Investor Beliefs and Asset Prices Under Selective Memory

Maximilian Voigt*

November 3, 2025

Click here for the latest version.

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. The findings 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 Nick Barberis, Francesca Bastianello (discussant), Giovanni Burro, Adrian Buss, Paula Cocoma, Chukwuma Dim, Paul Fontanier, Katrin Gödker, Andreas Grunewald, Leyla Han (discussant), Jonas Happel, Björn Höppner, Jonas Jensen, Petri Jylhä (discussant), Yigitcan Karabulut, Paul Karehnke, Heiko Karle, Emanuel Mönch, Stefan Nagel (discussant), Cameron Peng, Francesco Sannino, Paul Schmidt-Engelbertz, Frederik Schwerter, Kaushik Vasudevan, and Kaspar Zimmermann; seminar participants at Aalto University, W. P. Carey School of Business at Arizona State University, ESCP Paris, emLyon, Frankfurt School, HEC Montreal, HEC Paris, Heidelberg University, Toulouse School of Economics, University of Vienna, Gies College of Business at University of Illinois Urbana-Champaign, and Yale SOM; and conference participants at the Cognitive Foundations of Finance Conference 2025, the Frankfurt-Bonn-Mannheim PhD conference 2023, the Finance Theory Group Summer School 2023, the Helsinki Finance Summit 2025, the NFA 2024, the SITE Asset Pricing Workshop 2025, and the WFA 2024. 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. Affiliation: HEC Montréal, maximilian.voigt@hec.ca.

1 Introduction

Understanding belief formation is central to understanding asset prices. In any equilibrium, the price of an asset reflects beliefs about future dividends and prices. The rational expectations approach assumes that investors understand the temporal fluctuation of dividends and prices, but survey evidence differs systematically from rational expectations (Adam and Nagel, 2023). Evidence from economics (Zimmermann, 2020; Bordalo et al., 2023; Enke et al., 2024; Charles, 2025; Charles and Sui, 2025; Gödker et al., 2025; Jiang et al., 2025) and psychology (Tversky and Kahneman, 1973; Schacter, 2008; Kahana, 2012) highlights the role of selective memory for belief formation and decision making. Although recent models incorporate memory to account for individual-level belief and choice puzzles (Mullainathan, 2002; Bordalo et al., 2020b, 2023; Wachter and Kahana, 2024), the effect of selective memory on general equilibrium asset prices is largely unexplored (Malmendier and Wachter, 2024).

I show that selective memory reconciles important facts about belief formation, survey data, and asset prices and yields novel predictions. 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 generates a persistent wedge between 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 belief formation (Kahana, 2012; Bordalo et al., 2020b, 2023; Enke et al., 2024; Charles, 2025; Chen and Huang, 2025; Jiang

[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; Enke et al., 2024; Conlon, 2025; Gödker et al., 2025). My focus is not on the neural encoding process, but on the economic consequences of selective recall for beliefs and prices.

et al., 2025), and Kahana et al. (2024) list similarity as a law of human memory. Subjective beliefs under similarity-weighted memory are consistent with empirical patterns: the agent expects high growth in good times (Nagel and Xu, 2022; Bordalo et al., 2024), expectations overreact to news (Bordalo et al., 2020a), and the subjective volatility of fundamentals is lower in good than in bad times (Lochstoer and Muir, 2022).

Incorporating the agent's beliefs into the asset pricing model yields empirically observed differences between 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), while the subjective risk premium is not predictable using aggregate valuation ratios but positively related to subjective volatility (Nagel and Xu, 2023). In contrast, objectively realized returns and risk premia are countercyclical (Shiller, 1981; Mehra and Prescott, 1985), predictable by aggregate valuation ratios (Campbell and Shiller, 1988), and do not vary with macroeconomic risk (Lettau and Ludvigson, 2010).

Beyond matching these patterns, the model yields new predictions. First, similarity-weighted memory allows temporally distant but contextually related past observations to shape beliefs and thus asset prices—consistent with recent empirical findings (Charles and Sui, 2025; Chen et al., 2025). Second, the model predicts systematic non-linear variation in higher moments of subjective beliefs, such as a convexity of perceived risk and of subjective risk premia, for which I present suggestive evidence.

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 theoretical analysis, I focus on learning from an infinite sample to identify systematic effects 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 (Fudenberg et al., 2024). 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 (Bordalo et al., 2020a); 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 above than below the fundamental mean. The agent's posterior mean is thus more likely above than below the fundamental mean after an upward revision, implying a predictably negative forecast error.

Variation in subjective volatility arises because recalling similar past experiences interferes with the recall of less similar ones. The economy has two states. With two states, the subjective volatility of log endowment growth depends on the perceived difference between the mean growth rates in each state, which varies over time. For intuition, suppose that endowment growth is always high in normal times, such that the agent can only recall high growth from normal times. Recessions can be either mild or severe. During good times, the agent recalls mild recessions but is oblivious to severe recessions. She thus perceives only a minor difference between the states and her subjective volatility is low during good times. During bad times, in contrast, she recalls severe recessions vividly, perceives a large difference between the states and her subjective volatility is high.

Extending the intuition to a more continuous setting yields a novel prediction of the model: the agent's subjective volatility is countercyclical, yet convex in the current context. Economically, the agent perceives the economy as relatively safe during average times, but as very risky during extremely good as well as extremely bad times. For intuition, consider an extension of the previous example in which the recession state is highly volatile, so that endowment growth can occasionally be high even in recessions. While the posterior mean of the normal state remains fixed, the posterior mean of the recession state can be below, equal to, or above the posterior mean of the normal state, depending on the current context. The perceived difference between the states is thus convex in the current context, leading to a convex subjective volatility. While this mechanism is specific to the two-state setting, the same economic effect arises robustly when the agent learns from a limited number of past observations: In extreme times, only a few past observations resemble the current context. The agent thus needs to estimate the parameters of the endowment growth process from a small recalled sample, and is highly uncertain about the estimates. In contrast, during average times, many similar past observations are recalled, yielding more precise estimates. Subjective parameter uncertainty is thus convex in the current context and priced similar to volatility under Epstein and Zin (1989)-preferences (Collin-Dufresne et al., 2016). I find empirical evidence suggesting that subjective riskiness is both countercyclical and higher in extreme than in average times.

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 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 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 not predictable by aggregate valuation ratios and positively related to the agent's perception of risk, consistent with evidence in Nagel and Xu (2023). I also find evidence suggesting that the subjective risk premium is convex in endowment growth, as predicted by the model.

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., 2024). Moreover, the agent updates beliefs as if the recalled experiences were all that ever occurred and perceives her new beliefs to be persistent, which leads to volatile realized 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 limited number of past observations, and can even become negative during times of extremely high sentiment (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.

Extensions. I demonstrate the applicability of the framework with two extensions. First, I introduce a persistent memory context (Howard and Kahana, 1999, 2002), which generates a recency effect—the agent recalls more recent experiences with higher probability than more distant ones—in addition to similarity and interference (Kahana, 2012; Barberis, 2018). The persistent context leads to persistence in beliefs and allows the model to match the empirical persistence of the price-dividend ratio (Campbell and Cochrane, 1999). Second, I incorporate a peak-end memory distortion, capturing the higher memorability of extreme and contextually similar experiences (Kahneman, 2000), consistent with the literature on experience effects (Malmendier and Nagel, 2011, 2016). The peak-end distortion increases the tendency to recall extreme outcomes, leading to a high perceived risk and thus a higher subjective risk premium—consistent with survey evidence (Adam and Nagel, 2023).

Related literature. This paper develops a tractable general equilibrium framework that links selective memory to asset prices, contributing to a growing literature that examines the effect of memory on belief formation and decision-making (Zimmermann, 2020; Charles, 2021; Goetzmann et al., 2022; Graeber et al., 2022; Burro et al., 2023; Enke et al., 2024; Charles, 2025; Chen and Huang, 2025; Gödker et al., 2025; Jiang et al., 2025). Consistent with evidence from psychology (Tulving and Schacter, 1990; Schacter, 2008; Kahana, 2012), the agent's recall of past observations is systematically biased toward experiences that resemble the current context. Using a representative survey of individual investors, Jiang et al. (2025) identify similarity-weighted memory as a key mechanism in investors' belief formation, and Charles and Sui (2025) show that beliefs constructed under similarity-weighted memory can explain an array of return expectations. The theoretical literature primarily analyzes the effect of memory 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., 2023; Fudenberg et al., 2024).

My paper also contributes to the literature on subjective beliefs in asset pricing (for an overview, see Adam and Nagel, 2023). Empirically, investor expectations tend to be 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). Existing models explain this pattern through 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, 2025), or learning from personal experience in overlapping generations (Ehling et al., 2018; Malmendier et al., 2020). I show that similarity-weighted memory microfounds procyclical cash-flow expectations, in line with evidence (Nagel and Xu, 2022; Bordalo et al., 2024). Moreover, similarity-weighted memory also leads to return extrapolation, and time-varying subjective volatility, unifying several empirical regularities with a single microfoundation.

Studies that incorporate memory into asset pricing focus mainly on fading memory to account for the evidence that lifetime experiences shape macroeconomic expectations (Malmendier and Nagel, 2011, 2016; Happel et al., 2023; Malmendier and Wachter, 2024), 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 dissipates over time. In contrast and consistent with evidence in Charles and Sui (2025) and Chen et al. (2025), past experiences have a long-lasting effect on the agent's beliefs and are never truly forgotten under selective memory.[2] In addition to the asset pricing results in Nagel and Xu (2022), I also obtain a convex subjective risk premium, for which I find suggestive evidence.

Selective memory is a central motivation for diagnostic expectations (Bordalo et al., 2022, 2023), 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 while retaining Bayesian learning, 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 (2024) provide a psychologically motivated theory of associative recall. Their focus is on decision-making in a partial equilibrium, which is distinct from my general equilibrium focus.

