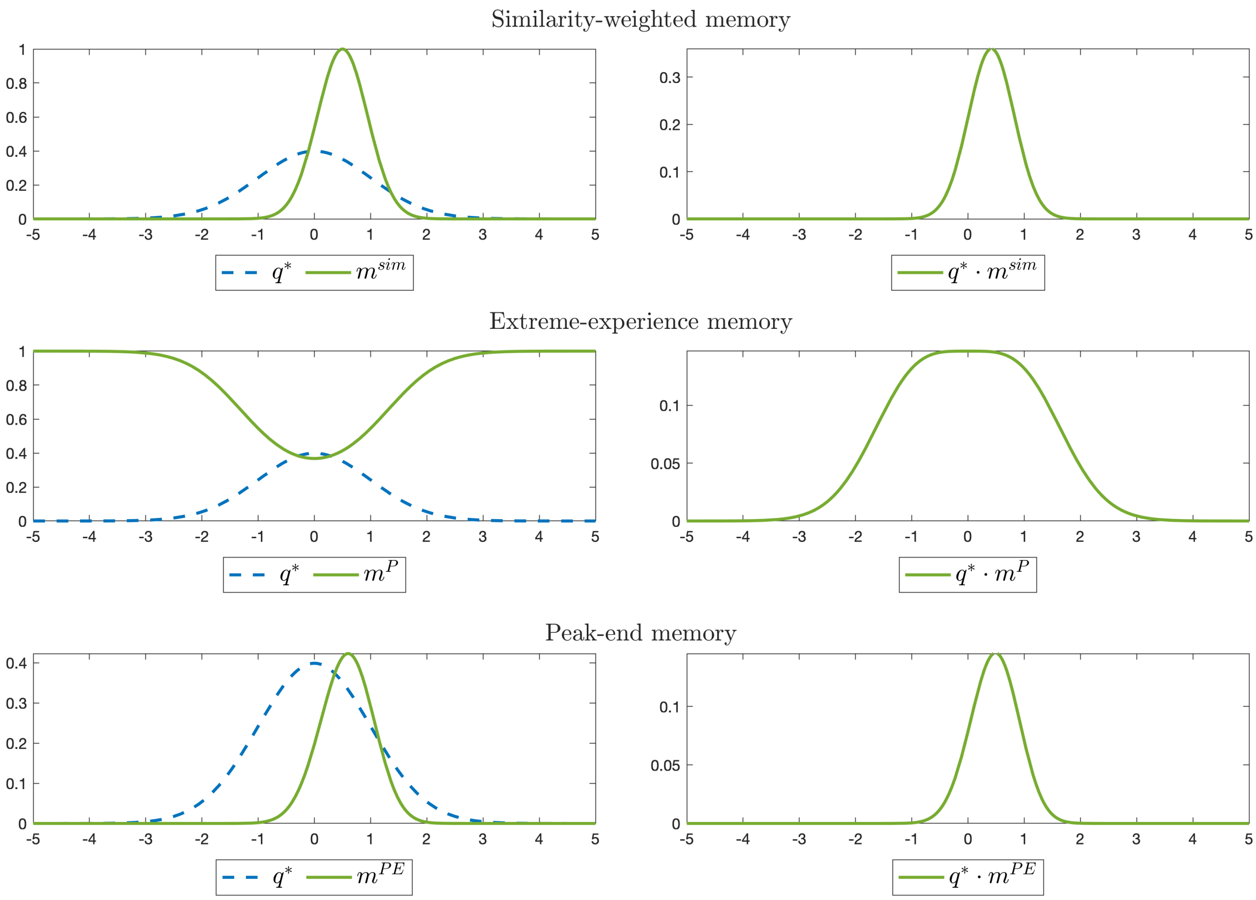

Figure D.3 shows the graphical construction of the memory-weighted true probability distribution. The first row shows the case of a similarity-weighted memory function, as considered in Section 3, the middle row shows the extreme-experience bias and the bottom row shows the peak-end memory function, which are both discussed in Section 4. The left column does always show the true probability distribution (blue dashed line), which is a standard normal distribution, and the respective memory function (green solid line). The right column shows the resulting memory-weighted true probability distribution, which is scaled to integrate to one. The memory-scrutiny parameter is $\kappa = 0.2$, and the current endowment growth is $g_t = 0.5$.

E Additional tables

Table E.1: Average asset pricing outcomes per quantile of log endowment growth

Quantile	No parameter uncertainty $\overline{g}_t$	$\overline{\hat{\mu}}_t$	$\overline{rp}_t$	$\overline{rp}_t^O$	Parameter uncertainty $\overline{g}_t$	$\overline{\hat{\mu}}_t$	$\overline{rp}_t$	$\overline{rp}_t^O$
1	-2.965	1.740	1.211	6.421	-2.954	1.719	1.289	18.639
2	0.812	1.775	1.201	2.084	0.813	1.769	1.081	4.674
3	1.920	1.785	1.199	0.819	1.920	1.784	1.069	1.962
4	3.010	1.795	1.196	-0.423	3.011	1.798	1.070	-0.812
5	6.090	1.822	1.190	-3.913	6.090	1.839	1.060	-8.414

Table E.1 reports average endwoment growth $\overline{g}_t$, posterior mean $\overline{\hat{\mu}}_t$, subjective risk premium $\overline{rp}_t$ and objective risk premium $\overline{rp}_t^O$ for each quantile of endowment growth. The moments are obtained from $10,000$ simulations of the model for 304 quarters. For the parameter uncertainty simulations, I use 120 quarters burn-in period (30 years) and I draw 10 realizations of the agent's memory for each of the $10,000$ economies. All averages are annualized by multiplying quarterly means by four.

Table E.2: Predictability and Coibion and Gorodnichenko (2015)-regressions under peak-end rule

	RP_{Subj}	RP_{Obj}	$\hat{b}_{CG}$
dp_t	-0.0002	9.5315	
$\left(\tilde{\mathbb{E}}_t - \mathbb{E}_{t-1}\right) g_{t+1}$			-0.3927

Table E.2 reports the mean estimates from regressions for $10,000$ simulations of the model for 304 quarters. The first row shows the mean coefficients when regressing subjectively expected and objectively obtained risk premia on the log dividend-price ratio, as in Nagel and Xu (2023). The price-dividend ratio is rescaled to unit standard deviation. The second row shows the mean estimate from Coibion and Gorodnichenko (2015)-regressions of the forecast error on the forecast revision. The agent's expectations are obtained under the peak-end memory distortion given in Equation 34.

F Asset pricing model and further asset-pricing results

In this appendix, I derive the asset-pricing results in Section 2.3 following Martin (2013). Consider the objective function in Equation 5 with $\psi \neq 1$[35] Under this formulation, the

[35]All results in this section extend to the case with unit EIS, $\psi = 1$. One can solve the case with $\psi = 1$ using the recursion in Hansen et al. (2008). The consumption-wealth ratio is constant for $\psi = 1$, and all other results generalize as the limit of $\eta \to \infty$.

stochastic discount factor becomes

$$M_{t+1} = \beta^{\eta} \left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\eta}{\psi}} (R_{w,t+1})^{\eta-1},$$

where the return on wealth, $R_{w,t+1}$, is

$$R_{w,t+1} = \frac{C_{t+1} + W_{t+1}}{W_t} = \frac{C_{t+1}}{C_t} \left(\frac{C_t}{W_t} + \frac{C_t}{W_t}\frac{W_{t+1}}{C_{t+1}}\right) = \frac{C_{t+1}}{C_t} \left(1 + CW\right),$$

where I conjecture that the consumption-wealth ratio CW is constant under the agent's beliefs.[36] I verify this conjecture below.

The price-dividend ratio of an asset that pays $D_t = C_t^{\lambda}$ under the agent's time-t beliefs is given by

$$\begin{aligned}\frac{P_t}{D_t} &= \mathbb{E}_t\left[\sum_{j=1}^{\infty} \beta^{j\,\eta} \left(\frac{C_{t+j}}{C_t}\right)^{-\frac{\eta}{\psi}} \left(\frac{C_{t+j}}{C_t}\right)^{\lambda} \left(\frac{C_{t+j}}{C_t}\right)^{\eta-1} (1+CW)^{j\,(\eta-1)}\right] \\ &= \sum_{j=1}^{\infty} \beta^{j\,\eta}\, \mathbb{E}_t\left[e^{(\lambda-\gamma)\,g_{t+1}}\right]^j \, e^{j\,(\eta-1)\,cw} = \frac{1}{e^{-\eta\,\log(\beta)+(1-\eta)\,cw-\mathcal{K}_t(\lambda-\gamma)}-1},\end{aligned}$$

if $-\eta\,\log(\beta)+(1-\eta)\,cw-\mathcal{K}_t(\lambda-\gamma) > 0$ and where I define $cw = \log\left(1+\frac{C}{W}\right)$. As in the main text, define the log dividend-yield as $dp_t = \log(1+\frac{D_t}{P_t}) = -\eta\,\log(\beta) + (1-\eta)\,cw - \mathcal{K}_t(\lambda-\gamma)$. The consumption-wealth ratio equals the dividend-price ratio for the wealth-portfolio with $\lambda = 1$, such that

$$cw = -\log(\beta) - \frac{1}{\eta}\mathcal{K}_t(1-\gamma),$$

which is constant under the agent's time-t beliefs as conjectured because the agent expects that $\mathcal{K}_{t+h}(k) = \mathcal{K}_t(k)$ for all $h \geq 1$. The dividend-price ratio is then

$$dp_t = -\log(\beta) + \left(1-\frac{1}{\eta}\right)\mathcal{K}_t(1-\gamma) - \mathcal{K}_t(\lambda-\gamma).$$

[36]Note that, by assumption, the agent knows that endowment growth is i.i.d..