My framework also draws on the literature on misspecified learning in statistics and economics (Berk, 1966; Esponda and Pouzo, 2016; Molavi, 2019; Heidhues et al., 2021; Molavi et al., 2024). Fudenberg et al. (2024) propose the concept of posterior beliefs as maximizers of the memory-weighted likelihood that is central to my characterization of subjective long-term beliefs.

[2]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, but experimental findings challenge the notion that experiences are truly forgotten. Instead, if the context of the original experience is reinstated, seemingly forgotten memories are typically recalled (Kahana, 2012).

2 Beliefs and asset prices under a general selective memory distortion

This section characterizes the agent's long-term beliefs under selective memory and embeds them into a standard asset pricing framework. Selective memory is my only departure from rational expectations, and all asset pricing implications arise from the agent's subjective long-run beliefs. The formulation here is general and accommodates various forms of selective memory—for instance, agents may be overly likely to recall past extreme experiences, or the recall of past experiences may depend on the current context.

Section 2.1 provides a non-technical outline of the selective-memory mechanism and its implications for belief formation. Section 2.2 then introduces the learning environment, and Section 2.3 characterizes the agent's subjective long-run beliefs. I focus on discrete distributions as in Fudenberg et al. (2024), and provide an extension to continuous distributions in Appendix G. Proposition 1 is new and characterizes subjective long-term beliefs under normally distributed fundamentals, as is relevant for applications in finance. Section 2.4 embeds the agent's subjective long-term beliefs into a standard consumption-based asset pricing framework.

2.1 From selective memory to subjective beliefs

I consider a standard Bayesian learning problem in which the agent holds prior beliefs, observes realizations from a fixed probability distribution, and updates beliefs using Bayes' rule. I focus on the case where the agent has already observed infinitely many realizations. Under some regularity conditions, the beliefs of a rational Bayesian agent converge to the true probability distribution. By the law of large numbers, the empirical histogram of realizations converges to the true probability distribution. The agent's long-run posterior beliefs concentrate on the distribution that best fits the histogram of observed realizations, and thus converge to the true probability distribution.

The agent considered here departs from full rationality because her recall of past observations is distorted by selective memory. A large literature in psychology and neuroscience shows that human memory is both stochastic—the recall of a specific past observation is probabilistic (Shadlen and Shohamy, 2016)—and selective—the recall of certain past experiences is more likely than the recall of others (Schacter, 2008; Kahana, 2012; Wachter and Kahana, 2024). I model stochastic and selective recall through a memory function that assigns each past observation a recall probability, potentially dependent on the context.

The agent then forms beliefs from recalled rather than from all past realizations. Within each bin of the empirical histogram, the share of recalled observations converges to the recall probability for that bin. The histogram of recalled realizations thus converges to a memory-weighted version of the empirical histogram. The agent's long-term beliefs are then given by the distribution that best fits this memory-weighted histogram, which is a memory-weighted version of the true distribution. Selective memory therefore distorts the agent's beliefs even when the agent has observed infinitely many realizations, and the distortion depends on the form of the memory function—–agents who overweight favorable outcomes, for example, will hold systematically optimistic beliefs. The following sections formalize the argument.

2.2 Learning framework

Economy. I study a representative agent endowment economy in discrete time. In every period t, the agent observes the state of the economy 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 $\Xi \in \Delta(S)$, where $\Xi(s)$ denotes the probability of state 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 $q^*_s(g) = Pr\left(g_t = g|s_t = s\right)$.[3] I assume

[3]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. (2024) here. Behaviorally, the restriction implies that the agent observes and recalls a discrete approximation of the continuously-distributed endowment growth.

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,[4] but must learn the (state-dependent) 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 $g_t = g$ conditional on $s_t = s$, 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. First, the agent is correctly specified $q^* \in Q$ (Esponda and Pouzo, 2016; Fudenberg et al., 2024), 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 $q_s(g) > 0$, which implies that none of the $q \in Q$ can be ruled out by the agent in finite time.

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 observes and recalls the current experience (g_t, s_t), but her memory of any past experience is distorted by the *memory function* $m_{(g_t,s_t)} : G \times S \mapsto [0, 1]$. 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 the current experience (g_t, s_t). The current experience shapes the agent's context, which can affect the probability of recalling past experiences.[5] 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

The results extend to continuous distributions (Appendix G).

[4]It is not essential that the agent knows Ξ. Section 3 shows results without this assumption.

[5]The model nests the rational-expectations benchmark with $m_{(g_t,s_t)}(g_\tau, s_\tau) = 1$ for all (g_τ, s_τ).

for each state s.

Beliefs. The agent forms Bayesian beliefs as if her recalled history H_t^R is all that occurred (naïveté).[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 subset of probability distributions in the agent's prior support $\mathcal{Q}$.

2.3 Subjective long-term beliefs

I now characterize the agent's subjective long-term beliefs. Define the *memory-weighted likelihood maximizer* (Fudenberg et al., 2024) conditional on the current experience (g_t, s_t),

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

The memory-weighted likelihood maximizer is the distribution in the agent's prior support that maximizes the likelihood of the recalled history H_t^R. Fudenberg et al. (2024) 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 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 (Berk, 1966;

[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.

Kleijn and Van Der Vaart, 2012), which implies that the agent's beliefs concentrate on the memory-weighted likelihood maximizer.

I next illustrate the effect of selective memory on the agent's subjective beliefs when log endowment growth is normally distributed conditional on the state, as is relevant for asset pricing. 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.[7] 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 and recalls the state, 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 $\hat{\theta}_{s,t} := (\hat{\mu}_{s,t}, \hat{\sigma}_{s,t}^2)$, and*

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

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

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

Equations (3)–(4) decompose the agent's posterior beliefs under selective memory for a normally distributed log endowment growth into fundamental components and memory-induced distortions.

The posterior mean in state s depends on two terms: the fundamental mean μ_s and an

[7] 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 $\bar{\sigma}^2 < \infty$.

adjustment arising from selective memory. The adjustment is the product of two factors: First, the forgetfulness term, $\mathbb{E}\left[\frac{t_s}{|H^R_{s,t}|}\right]$, measures the expected ratio of realized to recalled observations in state s. Second, the selectivity term, $\text{Cov}\left[g, \mathbb{1}_{\{g\in H^R_{s,t}\}}\right]$, captures the directional distortion in recall: it is positive whenever the agent is more likely to recall high than low realizations of g, and negative vice-versa.[8] The agent is thus too optimistic ($\hat{\mu}_{s,t} > \mu_s$) whenever the selectivity term is positive, meaning that the agent recalls a positively biased history of log endowment growth; and too pessimistic ($\hat{\mu}_{s,t} < \mu_s$) whenever the selectivity term is negative. In addition, the agent learns the fundamental mean whenever the selective memory mechanism is unbiased, such that the covariance-term is zero. The effect of selectivity is scaled by the degree of forgetfulness. If the agent recalls nearly all past realizations of g ($|H^R_{s,t}| \approx t_s$), the adjustment is small because the agent's recalled history is not strongly distorted by selective memory. If the agent instead recalls a few past realizations of g, it is important how the recalled observations are selected by the memory mechanism, and the selectivity term has a large effect on the posterior mean.

The posterior variance (Eq. (4)) is anchored at the true variance σ_s^2. The second term, $(\hat{\mu}s, t - \mu_s)^2$, is the usual term to correct for a potentially biased mean. The third term, $\mathbb{E}\left[\frac{t_s}{|H^R_{s,t}|}\right] \cdot \text{Cov}\left[(g - \hat{\mu}_{s,t})^2, \mathbb{1}_{\{g\in H^R_{s,t}\}}\right]$, reflects selectivity with respect to the tails: it is positive when the agent is more likely to recall extreme past observations, and negative when the agent is more likely to recall past observations that are close to the posterior mean. The agent perceives the economy as more risky than it fundamentally is whenever the selectivity term is positive, meaning that the agent is more likely to recall past observations in the tails of the distribution; while she perceives the economy as less risky than it fundamentally is whenever the selectivity term is negative.

In addition to generating overly optimistic or pessimistic beliefs, as well as distorted risk perceptions, selective memory can also lead to time-varying subjective beliefs if the selectivity terms vary, for example, with the current context (g_t, s_t). Proposition 1 is new and central

[8]Note that $\mathbb{1}_{\{g\in H^R_t\}}$ is a random variable with $\mathbb{E}\left(\mathbb{1}_{\{g\in H^R_t\}}\right) = m_{(g_t,s_t)}(g, s_\tau)$.

to the asset-pricing applications that follow. It holds without additional restrictions on the memory function and the applications in this paper introduce specific functional forms that determine the magnitude and sign of the selectivity and forgetfulness terms.

2.4 Asset pricing framework

I now 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).

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

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

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 the expected lifetime utility V_t under her subjective expectations $\tilde{\mathbb{E}}_t(\cdot)$ formed under the posterior 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, such that the economy is as in Martin (2013).

Consider an asset that pays a dividend stream $\{D_{t+k}\}_{k\geq 0}$ with $D_{t+k} = C_{t+k}^{\lambda}$ for some 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 $\hat{er}_t = \log\left(\tilde{\mathbb{E}}_t\ R_{t+1}\right)$. Similarly, the *log risk-free rate* is the log expected return on the riskless bond, and the *subjective risk premium* on the λ-asset is the difference between the log subjective expected return and the log risk-free rate.[9] Martin (2013) shows that

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

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

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

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

where $\mathcal{K}_t(k)$ is the cumulant-generating function under the agent's subjective long-term 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

$$\mathcal{M}_t(k) := \tilde{\mathbb{E}}_t\left(e^{k\,g_{t+1}}\right), \text{ and} \tag{10}$$

$$\mathcal{K}_t(k) := \log\left(\mathcal{M}_t(k)\right) = \log\tilde{\mathbb{E}}_t\left(e^{k\,g_{t+1}}\right), \text{ respectively.} \tag{11}$$

Both the moment-generating and cumulant-generating function provide expressions for the moments of log endowment growth under the agent's subjective belief b_t and we can recover the subjective moments of log endowment growth using $\mathcal{M}_t(k)$ or $\mathcal{K}_t(k)$. The first and second cumulant yield the agents' subjective mean $\hat{\mu}_t$ and variance $\hat{\sigma}_t^2$, whereby the n-th cumulant can be obtained by differentiating $\mathcal{K}_t(k)$ n times and evaluating the result at zero.

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 around zero, $\mathcal{K}_t(k) \approx k\,c_1 + \frac{1}{2}k^2\,c_2 = k\,\hat{\mu}_t + \frac{1}{2}k^2\,\hat{\sigma}_t^2$:

$$r_t^f = -\log(\beta) + \gamma\,\hat{\mu}_t - \frac{1}{2}\gamma^2\,\hat{\sigma}_t^2, \tag{12}$$

$$dp_t = -\log(\beta) - (\lambda-\gamma)\,\hat{\mu}_t - \frac{1}{2}(\lambda-\gamma)^2\hat{\sigma}_t^2, \tag{13}$$

[9] In equilibrium, the risk-free rate and asset prices in period t are determined under the agent's subjective measure and are therefore realized. Realized returns and risk premia may deviate from subjective expectations because they depend on next period's dividend and price.

$$\hat{er}_t = -\log(\beta) + \gamma\,\hat{\mu}_t + \lambda\,\gamma\,\hat{\sigma}_t^2 - \frac{1}{2}\,\gamma^2\,\hat{\sigma}^2, \tag{14}$$

$$\hat{rp}_t = \lambda\,\gamma\,\hat{\sigma}_t^2. \tag{15}$$

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.[10] 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 $\gamma\,\hat{\mu}_t$.

The agent's posterior variance—the agent's subjectively perceived risk in the economy—also 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, $\hat{rp}_t = \lambda\,\gamma\,\hat{\sigma}_t^2$, which is increasing in the posterior variance. The expected return increases in the posterior variance if the risk premium effect dominates the risk-free rate effect ($\lambda > \frac{1}{2}\,\gamma$).

Qualitatively, selective memory yields several empirical patterns in asset prices. For example, procyclical expected returns (Greenwood and Shleifer, 2014) arise when the agent's posterior mean varies with the business cycle while her perceived variance does not. Accord-

[10]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.

ing to Proposition 1, this occurs if the agent is more likely to recall high endowment growth rates in good times and low growth rates in bad times. Similarly, time-varying but acyclical risk premia (Nagel and Xu, 2023) emerge when the posterior variance—which is driven by the propensity to recall extreme past outcomes—fluctuates over time without systematic comovement with the state of the economy. The next section imposes additional structure on the memory distortion to examine the asset pricing implications in general equilibrium.

3 Similarity-weighted memory

In this section, I introduce similarity-weighted selective memory and impose additional structure on the economy. For tractability, I assume that the agent's memory context—which cues the recall of similar past experiences—is determined by today's log endowment growth (Section 3.1). Using Proposition 1, Section 3.2 characterizes the agent's long-term beliefs under similarity-weighted memory, while Section 3.3 derives the asset-pricing implications. Simulation results are in Section 3.4.

The memory and belief mechanism in this section works as follows: The economy is either in a normal state or in a recession. To gain intuition, suppose endowment growth is always high in normal times but can be high or low in recessions. The agent observes the current log endowment growth (the context) and retrieves past experiences with a similar endowment growth. As growth is always high in normal times, the agent can only recall normal times with a high growth, and her beliefs about normal times are essentially unaffected by similarity-weighted memory. In contrast, the recall of recession periods depends on the current context: when today's growth is high (low), the agent recalls more high (low) growth realizations from past recessions. Put differently, the agent becomes oblivious of bad recessions in good times but vividly recalls them in bad times.[11] Consequently, her posterior mean covaries positively with current growth because recessions are perceived to be

[11]This example extends the "thirsty traveler" example from Bordalo et al. (2020b): at airports, water prices are always high, so travelers recall only high prices; downtown, prices vary widely, so similarity-weighted memory systematically biases retrieved downtown prices.

worse if contemporaneous endowment growth is low than if it is high. The same mechanism generates variation in the subjectively perceived riskiness of the economy, which depends on how severe recessions are perceived to be relative to normal times. Variation in the agent's subjective beliefs due to similarity-weighted memory drives variation in asset prices in the representative agent economy analyzed here.

Similarity is a fundamental property of human memory (Kahana, 2012; Kahana et al., 2024; Wachter and Kahana, 2024) and influences both individual decision-making and market outcomes. Similarity-weighted memory has been shown to shape asset prices in experimental settings (Enke et al., 2024), while field studies document systematic effects on investor beliefs, trading behavior, and analyst forecasts (Chen and Huang, 2025; Jiang et al., 2025). Using a range of firm-level and macroeconomic variables, Charles and Sui (2025) measure the similarity between each quarter and the preceding sixty quarters and show that beliefs reconstructed from a similarity-weighted memory mechanism closely match survey-based return expectations, higher-moment beliefs, individual trading patterns, and recall behavior in earnings calls. At the aggregate level, Chen et al. (2025) demonstrate that average next-month returns over the twenty-five most similar historical episodes—as measured by textual similarity in news—predict stock-market returns out of sample. Together with the results in this section, these empirical findings suggest that similarity-weighted memory can explain salient facts about aggregate asset prices.

3.1 Structural assumptions

The assumptions in Section 2.2 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 focus on i.i.d. transitions by imposing $\pi_{11} = \pi_{21} =: \pi_1$ and $\pi_{12} = \pi_{22} = 1 - \pi_1 =: \pi_2$. Conditional on state $s_t = s$, log endowment growth is normally distributed

$$g_t = \mu_{s_t} + \sigma_{s_t} \epsilon_t, \quad \epsilon_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1), \tag{16}$$

with state-dependent mean μ_s and variance σ_s^2. 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.

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 as in Proposition 1. She also learns the transition probabilities π_{ij} from the recalled history H_t^R which is distorted by similarity-weighted selective memory

$$m_{(g_t,s_t)}^{\text{sim}}(g_\tau, s_\tau) = \exp\left[-\frac{(g_\tau - g_t)^2}{2\,\kappa}\right], \tag{17}$$

where $\kappa > 0$ captures the *scrutiny* with which the agent examines her memory database. A higher scrutiny implies that the agent recalls almost all past observations.

Discussion. The i.i.d. two-state specification of log endowment growth serves three purposes. First, Markov-switching models have long been used to capture aggregate endowment dynamics due to their flexibility and tractability (Mehra and Prescott, 1985; Rietz, 1988; Barro, 2006). Although Johannes et al. (2016) find that an i.i.d. model has a low posterior probability, I retain this structure to ensure that any time variation in beliefs and asset prices arises solely from similarity-weighted selective memory, and not from the properties of the economy. Second, the two-state setup generates variation in both the subjective mean and variance, whereas a one-state model only yields variation of the posterior mean (Proposition 2). The time-variation of the posterior variance qualitatively mirrors Bayesian parameter uncertainty that emerges when an agents with similarity-weighted memory forms beliefs from a limited number of past observations, but I can obtain the former in closed form using the i.i.d. two-state structure.[12] Third, the two-state structure links the model to

[12]A model with an i.i.d. log-normal endowment growth is in Appendix F.3.

recent work on learning within and across categories (Ba et al., 2024; Bastianello and Imas, 2025; Conlon and Kwon, 2025). I apply the same cognitive distortion—similarity-weighted selective memory—to both dimensions, and leave heterogeneity in distortions across versus within categories to future work.

Finally, Equation (17) models similarity as the numerical distance between today's log endowment growth, g_t, and past endowment growth, g_τ. Today's endowment growth thereby encodes the memory context (Kahana, 2012), which I generalize in Section 4. The functional form of Equation (17) is standard (Kahana, 2012; Charles and Sui, 2025; Jiang et al., 2025), provides tractable analytical results (see Proposition 2), and is not essential for the qualitative predictions of the model (see Proposition 1).

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, to then analyze overreaction of the agent's forecast and discuss the predictability of belief revisions.

Subjective beliefs. Proposition 2 characterizes the beliefs of an agent with similarity-weighted selective memory.

Proposition 2 (Subjective beliefs)**.** *Under similarity-weighted memory as in Equation (17), almost surely,*

$$\hat{\mu}_{s,t} = \frac{\kappa}{\kappa + \sigma_s^2}\,\mu_s + \frac{\sigma_s^2}{\kappa + \sigma_s^2}\,g_t = (1 - \alpha_s)\,\mu_s + \alpha_s\,g_t, \tag{18}$$

$$\hat{\sigma}_{s,t}^2 = (1 - \alpha_s)\,\sigma_s^2, \text{ and} \tag{19}$$

$$\hat{\pi}_{s,t} = \frac{\pi_s\,\mathcal{M}_{s,t}}{\pi_s\,\mathcal{M}_{s,t} + \pi_{s'}\,\mathcal{M}_{s',t}} \text{ with } \mathcal{M}_{s,t} = \sqrt{(1 - \alpha_s)}\,e^{-\frac{(\mu_s - g_t)^2}{2\,(\kappa + \sigma_s^2)}}, \tag{20}$$

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*, and* s' *generically denotes "other" state.*