Using the constant dividend-price ratio under the agent's beliefs, the subjective expected return on any asset is

$$\tilde{\mathbb{E}}_t\left[R_{t+1}\right] = \tilde{\mathbb{E}}_t\left[\frac{D_{t+1}}{D_t}\right]\left(1 + \frac{D_t}{P_t}\right) = \tilde{\mathbb{E}}_t\left[e^{\lambda\, g_{t+1}}\right] e^{dp_t}$$

and the log of the expected return is

$$er_t = -\log(\beta) + \mathcal{K}_t(\lambda) + \left(1 - \frac{1}{\eta}\right)\mathcal{K}_t(1-\gamma) - \mathcal{K}_t(\lambda - \gamma).$$

The risk-free rate is found by setting $\lambda = 0$,

$$r_t^f = -\log(\beta) + \left(1 - \frac{1}{\eta}\right)\mathcal{K}_t(1-\gamma) - \mathcal{K}_t(-\gamma),$$

and the risk premium on any asset is

$$rp_t = \mathcal{K}_t(\lambda) + \mathcal{K}_t(-\gamma) - \mathcal{K}_t(\lambda - \gamma).$$

Note that, under the agent's i.i.d. beliefs, the risk premium is independent of the elasticity of intertemporal substitution. Moreover, since the agent recalls an infinite history of observations, she has no parameter uncertainty that could be priced under Epstein-Zin preferences (Collin-Dufresne et al., 2016).

In addition, consider the objectively expected return under the econometrician's filtration. It is

$$\mathbb{E}\left(R_{t+1}\right) = \frac{D_t}{P_t}\,\mathbb{E}\left[\frac{D_{t+1}}{D_t}\right]\left(1 + \mathbb{E}\left[\frac{P_{t+1}}{D_{t+1}}\right]\right),$$

where I can no longer use the observation that the dividend-price ratio in period t equals the dividend-price ratio in period $t+1$. The exact present-value relation is non-linear in the expected revision of the agent's beliefs such that I apply a Campbell and Shiller (1988) approximation, as used in Campbell (1991). Denote the log-return on any asset as $r_{t+1} := \log(R_{t+1})$. It is

$$r_{t+1} - \tilde{\mathbb{E}}_t(r_{t+1}) = \lambda\ (\mathbb{E}_{t+1} - \mathbb{E}_t) \sum_{j=0}^{\infty} \bar{p}^j\, g_{t+1+j} - (\mathbb{E}_{t+1} - \mathbb{E}_t) \sum_{j=1}^{\infty} \bar{p}^j r_{t+1+j},$$

where $\bar{p} = \frac{1}{1+exp(\bar{p-d})} \approx 0.95$ annually, see Campbell (2017). The expected log-return—determined in equilibrium by an agent with selective and stochastic memory—is $\tilde{\mathbb{E}}_t\,(r_{t+1}) = \lambda\, \tilde{\mathbb{E}}_t(g_{t+1}) + dp_t$, such that we can rewrite the unexpected log return as

$$r_{t+1} - \tilde{\mathbb{E}}_t(r_{t+1}) = \lambda \left(g_{t+1} - \tilde{\mathbb{E}}_t(g_{t+1}) \right) - \frac{\bar{p}}{1-\bar{p}}\, (dp_{t+1} - dp_t)\,.$$

As a next step, rewrite the expected log-return using the (observable) risk-free rate to find

$$r_{t+1} - r_t^f = \lambda g_{t+1} + \mathcal{K}_t(-\gamma) - \mathcal{K}_t(\lambda - \gamma) - \frac{\bar{p}}{1-\bar{p}}\, (dp_{t+1} - dp_t)\,.$$

Taking objective expectations and noting that time-t quantities are observable then yields

$$\mathbb{E}\,(r_{t+1}) - r_t^f = \lambda \mathbb{E}(g_{t+1}) + \mathcal{K}_t(-\gamma) - \mathcal{K}_t(\lambda - \gamma) - \frac{\bar{p}}{1-\bar{p}}\, (\mathbb{E}\,(dp_{t+1}) - dp_t)\,,$$

with

$$\mathbb{E}(dp_{t+1}) = -\log(\beta) + \left(1 - \frac{1}{\eta}\right) \mathbb{E}\,(\mathcal{K}_{t+1}\,(1-\gamma)) - \mathbb{E}\,(\mathcal{K}_{t+1}\,(\lambda - \gamma))\,.$$

In general, we cannot obtain the expectation of the cumulant-generating function under the agents beliefs in closed-form. Therefore, I use a second-order Taylor approximation around $\tilde{\mathcal{M}}(k) := \mathbb{E}\,(\mathcal{M}_{t+1}(k))$, as

$$\mathbb{E}\,(\mathcal{K}_{t+1}(k)) = \mathbb{E}\,(\log \mathcal{M}_{t+1}(k)) \approx \log\left(\tilde{\mathcal{M}}(k)\right) + \frac{1}{2} \left(\frac{\mathbb{E}\,(\mathcal{M}_{t+1}(k)^2)}{\tilde{\mathcal{M}}(k)^2} - 1 \right).$$

I now highlight how these results can be applied to the similarity-weighted memory discussed in Section 3 of the main text.

Case 1: Log-normally distributed endowment growth

First, let us consider the case of log-normal endowment growth,

$$g_t = \mu + \sigma\,\epsilon_t,$$

where the agent learns about the mean under similarity-weighted memory. Her posterior belief for the mean is then $\mu_t = (1-\alpha)\,\mu + \alpha\, g_t$ and $\alpha = \sigma^2/(\kappa + \sigma^2)$ as in the main text.

The moment-generating and cumulant-generating functions under the agent's time-t beliefs are

$$\mathcal{M}_t(k) = \tilde{\mathbb{E}}_{\mathrm{t}}\left(e^{k\,g_t}\right) = e^{k\mu_t + \frac{1}{2}\,k^2\,\sigma^2}$$
$$\mathcal{K}_t(k) = = \log\left(\mathcal{M}_t(k)\right) = k\mu_t + \frac{1}{2}\,k^2\,\sigma^2.$$

The objective expectation of the agent's cumulant-generating function is then simply

$$\mathbb{E}\left(\mathcal{K}_t(k)\right) = k\,\mu + \frac{1}{2}\,k^2\,\sigma^2 = \mathcal{K}^*(k).$$

Inserting into the previous equations, I find

$$\begin{aligned}
r_t^f &= -\log(\beta) + \frac{1}{\psi}\,\mu_t - \frac{1}{2}\,\sigma^2\left(\gamma - \frac{1-\gamma}{\psi}\right),\\
dp_t &= -\log(\beta) + \left(\frac{1}{\psi} - \lambda\right)\mu_t - \frac{1}{2}\,\sigma^2\left(\gamma - \frac{1-\gamma}{\psi} + \lambda\,(\lambda - 2\,\gamma)\right),\\
rp_t &= \lambda\,\gamma\,\sigma^2.
\end{aligned}$$

The expected risk premium under the econometrician's objective expectations operator is

$$\mathbb{E}\left(r_{t+1}\right) - r_t^f = \left(\frac{1}{1-\bar{p}}\,\lambda - \frac{\bar{p}}{1-\bar{p}}\frac{1}{\psi}\right)(\mu - \mu_t) + rp_t - \frac{1}{2}\,\lambda^2\,\sigma^2.$$

Intuitively, the first term comes from the expected revision of beliefs under the econometrician's filtration. If $\lambda > \bar{p}\,\frac{1}{\psi}$, the expected risk premium will be low when the agent is too optimistic ($\mu_t > \mu$). Additionally, the expected risk premium depends on the subjective risk premium under which the agent priced the asset, and a Jensen's inequality adjustment.

Case 2: Two-state Markov process

Next, let us consider endowment growth as in the main text,

$$g_t = \mu_s + \sigma_s\,\epsilon_t,$$

where $s_t \in \{1, 2\}$ follows a two-state observable Markov chain with constant transition probabilities that ensure that endowment growth is i.i.d. Equation 14 in the main text gives the cumulant-generating function under the agent's beliefs. The momemt-generating function under the agent's beliefs follows from $\mathcal{M}_t(k) = \exp\left[\mathcal{K}_t(k)\right]$, and Equation 27 gives the expected moment-generating function under the econometrician's beliefs. Thus, the asset-pricing quantities under Epstein-Zin preferences are given as above.