Figure 1: Posterior beliefs under similarity-weighted memory

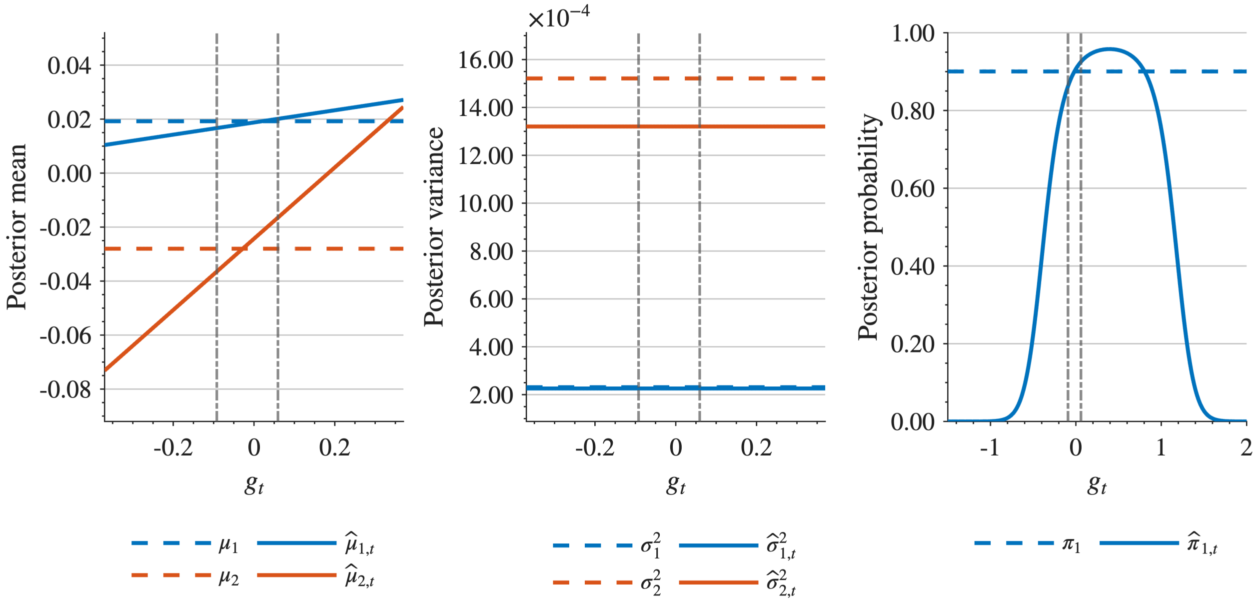

Figure 1 shows how the agent's posterior beliefs depend on the current context g_t. The solid lines in each panel show the posterior beliefs under similarity-weighted memory as given in Proposition 2, while the dashed lines show the values of the parameters that an agent with rational expectations learns. The parameters of the log endowment growth process are as in Table 1, and the vertical dashed lines show the 0.5% and 99.5% quantiles of the log endowment growth process.

Under similarity-weighted memory, the agent's subjective beliefs take simple, closed-form expressions, and Figure 1 shows how the agent's posterior beliefs depend on the current context g_t.

The posterior mean in state s is a convex combination of the fundamental mean μ_s and the agent's context g_t (Equation (18)). If today's endowment growth is high, the agent recalls more past experiences with a high endowment growth due to similarity. The agent will therefore be more (less) optimistic if this period's endowment growth is high (low), as shown in the left panel of Figure 1.

The sensitivity of the agent's state-dependent posterior mean to the current context g_t, given by α_s, depends on the state-dependent variance σ_s^2 and on memory scrutiny κ. If $\sigma_s^2 \to 0$, the distribution of log endowment growth in state s is concentrated at μ_s. All observations that the agent can possibly recall are then 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 agent's experiences in state s are very dispersed and similarity has a strong effect on the recalled experiences, leading to $\hat{\mu}_{s,t} \approx g_t$. I assume $\sigma_1^2 < \sigma_2^2$, such that the agent's posterior mean about state $s = 1$ is less sensitive to the context g_t than the posterior mean about state $s = 2$ (left panel of Figure 1).

A similar intuition applies to the scrutiny parameter κ. If scrutiny is high ($\kappa \to \infty$), similarity becomes irrelevant because the agent recalls all past observations, such that $\hat{\mu}_{s,t} \approx \mu_s$. If scrutiny is low ($\kappa \to 0$), similarity is very important for the recall probability and past experiences with log endowment growth that differ from the current context will not be recalled, such that $\hat{\mu}_{s,t} \approx g_t$.

The agent's state-dependent posterior variance (Equation (19)) is independent of the context g_t and smaller than σ_s^2 since $\alpha_s \in (0, 1)$ (see the middle panel of Figure 1). Similarity-weighted memory symmetrically overweights observations that are close to the current context g_t and underweights experiences that are very different. Therefore, the recalled state-dependent distribution of log endowment growth is more concentrated than the underlying state-dependent distribution, leading to a low perceived state-dependent variance.

The posterior probability in Equation (20) resembles the update of a Bayesian agent who learns about the probability of state s given signal g_t, with likelihood terms given by $\mathcal{M}_{s,t}$. If the current context is close to the mean of state s, the agent recalls more past experiences from state s than from state s' relative to the true probability π_s, because most of the observations that are close to g_t are from state s. In contrast, if the context is very extreme, the agent tends to recall more past experiences from the state with the higher variance, such that the perceived probability of the high-variance state is higher than the true probability. In the setting here, the posterior probability of the high-mean, low-variance state ($s = 1$) is a concave function of the context g_t, declining as g_t moves further from μ_1, as shown in the right panel of Figure 1.

I next discuss properties of the agent's unconditional time-t belief about log endowment growth, which determines asset prices. From now on, I assume that the agent knows the state-dependent variances, σ_1^2 and σ_2^2, and the transition probability π_1, but learns about the state-dependent means, μ_1 and μ_2. The results are qualitatively unchanged if the agent instead learns about variances and transition probabilities (see Figure B.1).

The unconditional distribution of log endowment growth is a mixture of the state-wise normal distributions, which is generally non-normal. I thus characterize the agent's perceived distribution using the *subjective* cumulant-generating function (CGF, Section 2.4):

$$\mathcal{K}_t(k) = \log\left[\mathcal{M}_t(k)\right] = \log\left[\pi_1\, e^{k\,\hat{\mu}_{1,t}+\frac{1}{2}\,k^2\,\sigma_1^2} + \pi_2\, e^{k\,\hat{\mu}_{2,t}+\frac{1}{2}\,k^2\,\sigma_2^2}\right]. \tag{21}$$

Under rational expectations, the corresponding CGF $\mathcal{K}^*(k)$ is obtained by replacing $\hat{\mu}_{s,t}$ with μ_s. Although higher-order cumulants significantly affect asset prices (Martin, 2013), I focus on the subjective mean and variance of log endowment growth here, which are typically measured in surveys and are the main drivers of the asset-pricing implications. The simulations in Section 3.4 incorporate all higher moments, and Figure B.1 shows how the first four subjective moments vary with g_t.

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

$$\hat{\mu}_t := \tilde{\mathbb{E}}_{\mathrm{t}}\left(g_{t+1}\right) = \pi_1\,\hat{\mu}_{1,t} + \pi_2\,\hat{\mu}_{2,t}. \tag{22}$$

Equation (22) shows that the 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). Similarity-weighted memory provides a microfoundation for procyclical cash-flow expectations (recall that all assets in the economy are levered claims on g_t), as found in survey data (Greenwood and Shleifer, 2014; Barberis, 2018; Adam and Nagel, 2023). In the present i.i.d. setting, the agents' beliefs are not persistent. Section 4 extends the model by incorporating a persistent context, which yields persistent expectations, as observed empirically (Bordalo et al., 2024). 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*

$$\mathbb{E}\left(\hat{\mu}_{1,t+1}|s_{t+1}=1\right)=\mu_1 \tag{23}$$

$$\mathbb{E}\left(\hat{\mu}_{1,t+1}|s_{t+1}=2\right)=\mu_1+\alpha_1\,\left(\mu_2-\mu_1\right), \tag{24}$$

and the average state-dependent posterior mean is

$$\mathbb{E}\left(\hat{\mu}_{1,t+1}\right)=\mu_1+\alpha_1\,\pi_2\,\left(\mu_2-\mu_1\right). \tag{25}$$

Equations (23)—(25) hold for $\hat{\mu}_{2,t+1}$ with the respective change of indices. Moreover, the average unconditional posterior mean of log endowment growth is

$$\mathbb{E}\left(\hat{\mu}_t\right)=\pi_1\,\mu_1+\pi_2\,\mu_2+\pi_1\,\pi_2\,\left(\mu_2-\mu_1\right)\,\left(\alpha_1-\alpha_2\right). \tag{26}$$

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. Thus, the posterior mean of state 1 is unbiased if $s_t=1$, but biased 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 (26) 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, overly optimistic about endowment growth.[13]

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

$$\hat{\sigma}_t^2 := \mathrm{Var}_t\left(g_{t+1}\right) = \pi_1\,\sigma_1^2 + \pi_2\,\sigma_2^2 + \pi_1\,\pi_2\,\left(\hat{\mu}_{1,t} - \hat{\mu}_{2,t}\right)^2. \tag{27}$$

Equation (27) shows that the perceived riskiness of the economy, $\hat{\sigma}_t^2$, depends on the state-specific variances and on the dispersion of state-wise posterior means. Intuitively, the economy appears riskier when the agent perceives recessions to be much worse than normal times ($|\hat{\mu}_{1,t} - \hat{\mu}_{2,t}|$ is large).

Proposition 4 (Subjective variance)**.** *The average posterior variance under similarity-weighted memory is*

$$\begin{aligned}\mathbb{E}\left[\hat{\sigma}_t^2\right] = &\left(\pi_1\,\sigma_1^2 + \pi_2\,\sigma_2^2\right)\left[1 + \pi_1\,\pi_2\,\left(\alpha_1 - \alpha_2\right)^2\right] \\ &+ \left(\mu_1 - \mu_2\right)^2\,\pi_1\,\pi_2\,\left[\pi_1\,\pi_2\left(\alpha_1 - \alpha_2\right)^2 + \left[1 - \left(\alpha_1\,\pi_2 + \alpha_2\,\pi_1\right)\right]^2\right],\end{aligned} \tag{28}$$

which is larger than the true variance σ^2 if

$$\frac{\left(\alpha_1 - \alpha_2\right)^2\,\left(\pi_1\,\sigma_1^2 + \pi_2\,\sigma_2^2\right)}{2\,\left(\pi_2\,\alpha_1 + \pi_1\,\alpha_2\right) - \left(\pi_2\,\alpha_1^2 + \pi_1\,\alpha_2^2\right)} \geq \left(\mu_1 - \mu_2\right)^2, \tag{29}$$

and bounded by

$$0 \leq \mathbb{E}\left[\hat{\sigma}_t^2\right] \leq 1.25\,\sigma^2.$$

Proposition 4 characterizes the average posterior variance, while Condition (29) identifies when similarity-weighted memory inflates the average perceived riskiness of the economy.