Moreover, let us use the second-order approximation of the expected cumulant-generating function. It is

$$\mathcal{M}_{t+1}(k)^2 = \pi_1^2\,e^{2\,k\,\hat{\mu}_{1,t+1} + k^2\,\sigma_1^2} + \pi_2^2\,e^{2\,k\,\hat{\mu}_{2,t+1} + k^2\,\sigma_2^2} + 2\,\pi_1\,\pi_2\,e^{k\,(\hat{\mu}_{1,t+1} + \hat{\mu}_{2,t+1}) + \frac{1}{2}k^2\left(\sigma_1^2 + \sigma_2^2\right)},$$

and I find the objective expectation as

$$\mathbb{E}\left(\mathcal{M}_{t+1}(k)^2\right) = \pi_1\,\mathbb{E}\left[\mathcal{M}_{t+1}(k)^2 | s_{t+1} = 1\right] + \pi_2\,\mathbb{E}\left[\mathcal{M}_{t+1}(k)^2 | s_{t+1} = 2\right],$$

with

$$\begin{aligned}
\mathbb{E}\left[\mathcal{M}_{t+1}(k)^2|s_{t+1}=1\right] =& \pi_1^2\,\left[e^{2\,k\,\mu_1+k^2\,\sigma_1^2\,(1+2\,\alpha_1^2)}\right]+\pi_2^2\,\left[e^{2\,k\,[(1-\alpha_2)\,\mu_2+\alpha_2\,\mu_1]+k^2\,\sigma_2^2+2\,k^2\,\alpha_2^2\sigma_1^2}\right] \\
&+2\,\pi_1\,\pi_2\left[e^{k\,[(1+\alpha_2)\,\mu_1+(1-\alpha_2)\,\mu_2]+\frac{1}{2}k^2\left(\sigma_1^2+\sigma_2^2\right)+\frac{1}{2}\,k^2\,(\alpha_1+\alpha_2)^2\,\sigma_1^2}\right] \\
\mathbb{E}\left[\mathcal{M}_{t+1}(k)^2|s_{t+1}=2\right] =& \pi_1^2\,\left[e^{2\,k\,[(1-\alpha_1)\,\mu_1+\alpha_1\,\mu_2]+k^2\,\sigma_1^2+2\,k^2\,\alpha_1^2\sigma_2^2}\right]+\pi_2^2\,\left[e^{2\,k\,\mu_2+k^2\,\sigma_2^2\,(1+2\,\alpha_2^2)}\right] \\
&+2\,\pi_1\,\pi_2\left[e^{k\,[(1-\alpha_1)\,\mu_1+(1+\alpha_1)\,\mu_2]+\frac{1}{2}k^2\left(\sigma_1^2+\sigma_2^2\right)+\frac{1}{2}\,k^2\,(\alpha_1+\alpha_2)^2\,\sigma_2^2}\right].
\end{aligned}$$

We can then insert the expressions into the objectively expected risk premium to derive numerical approximations of the expected risk premium, as done to construct Figure 2.

G Estimation and simulation procedures used in Section 3.4

G.1 Data and estimation

In this section, I describe the data and estimation procedure that I use to obtain the parameters for the simulations. The parameter values are given in Table 1.

The data used to estimate the parameters of endowment growth is the quarterly nominal consumption (nondurable and service) from BEA's Table 7.1 from Q1 1947 until Q1 2023. I transform the nominal data to real endowment growth taking the chain-weighted Tornqvist index of BEA's data into account. In addition, I use dividends to estimate the leverage parameter λ. I obtain aggregate quarterly dividends using the lagged total market value of the CRSP value-weighted index and the difference between returns without and with dividends. I deflate dividends using the Consumer Price Index (CPI) series in Shiller's data. The average annual endowment growth is 1.77% (4.22% for dividend growth), and the volatility of endowment growth is 1.89%.

Next, I estimate the parameters of the endowment growth process using Bayesian methods similar to Johannes et al. (2016). I assume a conjugate normal/inverse gamma prior for endowment growth in each state:

$$
\begin{aligned}
p(\mu_i, \sigma_i^2) &\sim \mathcal{NIG}(a_i, A_i, b_i/2, B_i/2) \\
p(\mu_i|\sigma_i^2) &\sim \mathcal{N}(a_i, A_i\, \sigma_1^2) \\
p(\sigma_i^2) &\sim \mathcal{IG}(b_i/2, B_i/2),
\end{aligned}
$$

and set the parameters of these distributions as

$$
\begin{aligned}
E(\mu_i) &= a_i \\
\mathrm{Var}(\mu_i) &= A_i\, \frac{B_i}{b_i - 2} = A_i\, E(\sigma_i^2),
\end{aligned}
$$

since the marginal distribution of μ_i is a scaled student-t distribution with $p(\mu_i) \sim t_{b_i}(a_i, A_i \frac{B_i}{b_i})$. The moments of the inverse-gamma distribution are

$$
\begin{aligned}
E(\sigma_i^2) &= \frac{B_i/2}{b_i/2 - 1} \\
\mathrm{Var}(\sigma_i^2) &= \frac{B_i/2}{(b_i/2 - 1)^2\, (b_i/2 - 2)} = E(\sigma_i^2)^2\, \frac{1}{b_i/2 - 2}.
\end{aligned}
$$

Thus, I find the parameters as follows:

$$
\begin{aligned}
a_i &= E(\mu_i) \\
A_i &= \frac{\mathrm{Var}(\mu_i)}{E(\sigma_i^2)} \\
b_i &= 2\, \frac{E(\sigma_i^2)^2}{\mathrm{Var}(\sigma_i^2)} + 4 \\
B_i &= E(\sigma_i^2)\, (b_i - 2)\,.
\end{aligned}
$$

In addition, I assume that the transition probabilities are independent of the parameters of endowment growth in each state and given by a Beta-distribution with $p(\pi_1) \sim \mathcal{B}(c_1, C_1)$. It is

$$
\begin{aligned}
E(\pi_1) &= \frac{c_1}{c_1 + C_1} \\
\mathrm{Var}(\pi_1) &= \frac{c_1\, C_1}{(c_1 + C_1)^2\, (C_1 + c_1 + 1)} \\
&= E(\pi_1)\ (1 - E(\pi_1))\ \frac{1}{C_1 + c_1 + 1} \\
&= E(\pi_1)\ (1 - E(\pi_1))\ \frac{1}{\frac{1}{E(\pi_1)}\, c_1 + 1}
\end{aligned}
$$

such that I find

$$
\begin{aligned}
c_1 &= \frac{E(\pi_1)^2\ (1 - E(\pi_1))}{\mathrm{Var}(\pi_1)} - E(\pi_1) \\
C_1 &= c_1 \left(\frac{1 - E(\pi_1)}{E(\pi_1)} \right).
\end{aligned}
$$

Table G.1 shows the parameters used for the estimation, which are close to the parameters used in Johannes et al. (2016) while imposing the restriction to i.i.d. endowment growth. Using the prior parameters given in Table G.1, I use a Markov-Chain-Monte-Carlo (MCMC) procedure to estimate the parameters of endowment growth.

Table G.1: Prior parameters for estimation

Parameter	Mean	St.Dev
μ_1	0.90%	0.17%
μ_2	0.00%	0.87%
σ_1^2	$(0.49\%)^2$	$(0.29\%)^2$
σ_2^2	$(2.89\%)^2$	$(1.49\%)^2$
π_1	95.40%	3.40%

Table G.1 reports parameters of the priors used to estimate the properties of an i.i.d. two state Markov-switching process for endowment growth. The values are chosen to match the values in Johannes et al. (2016).

Intuitively, the MCMC is solving the conjugate Bayesian posterior with the prior distributions given above. The algorithm iteratively varies the parameters of the model, and computes the log-likelihood of the posterior on the BEA endowment growth data. I then select the model that has the highest log-likelihood over 10,000 iterations, which corresponds to the Bayeisan maximum a posteriori (MAP) estimate. The parameters of the MAP are given in Table 1.

Finally, I estimate the leverage parameter λ by regressing the aggregate log dividend growth on the aggregate log endowment growth. We assumed that, for any asset, $D_t = C_t^\lambda$, such that

$$\frac{D_{t+1}}{D_t} = \left(\frac{C_{t+1}}{C_t}\right)^\lambda,$$

which implies

$$\log\frac{D_{t+1}}{D_t} = \lambda\, g_{t+1},$$

and I consequently run the regression

$$\log\frac{D_{t+1}}{D_t} = a + b\, g_{t+1} + \epsilon_{t+1}.$$

Empirically, I find $\hat{b} = 3.29$, which is close to the parameters used in the literature. Collin-Dufresne et al. (2016) and Nagel and Xu (2022) use a leverage parameter $\lambda = 3$ under a different dividend-growth process, such that I choose $\lambda = 3$ in simulations for comparability.

G.2 Parameter uncertainty

In this section, I highlight how parameter uncertainty emerges and how it affects the asset pricing results. Since closed-form solutions exist for log-normal endowment growth, it is instructive to analyze this case first. Thereafter, I outline how I simulate the asset pricing quantities for the two-state Markov-switching process analyzes in Section 3.

Case 1: Log-normally distributed endowment growth

Let consumption growth g_t be given by

$$g_t = \mu + \sigma\,\epsilon_t \quad \epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}(0,1).$$

The agent does not know the mean growth rate μ, but must learn it from the recalled observations. In any period t, she recalls $|H_t^R| = k_t$ past observations of endowment growth. Her prior for the mean endowment growth is $\mu \sim \mathcal{N}\left(\mu_0, \frac{\sigma^2}{\nu}\right)$, where ν scales the informativeness of the prior. The Bayesian posterior of the agent is then given by

$$\mu \sim \mathcal{N}\left(\mu_t, z_t\sigma^2\right),$$

where

$$\begin{aligned} z_t^{-1} &= k_t + \nu \\ \mu_t &= \frac{1}{z_t^{-1}}\left(\nu\,\mu_0 + \sum_{\tau \in r_t} \Delta g_\tau\right). \end{aligned}$$

The agent is naïve with respect to her memory distortions and thus believes that she will surely recall $k_t + 1$ observations next period. The agent's perceived belief and endowment growth dynamics are thus

$$\begin{aligned} g_{t+1} &= \mu_t + \sqrt{1+z_t}\,\sigma\,\tilde{\epsilon}_{t+1} \\ \epsilon_{t+1} &= \frac{\Delta c_{t+1} - \mu_t}{\sqrt{1+z_t}} \\ z_{t+1}^{-1} &= z_t^{-1} + 1 \\ \mu_{t+1} &= \mu_t + \frac{z_t}{\sqrt{1+z_t}}\,\sigma\,\tilde{\epsilon}_{t+1}. \end{aligned}$$

We can derive asset-prices under parameter uncertainty in closed form for $\psi = 1$. Using the value function iteration as in Hansen et al. (2008), the log of the wealth-consumption ratio $wc_t = \log(W_t/C_t)$ is

$$wc_t = \frac{\beta}{1-\gamma}\,\log\left(\tilde{\mathbb{E}}_t\, e^{(1-\gamma)(vc_{t+1}+g_{t+1})}\right).$$

I conjecture that $wc_t = a_t + B\,\mu_t$, as in Collin-Dufresne et al. (2016), which yields

$$\begin{aligned} wc_t &= \frac{\beta}{1-\gamma}\,\log\left(\tilde{\mathbb{E}}_t\, e^{(1-\gamma)(a_{t+1}+B\,\mu_{t+1}+g_{t+1})}\right) \\ &= \frac{\beta}{1-\gamma}\,\log\left(\tilde{\mathbb{E}}_t\, e^{(1-\gamma)\left(a_{t+1}+(B+1)\,\mu_t+\left(B\,\frac{z_t}{\sqrt{1+z_t}}+\sqrt{1+z_t}\right)\sigma\tilde{\epsilon}_{t+1}\right)}\right) \\ &= \beta\left(a_{t+1}+(B+1)\,\mu_t+\frac{1}{2}\,(1-\gamma)\,\left(B\,\frac{z_t}{\sqrt{1+z_t}}+\sqrt{1+z_t}\right)^2\sigma^2\right). \end{aligned}$$

Thus, I find that

$$\begin{aligned} B &= \frac{\beta}{1-\beta}, \\ a_t &= \beta a_{t+1} + \frac{1}{2}\beta(1-\gamma)\,\left((B+1)z_t+1\right)^2\frac{1}{1+z_t}\,\sigma^2 \\ &= \sum_{j=0}^{\infty}\beta^{j+1}\,\frac{1}{2}\beta(1-\gamma)\,\left((B+1)z_{t+j}+1\right)^2\frac{1}{1+z_{t+j}}\,\sigma^2. \end{aligned}$$

The log-SDF in the $\psi = 1$-case is then

$$\begin{aligned} m_{t+1} =& \log\left(\beta\left(\frac{C_t}{C_{t+1}}\right)\frac{V_{t+1}^{1-\gamma}}{\tilde{\mathbb{E}}_t\left[V_{t+1}^{1-\gamma}\right]}\right) \\ =& \log\left(\beta e^{-\Delta c_{t+1}}\,\frac{e^{((1-\gamma)\,(vc_{t+1}+g_{t+1}))}}{\tilde{\mathbb{E}}_t\left(e^{((1-\gamma)\,(vc_{t+1}+g_{t+1}))}\right)}\right) \\ =& \log\left(\beta e^{-\Delta c_{t+1}}\,\frac{e^{\left((1-\gamma)\,\left(B\,\mu_{t+1}+\sqrt{1+z_t}\,\sigma\,\tilde{\epsilon}_{t+1}\right)\right)}}{\tilde{\mathbb{E}}_t\left(e^{\left((1-\gamma)\,\left(B\,\mu_{t+1}+\sqrt{1+z_t}\,\sigma\,\tilde{\epsilon}_{t+1}\right)\right)}\right)}\right) \\ =& \log(\beta)-\mu_t-\sqrt{1+z_t}\,\sigma\,\tilde{\epsilon}_{t+1}+(1-\gamma)\,\left((B+1)\,z_t+1\right)\frac{1}{\sqrt{1+z_t}}\,\sigma\,\tilde{\epsilon}_{t+1} \\ &-\frac{1}{2}\,(1-\gamma)^2\,\left((B+1)\,z_t+1\right)^2\,\frac{1}{1+z_t}\,\sigma^2 \end{aligned}$$

$$= \mu_{m,t} - \mu_t - \zeta_t \, \sigma \, \tilde{\epsilon}_{t+1},$$

where

$$\mu_{m,t} = \log(\beta) - \frac{(1-\gamma)^2}{2}((B+1)\, z_t + 1)^2 \, \frac{1}{1+z_t}\sigma^2,$$
$$\zeta_t = [(1+z_t) - (1-\gamma)\; ((B+1)\, z_t + 1)]\; \frac{1}{\sqrt{1+z_t}}.$$

Shocks to the log SDF are thus

$$m_{t+1} - \tilde{\mathbb{E}}_{\mathrm{t}}(m_{t+1}) = -\zeta_t, \sigma \, \tilde{\epsilon}_t,$$

and the price of risk—defined as the conditional volatility of the log SDF—is given by $\zeta_t \sigma > \gamma \, \sigma$, which is the price of risk without parameter uncertainty.

Joint log-normality of endowment growth and the SDF then gives

$$0 = \tilde{\mathbb{E}}_{\mathrm{t}}(m_{t+1}) + \tilde{\mathbb{E}}_{\mathrm{t}}(r_{c,t+1}) + \frac{1}{2}\,\mathrm{Var}_t(m_{t+1}) + \frac{1}{2}\mathrm{Var}_t(r_{c,t+1}) + \mathrm{Cov}_t(m_{t+1}, r_{c,t+1}),$$

where $r_{c,t+1}$ is the log-return on the consumption-claim. Note that an EIS of $\psi = 1$ yields a constant wealth-consumption ratio, $WC = \frac{\beta}{1-\beta}$, and the log-return on the consumption claim is

$$r_{c,t+1} = \log\left(\frac{C_{t+1}}{C_t}\,(1 + WC^{-1})\right) = g_{t+1} - \log(\beta)$$

The expected log-return on the consumption-claim is

$$\tilde{\mathbb{E}}_{\mathrm{t}}(r_{c,t+1}) = \mu_t - \log(\beta) = -\,\tilde{\mathbb{E}}_{\mathrm{t}}(m_{t+1}) - \frac{1}{2}\mathrm{Var}_t(m_{t+1}) - \frac{1}{2}\mathrm{Var}_t(r_{c,t+1}) - \mathrm{Cov}_t\,(m_{t+1}, r_{c,t+1})$$

The log risk-free rate is

$$
\begin{aligned}
r_t^f &= -\tilde{\mathbb{E}}_{\mathrm{t}}(m_{t+1}) - \frac{1}{2}\operatorname{Var}_t(m_{t+1}) \\
&= \mu_t - \log(\beta) - \frac{1}{2}(1+z_t)\,\sigma^2 + (1-\gamma)\left[(B+1)z_t + 1\right]\sigma^2,
\end{aligned}
$$

and the risk premium on the consumption-claim is

$$
\begin{aligned}
\tilde{\mathbb{E}}_{\mathrm{t}}(r_{c,t+1}) - r_t^f &= -\frac{1}{2}\operatorname{Var}_t(r_{c,t+1}) - \operatorname{Cov}_t(m_{t+1}, r_{c,t+1}) \\
&= -\frac{1}{2}(1+z_t)\,\sigma^2 + \left[(1+z_t) + (\gamma-1)((B+1)z_t+1)\right]\sigma^2 \\
&= (\gamma-1)\,B\,z_t\sigma^2 + \left(\gamma - \frac{1}{2}\right)(z_t+1)\,\sigma^2.
\end{aligned}
$$

Alternatively, and as in the main text, let us consider the expected return on the consumption-claim as

$$
\tilde{\mathbb{E}}_{\mathrm{t}}\left(R_{c,t+1}\right) = \tilde{\mathbb{E}}_{\mathrm{t}}\left(e^{r_{c,t+1}}\right) = \frac{1}{\beta}e^{\mu_t + \frac{1}{2}(1+z_t)\sigma^2},
$$

and the log of the expected return is

$$
er_c = \log(\tilde{\mathbb{E}}_{\mathrm{t}}\left(R_{c,t+1}\right)) = \tilde{\mathbb{E}}_{\mathrm{t}}\left(e^{r_{c,t+1}}\right) = -\log(\beta) + \mu_t + \frac{1}{2}(1+z_t)\sigma^2.
$$

The risk premium on the consumption-claim is then

$$
\begin{aligned}
er_c - r_t^f &= -\log(\beta) + \mu_t + \frac{1}{2}(1+z_t)\sigma^2 - \mu_t + \log(\beta) + \frac{1}{2}(1+z_t)\,\sigma^2 - (1-\gamma)\left[(B+1)z_t+1\right]\sigma^2 \\
&= (1+z_t)\sigma^2 - (1-\gamma)\left[(B+1)z_t+1\right]\sigma^2.
\end{aligned}
$$

Next, let us derive objective risk premium by taking the expectation of the return on the consumption-claim. The econometrician knows the true underlying process, such that

$\mathbb{E}(R_{c,t+1}) = \frac{1}{\beta}\, e^{\mu+\frac{1}{2}\,\sigma^2}$ and the objective risk premium is

$$\log(\mathbb{E}(R_c)) - r_t^f = \underbrace{(\mu - \hat{\mu}_t)}_{\text{Belief wedge}} + \underbrace{(1 + z_t)\ \sigma^2 - (1-\gamma)\,[(B+1)z_t + 1]\ \sigma^2}_{\text{Subjective risk premium}} \quad \underbrace{-\frac{1}{2}\, z_t\, \sigma^2}_{\text{Jensen's inequality}}$$

The objective risk premium depends on three components: First, the wedge between the true mean endowment growth and the agent's expectation, $(\mu - \hat{\mu})$. Intuitively, if the agent's posterior mean $\hat{\mu}_t$ is too high, the agent drives up the price of the asset and objective returns next period will be low. The second component is the agent's subjective risk premium, which determines prices in equilibrium and thus therewith affect expected returns. The third component is a Jensen's inequality adjustment. We can similarly derive the objective risk premium on the dividend-paying asset.