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

Because the left-hand side of Condition (29) is always positive, the condition tends to hold when the state means are close, i.e., $(\mu_1 - \mu_2) \approx 0$. Moreover, due to the difference in α_s, the left-hand side of Condition (29) increases in the difference of the state-specific variances, σ_1^2 and σ_2^2. Thus, on average, the perceived riskiness of the economy under similarity-weighted memory exceeds the fundamental riskiness if σ_2^2 is much larger than σ_1^2, while $\mu_1 \approx \mu_2$.

As shown in Proposition 2, the sensitivity of each posterior mean to the current context, $\alpha_s = \sigma_s^2/(\kappa + \sigma_s^2)$, increases with the state's variance. Hence, when $\mu_1 \approx \mu_2$ but $\sigma_2^2 >> \sigma_1^2$, the posterior mean of the high-variance state $s = 2$ reacts more strongly to g_t than that of the low-variance state $s = 1$, as in Figure 1. The asymmetric reaction of the state-wise posterior means to the current context raises their expected squared difference, $(\hat{\mu}_{1,t} - \hat{\mu}_{2,t})^2$, which amplifies the agent's perceived aggregate riskiness.

Define $g^* = \frac{(1-\alpha_1)\,\mu_1 - (1-\alpha_2)\,\mu_2}{\alpha_2 - \alpha_1}$ as the level of log endowment growth at which the two posterior means coincide. The difference between the posterior means $\hat{\mu}_{1,t} - \hat{\mu}_{2,t}$ decreases in log endowment growth for $g_t < g^*$, equals zero for $g_t = g^*$, and increases in g_t whenever $g_t > g^*$. The unconditional posterior variance is a convex function of the context g_t, with a minimum at g^* (see Figure B.1). Behaviorally, the convexity of the unconditional posterior variance means that the agent perceives the economy as relatively safe during "average" times—close to g^*—, but as very risky during extremely good as well as extremely bad times. To the best of my knowledge, the convexity of the perceived riskiness of the economy is a new prediction of similarity-weighted memory.[14]

Forecast overreaction. A property of rational forecasts is that they should not systematically overreact or underreact to new information. Coibion and Gorodnichenko (2015)

[14]The convexity result relies on $\sigma_1^2 < \sigma_2^2$, as found in the data. A similar pattern arises when endowment growth is log-normal and the agent learns from a limited sample of past observations. In that setting, a Bayesian learner faces parameter uncertainty that is priced under Epstein-Zin preferences (Collin-Dufresne et al., 2016). With similarity-weighted memory, the agent recalls relatively many observations near the center of the distribution—where data is dense—and relatively few observations in the tails. As a result, parameter uncertainty, and hence perceived risk, is also convex in g_t: the agent feels most certain in "average" times and most uncertain during extremely good and bad times. Because closed-form solutions cease to exist under limited-sample learning, the framework here focuses on the two-state setting that provides a tractable analogue.

propose the following regression to test for over- or underreaction of forecasts:

$$\underbrace{g_{t+h} - \tilde{\mathbb{E}}_{\mathrm{t}}\left(g_{t+h}\right)}_{\text{Forecast error}} = a_{CG} + \beta_{CG} \underbrace{\left[\tilde{\mathbb{E}}_{\mathrm{t}}\left(g_{t+h}\right) - \mathbb{E}_{t-1}\left(g_{t+h}\right)\right]}_{\text{Forecast revision}} + u_{t+h}. \tag{30}$$

Rationality implies $a_{CG} = 0$ and $\beta_{CG} = 0$, 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.[15] Otherwise, a rational agent would adjust her forecast. If $\beta_{CG} < 0$, beliefs overreact as the belief revision is too strong on average, while $\beta_{CG} > 0$ captures underreaction. Proposition 5 shows that I find overreaction in Coibion and Gorodnichenko (2015)-regressions for an agent with similarity-weighted memory.

Proposition 5 (Forecast overreaction)**.** *The long-term beliefs of an agent with similarity-weighted memory overreact as measured by Coibion and Gorodnichenko (2015)-regressions in Equation (30), since*

$$\beta_{CG} = -\frac{1}{2} < 0, \qquad \textit{and} \qquad a_{CG} = -\pi_1\, \pi_2 \left(\mu_2 - \mu_1\right) \left(\alpha_1 - \alpha_2\right). \tag{31}$$

The intuition for overreaction under similarity-weighted memory is as follows: An upward forecast revisions means that today's growth rate is higher than yesterday's, $g_t > g_{t-1}$. A high g_t makes it more likely that the agent's posterior mean is above the fundamental mean, $\Pr(g_t \geq \mu_1 | g_t > g_{t-1}) > 0.5$, due to similarity-weighted memory. Since the true process is i.i.d., the best objective forecast for next period's growth, however, is the true mean. Therefore, the average forecast error is predictably negative following an upward revision of the agent's forecast.

Belief predictability. As a final step, I argue that the agent's belief revisions are predictable. Asset prices are determined by the agent's subjective beliefs, such that predictable belief revisions lead to return predictability.

[15] Note that the Coibion and Gorodnichenko (2015)-regression is not well identified for an agent without memory distortions in our setting, as such an agent never revises the forecast.

An econometrician with access to the same realized, infinitely long history H_t as the agent will learn the true parameters of the data-generating process and can use them to forecast the agent's beliefs. The objective expectation of the subjective moment-generating function—that is, the agent's expected next-period belief—is

$$\begin{aligned}\tilde{\mathcal{M}}(k) := \mathbb{E}\left[\mathcal{M}_{t+1}(k)\right] =&\pi_1\,\mathbb{E}\left[e^{k\,\alpha_1\,g_{t+1}}\right]\; e^{k\,(1-\alpha_1)\,\mu_1+\frac{1}{2}\,k^2\,\sigma_1^2}+\\ &\pi_2\,\mathbb{E}\left[e^{k\,\alpha_2\,g_{t+1}}\right]\; e^{k\,(1-\alpha_2)\,\mu_2+\frac{1}{2}\,k^2\,\sigma_2^2}\\ =&\pi_1\,e^{\mathcal{K}^*(k\,\alpha_1)+k\,(1-\alpha_1)\,\mu_1+\frac{1}{2}\,k^2\,\sigma_1^2}\\ &+\pi_2\,e^{\mathcal{K}^*(k\,\alpha_2)+k\,(1-\alpha_2)\,\mu_2+\frac{1}{2}\,k^2\,\sigma_2^2}.\end{aligned} \tag{32}$$

Equation (32) shows that the expected belief of an agent with similarity-weighted memory is constant over time as a result of the i.i.d. assumption on the agent's context g_t. 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 posterior state-wise means are higher than the true means, then an econometrician expects a downward revision of the agent's beliefs in the next period. The mean-reversion of the agent's beliefs translates into mean-reversion of asset prices, and thus into predictably countercyclical returns (Shiller, 1981).

Summary. The agent's subjective beliefs under similarity-weighted memory are time-varying, although the agent has access to infinite data. Consistent with empirical evidence, the agent's posterior mean varies procyclically and overreacts to new information. In addition, the framework predicts that the agent perceives the economy as excessively risky during extremely good as well as bad times. Moreover, I find that the agent's beliefs are predictably mean-reverting. The next section embeds these beliefs into the asset pricing model.

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. 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 (Equation (21)) and the results from Section 2.4. I focus on power utility for intuition. The results for Epstein and Zin (1989)-preferences are in Appendix A.6.

Proposition 6 (Asset prices under similarity-weighted memory)**.** *Focus on power utility ($\psi = 1/\gamma$). Under similarity-weighted memory as in Equation* (17) *and an i.i.d. two-state Markov-switching process for log endowment growth, it is*

$$r_t^f = -\log(\beta) - \log\left(\pi_1\, e^{-\gamma\,\hat{\mu}_{1,t}+\frac{1}{2}\gamma^2\,\sigma_1^2} + \pi_2\, e^{-\gamma\,\hat{\mu}_{2,t}+\frac{1}{2}\,\gamma^2\,\sigma_2^2}\right), \tag{33}$$

$$dp_t = -\log(\beta) - \log\left(\pi_1\, e^{(\lambda-\gamma)\,\hat{\mu}_{1,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_1^2} + \pi_2\, e^{(\lambda-\gamma)\,\hat{\mu}_{2,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_2^2}\right), \tag{34}$$

$$\hat{er}_t = dp_t + \log\left(\pi_1\, e^{\lambda\,\hat{\mu}_{1,t}+\frac{1}{2}\,\lambda^2\,\sigma_1^2} + \pi_2\, e^{\lambda\,\hat{\mu}_{2,t}+\frac{1}{2}\,\lambda^2\,\sigma_2^2}\right), \tag{35}$$

$$\begin{aligned}\hat{rp}_t = {}& \log\left(\pi_1\, e^{-\gamma\,\hat{\mu}_{1,t}+\frac{1}{2}\gamma^2\,\sigma_1^2} + \pi_2\, e^{-\gamma\,\hat{\mu}_{2,t}+\frac{1}{2}\,\gamma^2\,\sigma_2^2}\right) + \log\left(\pi_1\, e^{\lambda\,\hat{\mu}_{1,t}+\frac{1}{2}\,\lambda^2\,\sigma_1^2} + \pi_2\, e^{\lambda\,\hat{\mu}_{2,t}+\frac{1}{2}\,\lambda^2\,\sigma_2^2}\right) \\ & - \log\left(\pi_1\, e^{(\lambda-\gamma)\,\hat{\mu}_{1,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_1^2} + \pi_2\, e^{(\lambda-\gamma)\,\hat{\mu}_{2,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_2^2}\right).\end{aligned} \tag{36}$$