Case 2: Markov-process

Let us now consider the Markov process for endowment growth as in Equation 10. Closed-form solutions for asset prices with parameter uncertainty cease to exist, such that I detail the numerical procedure to obtain asset pricing equations in this Appendix. The procedure follows Collin-Dufresne et al. (2016).

The agent knows the state-dependent variance σ_s^2, but must learn the state-dependent means μ_s from her recalled history of log endowmentgrowth. In period t, the agent recalls $|H_{1,t}^R| = k_{1,t}$ endowment growth observations from state 1 and $|H_{2,t}^R| = k_{2,t}$ observations from state 2 and forms a Bayesian posterior about the mean in each state. The agent has conjugate, normally distributed prior beliefs about the state-dependent mean growth-rates, $\mu_s \sim \mathcal{N}\left(\hat{\mu}_{s,0}, \frac{\sigma_s^2}{\nu_s}\right)$, where ν_s scales the informativeness of the prior. The agent's posterior upon recalling the state-dependent history $H_{s,t}^R$ is

$$\mu \sim \mathcal{N}\left(\hat{\mu}_{s,t}, z_{s,t}\, \sigma_s^2\right),$$

with

$$z_{s,t} = \left(k_{s,t} + \nu_s\right)^{-1},$$
$$\hat{\mu}_{s,t} = z_{s,t}\left(\nu_s\,\hat{\mu}_{s,0} + \sum_{\tau \in H^R_{s,t}} g_\tau\right).$$

The agent's perceived dynamics of endowment and her beliefs are[37]

$$g_{t+1} = \hat{\mu}_{s,t} + \sqrt{1 + z_{s,t}}\,\sigma_s\,\tilde{\epsilon}_{t+1}$$
$$z^{-1}_{t+1,s} = z^{-1}_{s,t} + \mathbb{1}_{s_{t+1}=s}$$
$$\hat{\mu}_{t+1,s} = \mu_{s,t} + \mathbb{1}_{s_{t+1}=s}\frac{z_{s,t}}{1+z_{s,t}}\left(g_{t+1} - \mu_{s,t}\right),$$

where $\mathbb{1}_{a=b}$ equals one if the condition in subscript is true and the belief about the state that does not occur in the next period is not updated. The state variables that describe the agent's beliefs are $X_t \equiv [\mu_{1,t}, \mu_{2,t}, k_{1,t}, k_{2,t}]$, and the state of the Markov chain s_t is an additional state variable of the economy.

The agent has Epstein-Zin preferences, such that the SDF (for $\psi \neq 1$) is

$$M_{t+1} = \beta^\theta \left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\theta}{\psi}} R^{\theta-1}_{w,t+1},$$

where $\theta = \frac{1-\gamma}{1-\frac{1}{\psi}}$ is a composite parameter and $R_{w,t+1} = \frac{W_{t+1}+C_{t+1}}{C_t}$ is the return on wealth. The return on wealth is determined in equilibrium as

$$\tilde{\mathbb{E}}_{\mathrm{t}}\left[\beta^\theta \left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\theta}{\psi}} R^{\theta}_{w,t+1}\right] = 1,$$

[37] I relate the discussion thus far to the memory model in the main text as follows: In each period, the recalled experiences are drawn from selective memory, but the agent forms beliefs under naïvete, such that her perceived endowment and belief dynamics do not take her memory distortion into account. Thus, in period t, the agent thinks that she is a rational Bayesian and forecasts her belief evolution consistently.

which, when inserting the expression for the return on wealth, yields

$$\left(\frac{W_t}{C_t}\right)^{\theta} = \beta^{\theta}\, \tilde{\mathbb{E}}_{\mathrm{t}}\left[e^{(1-\gamma)\, g_{t+1}}\, \left(\frac{W_{t+1}}{C_{t+1}} + 1\right)^{\theta}\right].$$

Note that the wealth-consumption ratio at time t is a function of the state variables at time t. Writing $\frac{W_{t+1}}{C_{t+1}} = WC_{t+1}$, it is $WC_{t+1} = WC(X_{t+1}, s_{t+1}) = WC(X_t, s_{t+1}, g_{t+1})$, where the last step clarifies that the evolution of the state variables under the agent's beliefs depends on next period's state and on the realized endowment growth.

Under a two-state Markov process with known transition probabilities, the expression for the wealth-consumption ratio can be rewritten as

$$\begin{aligned} WC(X_t, s_t)^{\theta} = {} & \beta^{\theta}\, \pi_1\, \tilde{\mathbb{E}}_{\mathrm{t}}\left(e^{(1-\gamma)\, g_{t+1}}\, \left(WC(X_t, s_{t+1}, g_{t+1}) + 1\right)^{\theta} | s_{t+1} = 1, s_t, X_t\right) + \\ & \beta^{\theta}\, \pi_2\, \tilde{\mathbb{E}}_{\mathrm{t}}\left(e^{(1-\gamma)\, g_{t+1}}\, \left(WC(X_t, s_{t+1}, g_{t+1}) + 1\right)^{\theta} | s_{t+1} = 2, s_t, X_t\right), \end{aligned}$$

where I separate the expectation using the law of iterated expectations. For the conditional inner expectations, we do not have closed-form solutions. The expression needs to be evaluated numerically, and I proceed as follows: As a first step, I find the wealth-consumption ratio for the known parameters case with $z_t = 0$. Note that (perceived) endowment growth is i.i.d., such that I can use the results from the main text to obtain:

$$WC_{\infty} = \frac{\beta\, e^{\frac{1}{\theta}\, \mathcal{K}(1-\gamma)}}{1 - \beta\, e^{\frac{1}{\theta}\, \mathcal{K}(1-\gamma)}},$$

where $\mathcal{K}(m) = \log \tilde{\mathbb{E}}_{\mathrm{t}}\left(e^{m\, g_{t+1}}\right)$ is the cumulant-generating function under the agent's beliefs.

As a second step, I solve for the boundary case where one mean is known (no parameter uncertainty) and the other mean is unknown. Let us assume that the agent has no parameter uncertainty around $\hat{\mu}_{1,\infty}$ and thus she does not learn when state 1 realizes. The wealth-consumption ratio is then

$$WC(X_t,s_t)^\theta = \beta^\theta\, \pi_1\, e^{(1-\gamma)\hat{\mu}_{1,\infty}+\frac{1}{2}(1-\gamma)^2\,\sigma_1^2}\, \left(WC(X_t,s_t)+1\right)^\theta +$$
$$\beta^\theta\, \pi_2\, \tilde{\mathbb{E}}_{\mathrm{t}}\left(e^{(1-\gamma)\,g_{t+1}}\, \left(WC(X_t,s_{t+1},g_{t+1})+1\right)^\theta \,|s_{t+1}=2,s_t,X_t\right).$$

We need to integrate out two sources of uncertainty under the agent's belief: The noise in endowment growth $\tilde{\epsilon}_{t+1}$ and the agent's uncertainty about her posterior mean $\hat{\mu}_{2,t}$ if state 2 occurs. I iterate backwards from the known parameter case and use a Gauss-Hermite quadrature to approximate the expectation. The numerical approximation for the expectation is

$$\tilde{\mathbb{E}}_{\mathrm{t}}\left(e^{(1-\gamma)\,g_{t+1}}\, \left(WC(X_t,s_{t+1},g_{t+1})+1\right)^\theta \,|s_{t+1},s_t,X_t\right)$$
$$\approx \sum_{j=1}^{J}\omega_\epsilon(j)\sum_{k=1}^{K}\omega_{\mu_{2,t}}(k)\left(e^{(1-\gamma)\,g_{t+1}}\, \left(WC(X_t,s_{t+1},g_{t+1})+1\right)^\theta \,|s_{t+1}=2,s_t,X_t\right),$$

where $w_\epsilon(j)$ is the quadrature weight for the standard-normal variable $\tilde{\epsilon}_{t+1}$, corresponding to the quadrature point $n_\epsilon(j)$, and $\omega_{\mu_{2,t}}(k)$ is the quadrature weight for the normally distributed posterior mean corresponding to quadrature point $n_{\mu_{2,t}}(k)$. The realized endowment growth in state 2 is then given by

$$g_{t+1}(k,j) = n_{\mu_{2,t}}(k) + \sigma_s\, n_\epsilon(j),$$

since the uncertainty about the mean that affects the perceived endowment growth is integrated out. Having solved for the *inner expectation*, I find $WC(X_t,s_t)$ as the fixed-point of the non-linear equation above.

As a third step, I iterate backwards from the boundary cases using the same quadrature-type method to approximate the agent's expectation. Since I find both inner expectations numerically, I do not need to solve for a fixed-point in order to find $WC(X_t,s_t)$.

Similarly, I can obtain the prices of dividend-paying assets. Recall that the return on any asset is, by definition,

$$R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{D_t}{P_t}\,\frac{D_{t+1}}{D_t}\,\left(\frac{P_{t+1}}{D_{t+1}} + 1\right) = e^{\lambda\, g_{t+1}}\,\frac{PD(X_t, s_{t+1}, g_{t+1}) + 1}{PD(X_t, s_t)},$$

where I used $\frac{P_{t+1}}{D_{t+1}} = PD(X_{t+1}, s_{t+1}) = PD(X_t, s_{t+1}, g_{t+1})$, as before. In equilibrium, we find the return on any asset as

$$\begin{aligned}
1 &= \tilde{\mathbb{E}}_{\mathrm{t}}\left[M_{t+1}\,R_{t+1}\right] \\
&= \tilde{\mathbb{E}}_{\mathrm{t}}\left[\beta^{\theta}\,\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\theta}{\psi}}\,R_{w,t+1}^{\theta-1}\,R_{t+1}\right].
\end{aligned}$$

Inserting, we thus can write the price-didend ratio of any asset as

$$\begin{aligned}
PD(X_t, s_t) =& \beta^{\theta}\,\tilde{\mathbb{E}}_{\mathrm{t}}\left[e^{(\lambda-\gamma)\,g_{t+1}}\,\left(\frac{WC(X_{t+1}, s_{t+1}) + 1}{WC(X_t, s_t)}\right)^{\theta-1}\,\left(PD(X_t, s_{t+1}, g_{t+1}) + 1\right)\right] \\
=& \beta^{\theta}\,\pi_1\,\tilde{\mathbb{E}}_{\mathrm{t}}\left[e^{(\lambda-\gamma)\,g_{t+1}}\,\left(\frac{WC(X_{t+1}, s_{t+1}) + 1}{WC(X_t, s_t)}\right)^{\theta-1}\,\left(PD(X_t, s_{t+1}, g_{t+1}) + 1\right)|s_{t+1} = 1\right] + \\
& \beta^{\theta}\,(1-\pi_1)\,\tilde{\mathbb{E}}_{\mathrm{t}}\left[e^{(\lambda-\gamma)\,g_{t+1}}\,\left(\frac{WC(X_{t+1}, s_{t+1}) + 1}{WC(X_t, s_t)}\right)^{\theta-1}\,\left(PD(X_t, s_{t+1}, g_{t+1}) + 1\right)|s_{t+1} = 2\right].
\end{aligned}$$

We can thus solve for the price-dividend ratio of any asset in the same way as we did for the wealth-consumption ratio.

In the main text, I analyzed the following asset pricing quantities under the agent's subjective beliefs:

$$\begin{aligned}
er_t &= \log\left(\tilde{\mathbb{E}}_{\mathrm{t}}\;R_{t+1}\right) \\
r_t^f &= \log\left(\tilde{\mathbb{E}}_{\mathrm{t}}\;R_{t+1}^f\right) \\
rp_t &= er_t - r_t^f,
\end{aligned}$$

as well as the following objective quantity

$$rp_t^o = \log\left(\mathbb{E}\, R_{t+1}\right) - r_t^f.$$

Using the wealth-consumption ratio and the price-dividend ratio as above, we can obtain the asset pricing quantities as follows:

$$\begin{aligned} r_t^f &= \log\left[\tilde{\mathbb{E}}_{\mathrm{t}}\left(\frac{PD(X_{t+1}, s_{t+1}|\lambda = 0) + 1}{PD(X_t, s_t|\lambda = 0)}\right)\right], \\ er_t &= \log\left[\tilde{\mathbb{E}}_{\mathrm{t}}\left(e^{\lambda\, g_{t+1}}\,\frac{PD(X_{t+1}, s_{t+1}) + 1}{PD(X_t, s_t)}\right)\right], \end{aligned}$$

where we obtain the price-dividend ratio of the riskless asset as above, and need to numerically approximate the expected return under the agent's beliefs using the same methods as before. The subjective risk premium is then found as the difference between the log expected return and the risk-free rate. Finally, I obtain the objective risk premium from the realized asset returns (objective expectations equal the average realized return). I simulate the endowment growth process multiple times. Having obtained the price-dividend ratio above, I can then compute the price of the asset in period t as $PD(X_t, s_t)\, D_t = PD(X_t, s_t)\, C_t^{\lambda}$. The realized return is found using the definition of the return as $R_{t+1} = \frac{P_{t+1}+D_{t+1}}{P_t}$.

Online Appendix

OA.1 Selective memory for continuous distributions

In this Online Appendix, I introduce the learning framework of Section 2 for the case of a continuous outcome distribution.

Economy. Let us assume that the realized signal $s_t = s$ induces a fixed and i.i.d. density q_s^* of log endowment growth g, such that $Pr\left(g \in [a, b]\right) = \int_a^b q_s^*(g)\, dg$ conditional on $s_t = s$. The density $q_s^* \in \mathcal{D}$, where $\mathcal{D}$ is the set of densities over $\mathbb{R}$.[1] I maintain the assumption that q_s^* belongs to the family of parametric probability densities, $q_s^* \in \{q_\theta : \theta \in \Theta\}$, with $\Theta \subseteq \mathbb{R}^k, k \in \mathbb{N}$ closed and convex.

Learning. To model uncertainty about the distribution of log endowment growth, I assume that the agent holds a prior belief b_0 over potential densities $q \in \mathcal{D}^S$, where $\int_a^b q_s(g)\, dg$ gives the probability of observing $g_t \in [a, b]$ under density q_s, and q assigns one density to every signal realization $s \in S$. The support of the prior contains all distributions that the agent initially considers possible. I assume that the agent knows that log endowment growth is generated by a parametric distribution, such that the prior support $Q \subseteq \{q_\theta : \theta \in \Theta\}^S \subset \mathcal{D}^S$. The assumptions on the prior from Section 2 continue to hold.[2]

Memory. The assumptions on the memory function remain as in the main text. The memory-function is applied to the densities, and $m_{(g_t, s_t)} : \mathbb{R} \times S \mapsto [0, 1]$.

Beliefs The agent forms Bayesian beliefs conditional on her recalled experiences, and Equation (1) determines the agent's beliefs.

Define the continuous *memory-weighted* likelihood maximizer conditional on this period's

[1]Formally, let us consider the probability space $(\Omega, \mathcal{F}, P)$ and the measurable space $(\mathcal{A}, \mathcal{B})$, with $\mathcal{A} \subseteq \mathbb{R}$ and $\mathcal{B}$ the respective Borel σ-algebra. Endowment growth is a measurable function that maps from Ω to $\mathcal{A}$, $g_t : \Omega \mapsto \mathcal{A}$. The density q_s^* is then constructed from the probability measure assigned to the preimage of each interval $[a, b]$ under g_t as $P(g_t^{-1}((a, b))) = \int_a^b q_s^*(g_t)\, dg_t$, the image measure. The set of all densities $\mathcal{D}$ is the set of all measureable functions $\xi : \Omega \mapsto \mathbb{R}$ that are non-negative almost everywhere and satisfy $\int_\Omega \xi(x)\, dx = 1$.

[2]The agent is correctly specified, $q^* \in Q$ and all measures in the prior support are mutually absolutely continuous. Consequently, each measure in the prior support can be obtained from any other measure by a Radon-Nikodym derivative.

experience as

$$\mathrm{LM}^c(g_t, s_t) = \operatorname*{argmax}_{q \in Q} \sum_{s \in S} \psi(s) \int_{-\infty}^{\infty} m_{(g_t,s_t)}(g,s)\, q_s^*(g)\, \log q_s(g)\, dg.$$

The following proof shows that the agent's belief concentrates on data-generating processes that maximize the likelihood of the recalled history as given by $\mathrm{LM}^c(g_t, s_t)$.

Proof. The proof follows the arguments presented in Fudenberg et al. (2023) and proceeds as follows: First, I show that the histogram of the agent's recalled experiences converges to the memory-weighted true probability density. Second, following Berk (1966), I argue that the agent's Bayesian posterior concentrates on maximizers of the (log-)likehood. Last, I show that the recalled history is almost surely large, such that the convergence results are meaningful. Combining those steps yields Proposition .

Step 1: Recall the notation. The history of experiences is $H_t = \{(g_\tau, s_\tau)\}_{\tau=-\infty}^{t}$. The agent recalls the experience from period $\tau \leq t$ with probability $m_{(g_s,t_t)}(g_\tau, s_\tau) \in [0,1]$. The recalled periods r_t are therefore a random subset of all experiences that occurred with distribution $\mathbb{P}\left[r_t | H_t, g_t, s_t\right] = \prod_{\tau \in r_t} m_{(g_s,t_t)}(g_\tau, s_\tau) \prod_{\tau \notin r_t} \left(1 - m_{(g_s,t_t)}(g_\tau, s_\tau)\right)$. Define the empirical joint distribution function of recalled growth rates and signals as

$$\hat{F}_t(g,s) = \frac{1}{|H_t^R|} \sum_{\tau \in r_t} \mathbb{1}\{g_\tau \leq g, s_\tau \leq s\},$$

while the true joint distribution function of experiences is given by $F(g,s)$. Without memory selectivity, $m_{(g_s,t_t)}(g_\tau, s_\tau) = c \in [0,1]\, \forall (g_\tau, s_\tau)$, the Glivenko-Cantelli lemma[3] ensures uniform almost sure convergence of the empirical joint distribution, $\hat{F}_t(g,s)$, to the true distribution, $F(g,s)$ as $t \to \infty$:

$$\sup_{g \in \mathbb{R}, s \in S} \left| \hat{F}_t(g,s) - F(g,s) \right| \overset{\text{a.s.}}{\to} 0.$$

[3] Formally, the Glivenko-Cantelli lemma holds for univariate distribution, but its extension to multivariate distribution follows from the generalizations by Vapnik–Chervonenkis, see Shorack and Wellner (1986).