The risk-free rate r_t^f, in Equation (33)—which is determined by the agent's subjective beliefs—increases with contemporaneous log endowment growth g_t due to the procyclicality of the agent's endowment growth expectations. The posterior state-wise means $\hat{\mu}_{1,t}$ and $\hat{\mu}_{2,t}$ both covary positively with g_t, so that the cumulant-generating function at $-\gamma$, $\mathcal{K}_t(-\gamma) = \log\left(\pi_1\, e^{-\gamma\,\hat{\mu}_{1,t}+\frac{1}{2}\gamma^2\,\sigma_1^2} + \pi_2\, e^{-\gamma\,\hat{\mu}_{2,t}+\frac{1}{2}\,\gamma^2\,\sigma_2^2}\right)$ declines in g_t. Since $\mathcal{K}_t(-\gamma)$ enters Equation (33) negatively, the risk-free rate increases in g_t. Intuitively, when the current context g_t is high, the agent selectively recalls past experiences with a high log endowment growth due to similarity and forms optimistic beliefs. Being optimistic, the agent has a low incentive to save for the future, such that the risk-free rate must be high to make saving attractive. The model thus yields a procyclical risk-free rate, as is consistent with evidence (Adam and Nagel, 2023). Moreover, the volatility of the agent's beliefs leads to a volatile risk-free rate as empirically observed by Jordà et al. (2019).

Second, the dividend-price ratio of the λ-asset dp_t declines with contemporaneous log endowment growth g_t for $\lambda > \gamma$. As discussed in Section 2.4, the asset's price increases with the agent's posterior mean if leverage λ exceeds the curvature of the utility function, given by γ.[16] Because similarity-weighted memory leads to procyclical cash-flow expectations, the model predicts countercyclical dividend-price ratios and, equivalently, procyclical price-dividend ratios. Empirically, as well as in the model, the price-dividend ratio is positively correlated with cash-flow expectations (De La O and Myers, 2021; Bordalo et al., 2024).

Third, the subjectively expected return, which depends on the dividend-price ratio dp_t and on the subjectively expected cash-flow growth $\mathcal{K}_t(\lambda)$, increases with contemporaneous log endowment growth g_t. The subjective cash-flow growth expectation $\mathcal{K}_t(\lambda)$ increases in the agent's posterior mean and thus in the contemporaneous log endowment growth g_t. When $\gamma > \lambda$, the dividend-price ratio also increases in g_t, yielding a procyclical subjective expected return. When $\gamma < \lambda$, instead, the dividend-price ratio decreases in g_t. However, $\lambda > \gamma > 0$ implies that the cash-flow expectations dominate the discount-rate effect, such that the subjectively expected return is procyclical in this case as well. Intuitively, the agent is optimistic after observing a high contemporaneous log endowment growth, such that the risky asset needs to yield a high expected return to induce investment.[17] Procyclical subjective return expectations are consistent with survey evidence (Amromin and Sharpe, 2014; Greenwood and Shleifer, 2014).

Fourth, the subjective risk premium $\hat{rp}_t$ depends on the convexity of the cumulant-

[16] 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.

[17] An increase in g_t rotates $\mathcal{K}_t(k)$. As the CGF is convex (Martin, 2013), the effect of an increase of g_t on the subjective expected return is not necessarily monotone, and Figure 2 shows that the expected return is indeed convex in g_t. The statements in the main text hold for a second-order approximation of $\mathcal{K}_t(k)$.

generating function over $[-\gamma, \lambda]$ (Martin, 2013). Under similarity-weighted memory, this convexity changes over time with g_t. Consider the second-order approximation of the subjective cumulant-generating function, $\hat{rp}_t \approx \lambda \, \gamma \, \hat{\sigma}_t^2$. The subjective risk premium of the agent inherits, up to a second-order approximation, the convexity of the unconditional posterior variance. Behaviorally, the agent demands a higher risk premium in exceptionally good as well as bad times, and a low risk premium in average times. The convexity of the subjective risk premium is a new, testable prediction of the model. In addition, the subjective risk premium thus varies with the subjective riskiness, and the convexity of the subjective risk premium implies that $\hat{rp}_t$ is approximately acyclical with respect to the price-dividend ratio—which varies roughly linearly with g_t. Both properties are consistent with empirical findings by Nagel and Xu (2023).

Figure 2 displays the risk-free rate, price-dividend ratio, expected return, and subjective risk premium. In contrast to the discussion so far, I consider a specification of Epstein and Zin (1989)-preferences with $\psi = 1.5 \neq 1/\gamma$.[18] The qualitative properties of asset prices discussed above for power utility continue to hold. Under similarity-weighted memory, the risk-free rate and the price-dividend ratio increase in the context g_t, while the subjectively expected return is a convex function of g_t due to the convexity of the posterior variance. Similarly, the subjective risk premium is a convex function of the context g_t. Under rational expectations, instead, the risk-free rate, price-dividend ratio, subjective expected return, and subjective risk premium are constant due to the i.i.d. assumption.

Objective asset prices. The agent's beliefs are predictable by an outside observer with access to the same data as the agent, as shown in Section 3.2. The objectively realized return and risk premium therefore deviate from their subjective counterparts. Using a Campbell-Shiller decomposition as in Campbell (1991), the objective risk premium is (Appendix D)

$$\mathbb{E}\left(r_{t+1}\right) - r_t^f = \lambda \, \mathbb{E}(g_{t+1}) + \mathcal{K}_t(-\gamma) - \mathcal{K}_t(\lambda - \gamma) - \frac{\bar{p}}{1-\bar{p}} \left(\mathbb{E}\left(dp_{t+1}\right) - dp_t\right), \qquad (37)$$

[18] Importantly, the parameters imply that the agent cares more about the resolution of uncertainty than about predictable variation in consumption, which becomes important in the simulations below.

Figure 2: Qualitative asset pricing implications of similarity-weighted memory

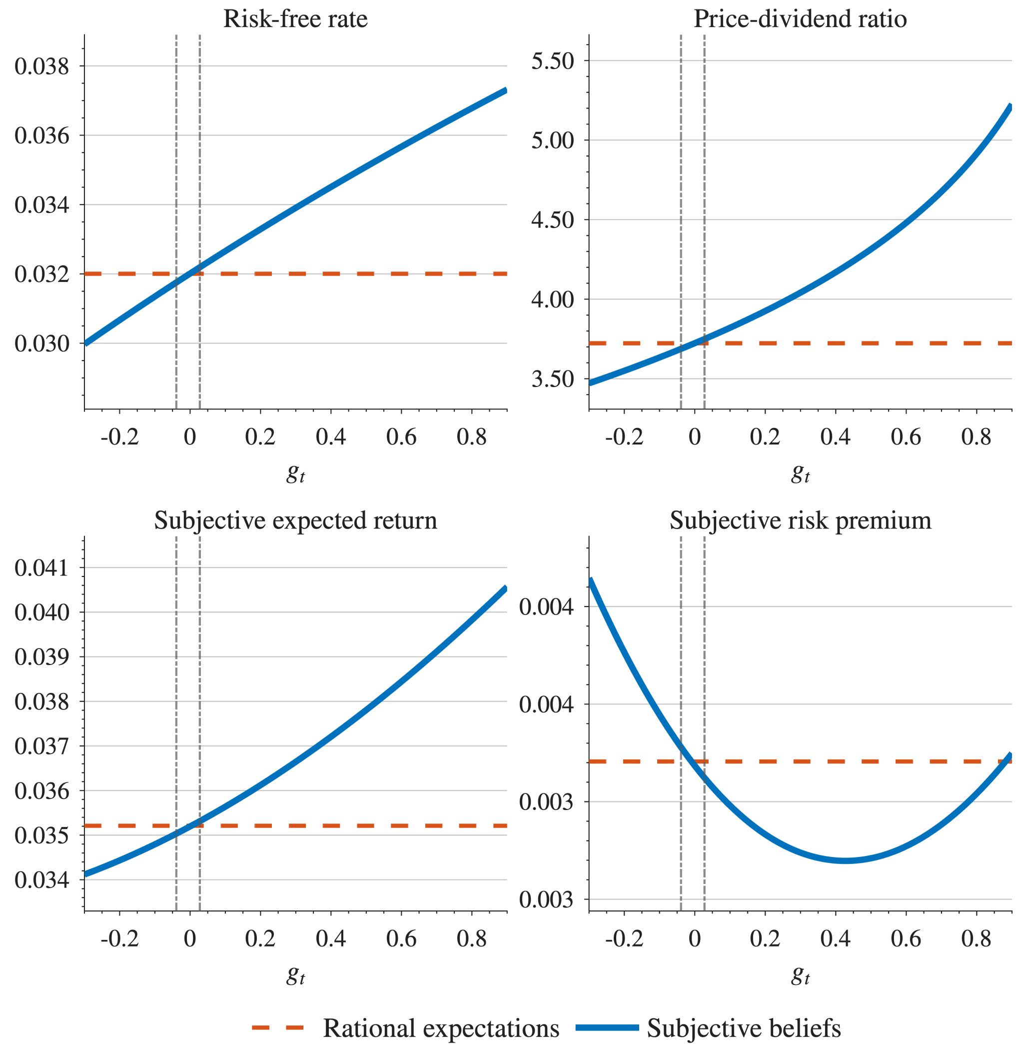

Figure 2 shows the risk-free rate, price-dividend ratio, subjectively expected return and subjectively expected risk premium for different values of the context g_t under rational expectations (dashed line) and under similarity-weighted memory (solid line). The parameters of the log endowment growth process are as in Table 1, and the preference parameters are $\beta = 0.97, \gamma = 10, \psi = 1.5, \kappa = 0.01$. The vertical dashed lines show the 0.5% and 99.5% quantiles of the log endowment growth process.