Adopting the proof of the Glivenko-Cantelli lemma, I now show that, in general, the empirical distribution of recalled experiences converges to the memory-weighted distribution. By the strong law of large numbers, the empirical joint distribution $\hat{F}_t(g,s)$ converges pointwise to $m_{(g_s,t_t)}(g,s) \cdot F(g,s)$, that is

$$\hat{F}_t(g,s) - m_{(g_s,t_t)}(g,s) \cdot F(g,s) \overset{\text{a.s.}}{\to} 0.$$

The convergence is also uniform. Denote $F_{m,t}(g,s) = m_{(g_s,t_t)}(g,s) \cdot F(g,s)$ and fix a grid of two-dimensional points $x_j = (g_j, s_j)$, $j = 1, ..., m$ with $x_j < x_{j+1}$ and such that $F_{m,t}(x_j) - F_{m,t}(x_{j-1}) = \frac{1}{m}$. For all $x \in \mathbb{R} \times S$, it exists a $k \in \{1, ..., m\}$ such that $x \in [x_{k-1}, x_k]$. It must then hold that

$$\begin{aligned}
\hat{F}_t(x) - F_{m,t}(x) &\leq \hat{F}_t(x_k) - F_{m,t}(x) \leq \hat{F}_t(x_k) - F_{m,t}(x_{k-1}) = \hat{F}_t(x_k) - F_{m,t}(x_k) + \frac{1}{m} \\
\hat{F}_t(x) - F_{m,t}(x) &\geq \hat{F}_t(x_{k-1}) - F_{m,t}(x) \leq \hat{F}_t(x_{k-1}) - F_{m,t}(x_k) = \hat{F}_t(x_{k-1}) - F_{m,t}(x_k) - \frac{1}{m}.
\end{aligned}$$

Consequently,

$$\sup_{x \in \mathbb{R} \times S} \left| \hat{F}_t(x) - F_{m,t}(x) \right| \leq \max_{k \in \{1,...,m\}} \left| \hat{F}_t(x_k) - F_{m,t}(x_k) \right| + \frac{1}{m}.$$

However, $\max_{k \in \{1,...,m\}} \left| \hat{F}_t(x_k) - F_{m,t}(x_k) \right| \overset{\text{a.s.}}{\to} 0$ by the pointwise convergence that follows from the strong law of large numbers and we can guarantee that for any $\epsilon > 0$ and m such that $1/m < \epsilon$, we find a T such that for all $t \geq$ T we have $\max_{k \in \{1,...,m\}} \left| \hat{F}_t(x_k) - F_{m,t}(x_k) \right| \leq \epsilon - \frac{1}{m}$, which establishes almost sure convergence.

We have established that the empirical joint (cumulative) distribution converges uniformly to the true joint distribution. As a next step, I show that also the empirical density converges. Since the distribution of signals is known, I focus on the marginal density of endowment growth, but the argument extends to the joint density. Define a partition of the real line d_k such that $d_{k+1} - d_k = h$. The histogram of growth rates is then

$$\hat{f}_t(g) = \sum_{s \in S} \frac{\hat{F}_t(d_{k+1}, s) - \hat{F}_t(d_k, s)}{h},$$

for $g \in [d_{k+1}, d_k]$. Note that the marginal memory-weighted distribution of endowment growth, $f_{m,g}(g)$, is (Lipschitz-)continuous and finite by assumption. If we let $h \to 0$, the continuity of the marginal distribution and the mean-value theorem ensure that $\mathbb{E}\left(\hat{f}_t(g)\right) \to f_{m,g}(g)$ as $|H_t^R| \to \infty$. Thus, the empirical histogram of growth rates is an unbiased estimator of the memory-weighted density. Moreover, note that the histogram of recalled experiences becomes deterministic for $|H_t^R| \to \infty$, since $\mathrm{Var}\left(\hat{f}_t(g)\right) = \frac{\Pr(d_k \leq g \leq d_{k+1})\,(1-\Pr[d_k \leq g \leq d_{k+1}])}{|H_t^R|\,h^2}$. These properties of the empirical histogram of recalled growth rates imply that

$$\hat{f}_t(g) \xrightarrow{p} f_{m,g}(g).$$

The agent's recalled growth rates converges in probability to the memory-weighted version of the true probability density, since the density exists by construction of $\mathcal{D}$. Moreover, if we restrict the set $\mathcal{D}$ to the class of uniformly integrable random variables, as considered in the applications of this paper, then the empirical density is uniformly integrable.

Step 2: As a next step, I show that the agent's posterior beliefs concentrate on those elements of the prior that maximize the likelihood. Intuitively, the Bayesian posterior is proportional to the prior times likelihood, but the prior is "washed out" for $t \to \infty$. The agent's beliefs thus concentrate on distributions that maximize the likelihood (see the Bernstein-von-Mises theorem).

For a many recalled observations $|H_t^R| \to \infty$, the log-likelihood of recalled experiences under a given distribution $q \in Q$ is

$$\log\left(\prod_{\tau \in r_t} q_{s_\tau}(g_\tau)\right) = \sum_{s \in S} \psi(s) \int_{-\infty}^{\infty} |H_t^R|\, \hat{f}_t(g)\, \log q_s\,(g)\; dg$$

$$= |H_t^R| \sum_{s \in S} \psi(s) \int_{-\infty}^{\infty} f_{m,g}(g) \log q_s(g) \; dg$$
$$= |H_t^R| \sum_{s \in S} \psi(s) \int_{-\infty}^{\infty} m_{(g_s,t_t)}(g,s) \, q_s^*(g) \log q_s(g) \; dg$$
$$= |H_t^R| \, L(q, H_t^R),$$

where I used the convergence of the empirical density $\hat{f}_t(g)$ to the memory-weighted true density from Step 1, and denote the log-likelihood of model q given the recalled history H_t^R by $L(q, H_t^R)$.

From Equation 1, the posterior odds ratio of two models $q, q' \in Q$ is given by

$$\frac{\prod_{\tau \in r_t} q_{s_\tau}(g_\tau) \, b_0(q)}{\prod_{\tau \in r_t} q'_{s_\tau}(g_\tau) \, b_0(q')} = \rho \frac{\exp\left[\log \prod_{\tau \in r_t} q_{s_\tau}(g_\tau)\right]}{\exp\left[\log \prod_{\tau \in r_t} q'_{s_\tau}(g_\tau)\right]}$$
$$= \rho \exp\left[|H_t^R| \left(L(q, H_t^R) - L(q', H_t^R)\right)\right].$$

The prior odds ratio, $\rho = \frac{b_0(q)}{b_0(q')}$, is fixed. However, for $L(q, H_t^R) > L(q', H_t^R)$, the posterior odds ratio diverges to ∞ for $|H_t^R| \, to \infty$, since the probability of model q' being correct goes to zero. Similarly, if $L(q, H_t^R) < L(q', H_t^R)$, the posterior odds ratio converges to 0 because the probability of q being correct goes to 0. Therefore, the agent's posterior beliefs concentrate on the maximizers of the memory-weighted likelihood as given in Equation 2. If the prior support contains the memory-weighted density, the agent's beliefs will then concentrate on the memory-weighted density.

Step 3: Last, I show that indeed $|H_t^R| \to \infty$ for $t \to \infty$, which follows from claim 1 in Fudenberg et al. (2023). The proof is replicated here for completeness. Formally, I want to show that for all $\hat{v} \in \mathbb{N}$ and $k \in \mathbb{N}$, $\mathbb{P}\left(|H_t^R| \leq k, \forall v \geq \hat{v} \;\right) = 0$. For $j \in \mathbb{N}$, it is

$$\mathbb{P}\left(|H_t^R| \leq k, \forall v \in \{\hat{v}, \hat{v} + j\}\right)$$
$$= \prod_{\tau = \hat{v}}^{\hat{v}+j} \sum_{h \in H_{\tau-1}} \mathbb{P}[h] \left(1 - \mathbb{P}\left[|H_\tau^R| > k | h\right]\right)$$

$$\leq \prod_{\tau=\hat{v}}^{\hat{v}+j} \left(\mathbb{P}\left[\{h \in H_{\tau-1} : |h| \leq k\}\right] + \sum_{h \in H_{\tau-1}:|h|>k} \mathbb{P}\left[h\right] \left(1 - \mathbb{P}\left[|H_\tau^R| > k|h\right]\right) \right).$$