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 $\bar{p} = \frac{1}{1+\exp\left(\overline{d-p}\right)}$ 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 $\mathbb{E}(g_{t+1}) = \pi_1\,\mu_1 + \pi_2\,\mu_2$ deviates from the subjectively expected endowment growth $\hat{\mu}_t = \pi_1\,\hat{\mu}_{1,t} + \pi_2\,\hat{\mu}_{2,t}$. Second, the econometrician expects a revision of the agent's beliefs to their long-term mean, which affects the (future) 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 sizable fluctuations of 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.[19] Under log-normality, it is

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

As in Nagel and Xu (2022), the objective risk premium has three components: (i) the subjective risk premium, which is the priced risk compensation required by the representative agent in equilibrium; (ii) a sentiment premium that is driven by the time-varying belief wedge $\mu - \hat{\mu}_t$ and depends on the leverage of the asset λ as well as on the inverse of the EIS ψ^{-1}; and (iii) a Jensen adjustment.

If the agent is too optimistic, $\hat{\mu}_t > \mu$, the econometrician expects a mean reversion of the agent's beliefs towards μ and a low return next period. In good times, the agent becomes overly optimistic about future cash-flows, such that the asset price is high. The agent's cash-flow expectation tends to be disappointed next period (see Proposition 5), such that returns are low following good times.

Figure 3 confirms these dynamics for the two-state setting. When the context g_t is high, the agent is very optimistic and the realized risk premium will, on average, be low. Equation (37) shows that the price-dividend ratio (the reciprocal of dp_t) negatively predicts the realized risk premium. The objective risk premium is low if g_t is high (Figure 3), which corresponds

[19] Appendix D shows how to approximate the expected subjective cumulant-generating function for the two-state Markov-switching process.

Figure 3: Objective and subjective risk premium

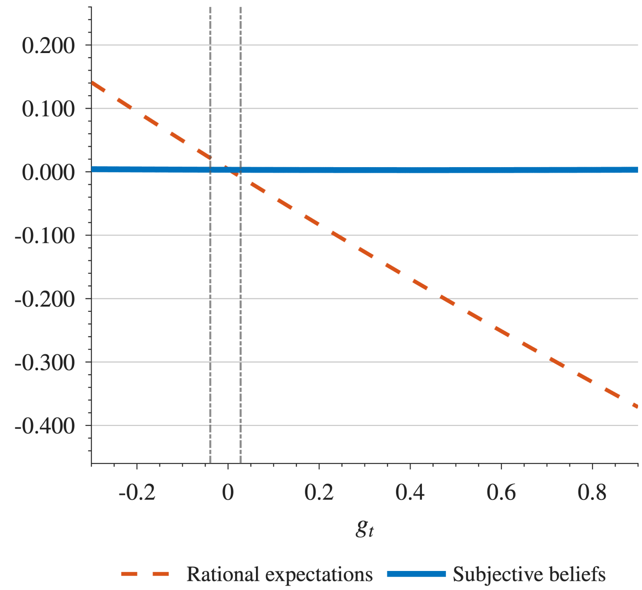

Figure 3 shows the objective (orange dashed line) and subjective (blue solid line) risk premium for different realizations of the contemporaneous log endowment growth g_t. The parameters of the log endowment growth process are as in Table 1, and the preference parameters are $\beta = 0.97, \gamma = 10, \psi = 1.5, \kappa = 0.01$. The vertical dashed lines show the 0.5% and 99.5% quantiles of the log endowment growth process.

to times in which the price-dividend ratio is high (Figure 2). The objective risk premium is thus predictably countercyclical using aggregate valuation ratios (Shiller, 1981; Campbell and Shiller, 1988; Nagel and Xu, 2023).

Importantly, Figure 3 also highlights that the subjective risk premium varies significantly less than the objective risk premium and is effectively acyclical. Empirically, the objective risk premium is two orders of magnitude more volatile than the subjective risk premium (Nagel and Xu, 2023). In the framework here, the acyclicality of the subjective risk premium in comparison to the cyclicality of the objective risk premium arises because the time-variation of the perceived riskiness of the economy is a second-order effect here, while variation of the first moment of beliefs is first-order. This mechanism offers an alternative behavioral microfoundation for the observed dichotomy of subjective and objective risk premia (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 and simulation

I now turn to a quantitative analysis of the model using simulations. After describing the (estimated) parameters of the endowment growth process and the agent's preference specification, I discuss the simulated asset prices. Extensions are in Section 4.

Parameters. Table 1 summarizes the parameters used for the simulations. There are two types of parameters: endowment growth and preference/belief parameters. I estimate the parameters of the i.i.d. two-state Markov-switching process for log endowment growth (Equation (16)) using Bayesian methods (see Appendix E and Johannes et al. (2016)). I measure endowment growth using the quarterly growth of services and non-durable consumption from the Bureau of Economic Analysis from Q1 1947 to Q2 2025 and report the *annualized* Bayesian maximum a-posterior parameters among $10,000$ parameter combinations in Table 1. To ensure equilibrium existence for a wide range of beliefs, I lower the mean growth rate in the recession state, μ_2, from -0.28% to -2.80%. As a result, the simulated economies exhibit an average annual growth rate of 1.47%, compared with 1.79% in the data. Furthermore, I set the leverage parameter $\lambda = 3$ as in Collin-Dufresne et al. (2016) and Nagel and Xu (2022). Estimating λ by regressing quarterly aggregate dividends obtained from CRSP on endowment growth, I find $\lambda \approx 3.31$.

The preference parameters are chosen to be close to values used in the literature while

Table 1: Calibration parameters

Parameter	Symbol	Value	Source
Endowment growth process			
Mean growth in			
State 1	μ_1	1.92%	Estimated
State 2	μ_2	−2.80%	Calibrated
Growth volatility in			
State 1	σ_1	1.52%	Estimated
State 2	σ_2	3.90%	Estimated
Probability of state 1	π_1	90.05%	Estimated
Leverage	λ	3.00	Collin-Dufresne et al. (2016)
Preferences and memory			
Risk aversion	γ	10	Jin and Sui (2022)
EIS	ψ	1.7	Gruber (2013)
Time discount factor	β	0.993	-
Memory scrutiny	κ	0.02	-

Table 1 reports the parameters used in the simulation. The parameters of the endowment growth process are estimated using Bayesian methods (Johannes et al., 2016) and annualized by multiplying means by four and standard deviations by 2. I mostly set the preference parameters to ensure equilibrium existence and use values from Jin and Sui (2022) and Gruber (2013). Memory scrutiny κ is set to be of the same magnitude as the volatility of endowment growth.

ensuring equilibrium existence in the i.i.d. economy considered here.[20] I set the coefficient of relative risk-aversion $\gamma = 10$ (Jin and Sui, 2022) and the elasticity of intertemporal substitution (EIS) to $\psi = 1.7$ (Gruber, 2013).[21] I set the discount factor $\beta = 0.993$, which is lower than in related work by Nagel and Xu (2022) and Jin and Sui (2022) to ensures equilibrium existence. Finally, I set the memory scrutiny parameter $\kappa = 0.02$, so that memory scrutiny is of the same order of magnitude as the volatility of endowment growth. Setting $\kappa = 0.02$ implies $\alpha_1 = 0.003$ and $\alpha_2 = 0.019$, such that the agent's posterior mean is

[20] I also restrict the agent's beliefs used to price assets to the range consistent with equilibrium existence. For the parameter values in Table 1, I compute the upper and lower bounds of the posterior means that support an equilibrium and restrict pricing-relevant beliefs to remain within this range (Collin-Dufresne et al., 2016). Using $\kappa = 0.02$, the restriction is binding in 5.4% of all simulated periods. *Subjective* beliefs and asset prices in Table 2 are not restricted to ensure equilibrium existence. It is an open question if survey beliefs are consistent with equilibrium considerations (Andre et al., 2025; Bastianello and Fontanier, 2025).

[21] Values of γ between 4 and 10 are standard in the long-run risk literature, and Mehra and Prescott (1985) argue that values up to and around ten are reasonable. Estimates of the EIS vary across studies, but most empirical work reports values above, but close to one (Beeler and Campbell, 2012). Setting ψ to one yields a higher risk-free rate of 3.86%, but leaves the subjective and objective risk premium unchanged.