Now, note that $\forall h \in H_{\tau-1} : |h| > k$ and for all objective histories H, there exists a constant $l \leq 1$ such that $\mathbb{P}\left[|H_\tau^R| > k|h\right] \leq l$, such that

$$\begin{aligned}
&\prod_{\tau=\hat{v}}^{\hat{v}+j} \left(\mathbb{P}\left[\{h \in H_{\tau-1} : |h| \leq k\}\right] + \sum_{h \in H_{\tau-1}:|h|>k} \mathbb{P}\left[h\right] \left(1 - \mathbb{P}\left[|H_\tau^R| > k|h\right]\right) \right) \\
\leq &\prod_{\tau=\hat{v}}^{\hat{v}+j} \left(\mathbb{P}\left[\{h \in H_{\tau-1} : |h| \leq k\}\right] + \sum_{h \in H_{\tau-1}:|h|>k} \mathbb{P}\left[h\right] (1 - l) \right) \\
= &\prod_{\tau=\hat{v}}^{\hat{v}+j} \left(\mathbb{P}\left[\{h \in H_{\tau-1} : |h| \leq k\}\right] + (1-l)\left(1 - \mathbb{P}\left[\{h \in H_{\tau-1} : |h| \leq k\}\right]\right) \right) \\
= &\prod_{\tau=\hat{v}}^{\hat{v}+j} 1 - l + l\,\mathbb{P}\left[\{h \in H_{\tau-1} : |h| \leq k\}\right]).
\end{aligned}$$

For a sufficiently large $\hat{v}$ and for all $v > \hat{v}$, the probability of histories having less than k observations is smaller than 1, or $\mathbb{P}\left[\{h \in H_{\tau-1} : |h| \leq k\}\right] < 1$, implying that $-l + l\,\mathbb{P}\left[\{h \in H_{\tau-1} : |h| \leq k\}\right] < 0$. Since $1 - x \leq e^{-x}$ for all $x \in \mathbb{R}$, it is

$$\begin{aligned}
\prod_{\tau=\hat{v}}^{\hat{v}+j} 1 - l + l\,\mathbb{P}\left[\{h \in H_{\tau-1} : |h| \leq k\}\right]) &\leq \prod_{\tau=\hat{v}}^{\hat{v}+j} \exp\left(-l + l\,\mathbb{P}\left[\{h \in H_{\tau-1} : |h| \leq k\}\right] \right) \\
&= \exp \sum_{\tau=\hat{v}}^{\hat{v}+j} \left(-l + l\,\mathbb{P}\left[\{h \in H_{\tau-1} : |h| \leq k\}\right] \right),
\end{aligned}$$

such that

$$\lim_{j\to\infty} \mathbb{P}\left(|H_t^R| \leq k, \forall v \in \{\hat{v}, \hat{v}+j\}\right) \leq \lim_{j\to\infty} \exp \sum_{\tau=\hat{v}}^{\hat{v}+j} \left(-l + l\,\mathbb{P}\left[\{h \in H_{\tau-1} : |h| \leq k\}\right] \right) = 0,$$

which shows the claim that $H_t^R \to \infty$ almost surely for $t \to \infty$.

OA.2 Extensions

OA.2.1 Similarity-weighted memory and log-normal endowment growth

In this Appendix, I briefly discuss the implications of similarity-weighted memory if the endowment growth process is log-normal. I first highlight the implications of similarity-weighted memory for the agent's subjective beliefs, to then discuss the implications for asset prices.

Consider the framework in Section 3 with $\mu_1 = \mu_2$ and $\sigma_1 = \sigma_2$, such that

$$g_t = \mu + \sigma \epsilon_t, \quad \epsilon_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1).$$

The agent learns about both parameters of endowment growth, the mean and the volatility, from her recalled observations. The agent's memory is distorted by the similiarty-weighted memory function in Equation 11.

The agent's long-term beliefs are as in Proposition 2, with

$$\begin{aligned} \hat{\mu}_t &= (1 - \alpha)\, \mu + \alpha\, g_t, \text{ and} \\ \hat{\sigma}_t^2 &= (1 - \alpha)\, \sigma^2, \end{aligned}$$

where $\alpha = \frac{\sigma^2}{\kappa + \sigma^2}$. The dynamics of the agent's posterior mean are as in Section 3, but the agent's posterior variance is always smaller than the fudamental variance because $\alpha \in (0, 1)$. Intuitively, under similarity-weighted memory, the agent is more likely to recall growth rates that are close to g_t, while the agent does not recall growth rates that are further in the tail of the distribution. In line with Proposition 1, the covariance between the distance of endowment growth from the subjective location parameter $\hat{\mu}_t$ and the probability of recall is negative under similarity-weighted memory, such that $\hat{\sigma}_t^2 < \sigma^2$. Moreover, the agent's posterior variance is not time-varying if endowment growth is log-normally distributed.

The cumulant-generating function of endowment growth under the agent's time-t belief is then given by

$$\mathcal{K}_t^{SL}(k) = \log \tilde{\mathbb{E}}_t \left(e^{k\, g_{t+1}}\right) = k\, \hat{\mu}_t + \frac{1}{2}\, k^2\, \hat{\sigma}_t^2 = \underbrace{\alpha\, k\, g_t}_{\text{Time-varying}} + \underbrace{(1-\alpha)\, \left[k\, \mu + \frac{1}{2}\, k^2\, \sigma^2\right]}_{\text{Fixed}}.$$

The subjective cumulant-generating function under similarity-weighted memory consists of two components: This period's endowment growth g_t—which receives weight α, and the true cumulant-generating function of endowment growth, with weight $(1-\alpha)$.

As a next step, I simulate the model $10,000$ times for 304 quarters and report average moments in Table 4. The parameters of the endowment growth process are as in Nagel and Xu (2022) with a quarterly mean endowment growth of $\mu = 0.44\%$ and a quarterly volatility of $\sigma = 1.31\%$. All other parameters are as in Table 1.

Table OA.1: Asset prices under similarity-weighted memory and log-normal endowment growth

Symbol	Mean	Std.	Corr. with g_t
Endowment growth and subjective beliefs			
g_t	1.758	2.618	1.000
$\hat{\mu}_t$	1.760	0.044	1.000
$\hat{\sigma}_t$	2.598	0.000	0.000
Subjective asset prices			
er_t	3.980	0.029	1.000
r_t^f	1.956	0.059	1.000
rp_t	2.205	< 0.001	-0.004
Objective asset prices			
rp_t	1.724	13.414	-1.000

Table OA.1 reports the model moments obtained from $10,000$ simulations of the model for 304 quarters. I annualize the quantities as follows: Means are multiplied by four and the standard deviations are multiplied by two. For the risk-free rate, I multiply the quarterly mean and the standard deviation by four.

The simulation results in Table OA.1 highlight that the agent's posterior mean is an unbiased estimate of the true mean endowment growth, while the agent's posterior variance is lower than the fundamental variance and constant over time. The asset pricing implications are as discussed in the main text, but the subjective risk premium is almost constant due to the constant posterior variance.

OA.2.2 State-dependent similarity-weighted memory

In this Appendix, I consider the effect of similarity-weighted memory if similarity also depends on the observable state. I focus on the implications of similarity-weighted memory on the agent's beliefs.

Assume that endowment growth is as in Section 3, but the memory function differs from that in Section 3 and is given by

$$m^S_{(g_t,s_t)}\left(g_\tau, s_\tau\right) = \exp\left[-\frac{\left(g_\tau - g_t\right)^2}{\left(2 - \left|s_t - s_\tau\right|\right)\kappa}\right]. \tag{.1}$$

If $s_t = s_\tau$, the memory function is as in Equation 11. On the contrary, if $s_t \neq s_\tau$, the memory function is $\exp\left[-\frac{(g_\tau - g_t)^2}{\kappa}\right]$. Under the memory function in Equation .1, the agent is more likely to remember past growth rates that are closer to today's endowment growth rate, and to remember growth rates that occurred in the same state as today's state.

Since the agent's recalled experiences consist of (g_τ, s_τ), we can proceed case-wise and analyze the agent's posterior beliefs conditional on today's state. It is

$$\hat{\mu}_{1,t} = \mu_1 + \begin{cases} \frac{\sigma_1^2}{\sigma_1^2+\kappa}\left(g_t - \mu_1\right) & \text{if } s_t = 1 \\ \frac{2\,\sigma_1^2}{2\,\sigma_1^2+\kappa}\left(g_t - \mu_1\right) & \text{if } s_t = 2 \end{cases}, \text{ and}$$

$$\hat{\mu}_{2,t} = \mu_2 + \begin{cases} \frac{2\,\sigma_2^2}{2\,\sigma_2^2+\kappa}\left(g_t - \mu_2\right) & \text{if } s_t = 1 \\ \frac{\sigma_2^2}{\sigma_2^2+\kappa}\left(g_t - \mu_2\right) & \text{if } s_t = 2 \end{cases}.$$

Note that $\frac{2\,\sigma_s^2}{2\,\sigma_s^2+\kappa} > \frac{\sigma_2^2}{\sigma_2^2+\kappa}$, such that the effect of similarity is stronger for the posterior mean about the state that is not currently observed. Thus, although the framework is i.i.d., we expect to observe predictable changes in the agent's posterior mean belief conditional on the current state even holding g_t fixed. In addition, the conditional posterior variance of the agent is given by

$$\hat{\sigma}_{1,t}^2 = \sigma_1^2 \cdot \begin{cases} \frac{\kappa}{\kappa+\sigma_1^2} & \text{if } s_t = 1 \\ \frac{2\,\kappa}{\kappa+2\,\sigma_1^2} & \text{if } s_t = 2 \end{cases},$$

$$\hat{\sigma}_{2,t}^2 = \sigma_2^2 \cdot \begin{cases} \frac{2\,\kappa}{\kappa+2\,\sigma_2^2} & \text{if } s_t = 1 \\ \frac{\kappa}{\kappa+\sigma_2^2} & \text{if } s_t = 2 \end{cases}.$$

Again, since similarity-based selectivity is stronger for the state that is currently not occurring, the agent's posterior variance of endowment growth in the "other" state is smaller than her posterior variance of endowment growth in the current state.