Figure 4: Posterior long-term beliefs

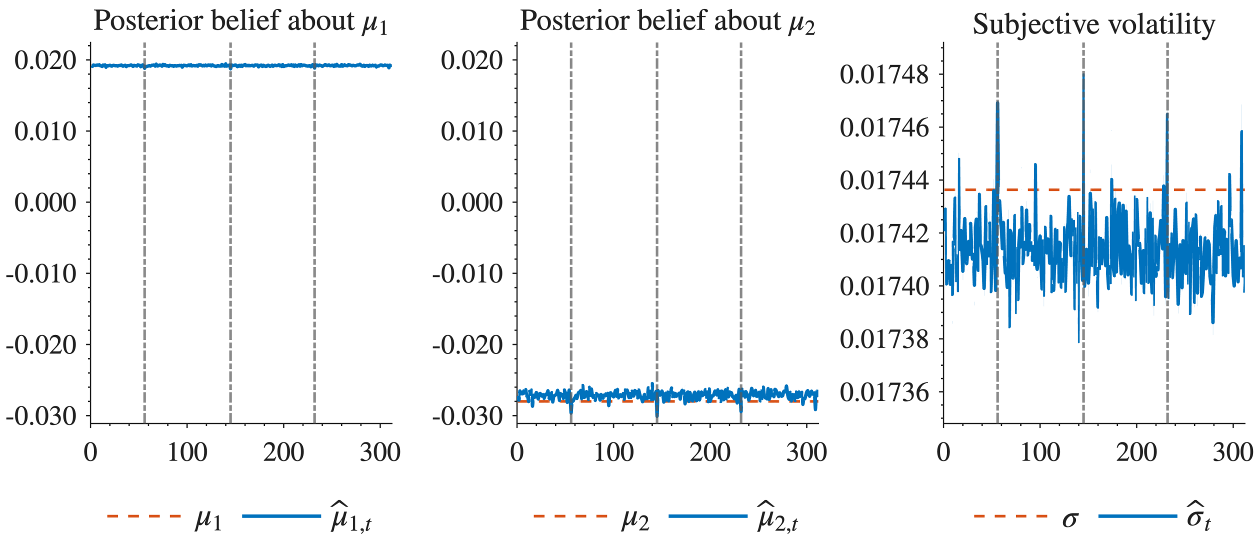

Figure 4 shows the annualized posterior beliefs of the agent for one realization of log endowment growth. The left (middle) panel shows the posterior mean of log endowment growth in the good state (bad state) and the right panel shows the subjective volatility. The dashed line in each panel plots the respective quantity under rational expectations, and the vertical lines mark realized bad states. The parameters are as in Table 1, except that I use $p_1 = 0.96$. Using $p_1 = 0.96$ implies that the subjective volatility is quantitatively different from that reported in Table 2.

not overly sensitive to g_t.[22]

Simulation. I simulate the model quarterly to match the frequency of consumption data.[23] Table 2 reports the annualized average moments from 50, 000 simulations of the model, each for 312 quarters, and shows that the model qualitatively matches salient facts about asset prices.

Figure 4 plots the agent's posterior beliefs for one realization of the endowment growth process. Together with Table 2, it highlights that the posterior mean of the more frequently observed good state, $\hat{\mu}_{1,t}$, is approximately unbiased and constant. In contrast, the posterior mean of the rare recession state $\hat{\mu}_{2,t}$ is biased upward and more volatile. The bottom row of Table C.1 reports the coefficient of a Coibion and Gorodnichenko (2015)-regression as in

[22]For $\kappa = 0.2\,(0.002,\ 0.0002)$, it is $\alpha_1 = 0.0003\,(0.0281,\ 0.2241)$ and $\alpha_2 = 0.0019\,(0.1598,\ 0.6553)$.

[23]A quarterly calibration implies that the agent's memory context updates quarterly. The model can be simulated at higher or lower frequencies, and also with temporal variation in the memory context.

Table 2: Simulation results with baseline parameters

Symbol	Subjective			Objective		
	Total	Normal	Recession	Total	Normal	Recession
Endowment growth						
Mean	1.457	1.919	-2.721	1.451	1.919	-2.794
Std $(\hat{\mu}_t)$	0.009	0.006	0.038	-	-	-
Volatility	2.019	2.019	2.022	2.017	-	-
$\text{corr}(\hat{\sigma}_t, g_t)$	-1.000	-1.000	-1.000	-	-	-
Asset prices						
$\overline{er_t}$	4.611	4.612	4.604	-	-	-
Std (er_t)	0.004	0.003	0.007	-	-	-
corr (er_t, g_t)	1.000	1.000	1.000	-	-	-
$\overline{r_t^f}$	-	-	-	3.334	3.336	3.321
$\text{Std}(r_t^f)$	-	-	-	0.011	0.008	0.021
$\text{corr}(r_t^f, g_t)$	-	-	-	0.987	0.991	0.980
$\overline{rp_t}$	1.277	1.276	1.283	1.076	0.806	3.528
$\text{Std}(rp_t)$	0.002	0.002	0.004	1.160	0.873	2.221
$\text{corr}(rp_t, g_t)$	-0.910	-0.934	-0.851	-1.000	-1.000	-1.000

Table 2 shows the moments obtained from $50,000$ simulations of the model for 312 quarters under the assumption that the agent has an infinite sample of observations. 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.

Proposition 5 and indicates overreaction of the agent's beliefs ($b_{CG} < 0$).[24]

The agent's posterior variance (right panel of Figure 4) tends to spike during recessions but also fluctuates in normal times. The average subjective volatility of the economy is low (2.019) but slightly higher in recessions (2.022) than in normal times (2.019). The strong negative correlation of the subjective volatility with endowment growth arises because almost all realizations of endowment growth lie on the decreasing part of the globally convex subjective volatility (Proposition 4).[25] Using a measure of the perceived riskiness from the Graham-Harvey CFO Survey compiled by Nagel and Xu (2023),[26] I also find countercyclical

[24]Using conventional t-Statistics, I cannot reject the null that $b_{CG} = 0$ in the baseline simulations. In unreported results, I confirm that increasing the sample size leads to $b_{CG} \to -0.5$ with high t-Statistics.

[25]Given the parameters, the subjective volatility is minimized at $g^* = \frac{(1-\alpha_1)\,\mu_1-(1-\alpha_2)\,\mu_2}{\alpha_2-\alpha_1} \approx 0.7387$.

[26]I thank the authors for providing the data.

variation of the perceived riskiness: the average perceived riskiness is higher in recessions (15.57) than in normal times (12.84, see Table E.2).[27]

A new prediction of the model is that the perceived riskiness of the economy is a convex function of economic conditions: the agent perceives the economy as risky in extraordinarily good as well as bad times. Regressing perceived volatility on log consumption growth and its square in the simulated data yields a significantly negative coefficient on the linear term and a significantly positive coefficient on the squared term. Empirically, the perceived volatility is higher in the lowest and highest quintiles of log consumption growth than in the middle quintiles (Table E.2), suggesting a convex relation between economic conditions and perceived riskiness of the economy. Moreover, running the same regression in survey data (Table E.3), I find a positive, marginally significant squared term and a negative, but insignificant linear term—broadly consistent with the model's prediction.

Table 2 shows that the risk-free rate and the subjectively expected return are procyclical (positive correlation with endowment growth), as is consistent with empirical findings. The average risk-free rate varies slightly around 3.33%, but is relatively constant across normal times (3.34%) and recessions (3.32%). Empirically, the log risk-free rate is at about 3.80% in postwar data (see Table E.2),[28] and varies across normal times (3.98%) and recessions (2.72%).

The subjective risk premium in the simulations is low (1.277%), almost constant (average standard deviation of 0.002), does barely vary across normal times (average of 1.276%) and recessions (average of 1.283%), and is not predictable by the dividend-price ratio (Table C.1). Using data by Nagel and Xu (2023), the subjective risk premium of individual investors—measured using log return expectations minus the log risk-free rate—is higher than in the simulations with 5.63%, but it also barely varies across normal times (5.59%) and recessions

[27] I measure *recessions* using the posterior probability of being in the good state under the maintained assumption of a two-state i.i.d. Markov chain with the same parameters as in Table 1. A period is a recession if the probability of being in the good state is below 0.5. The recession measure has a high overlap with NBER recession periods, but assigns quarters with an extremely high consumption growth to a recession, as is consistent with the model.

[28] I use the 3-month T-bill rate from FRED.

(6.46%, see Table E.2). In the model, subjective risk premia inherit the time-variation of the subjective volatility due to the positive risk-return trade-off under the subjective measure (Nagel and Xu, 2023). Empirically, I find a significantly positive coefficient on the squared term when regressing the subjective risk premium on log consumption growth and its square. This indicates a convex relationship between the subjective risk premium and economic conditions (see Table E.3), consistent with the model's prediction.

The objective risk premium is more volatile than the subjective risk premium (standard deviation of 1.160%) and slightly lower than the subjectively expected risk premium with an average of 1.076% in the simulations. The countercyclicality of the objective risk premium is pronounced, since the objective risk premium is lower in normal times (0.806%) than in recessions (3.528%).[29] Table C.1 shows that the objective risk premium is predictably countercyclical using the dividend-price ratio ($\hat{b}_{obj} = 0.60$). The comparably high volatility and the predictability of the objective risk premium are consistent with empirical findings (Campbell and Shiller, 1988; Nagel and Xu, 2023).

Table C.2 in Appendix C shows the average subjective and objective risk premium for each quintile of log endowment growth. The subjective risk premium is almost constant across quintiles, but the objective risk premium is high in the lowest quintile of endowment growth (4.36%), and becomes negative in the highest quintile (−1.75%), 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 generate key qualitative patterns found in the data, but the realized risk premium is comparably low. As shown in Section 3.3, the objective risk premium depends on the predictability of the agent's subjective beliefs. When the agent has already observed an infinitely long history and the memory context depends solely on contemporaneous endowment growth, the agent's beliefs are too stable to generate a high realized risk premium. Combined with i.i.d. fundamentals,

[29] Empirically, the objective risk premium is higher than in the simulations, with an average of 6.07% that varies countercyclically (10.0% in recessions and 5.63% in normal times).

the model yields fast mean reversion and little persistence in the risk premium, the price-dividend ratio, or the subjective cash-flow beliefs—in contrast to empirical findings.

Overall, relaxing rational expectations through similarity-weighted memory can qualitatively account for the observed cyclicality, predictability, and sensitivity to risk of both subjective and objective asset prices. The model especially generates predictability and time-variation that is consistent with empirical results, even though all dynamics in the model are driven by i.i.d. fundamentals. The i.i.d. dynamics, however, also imply that the model does not match other important empirical moments. The next section considers extensions of the framework that yield additional dynamics.

4 Extensions and discussion

4.1 Limited history

In this section, I consider an extension in which the agent learns from a finite sample of past observations. Finite-sample learning introduces Bayesian parameter uncertainty, which endogenously generates long-run risk and a high price of risk under Epstein and Zin (1989)-preferences (Collin-Dufresne et al., 2016). I show that combining parameter uncertainty with similarity-weighted memory produces a realistic risk premium while retaining the qualitative patterns documented for the baseline model. Moreover, limited data induces persistence in beliefs and asset prices, as the agent updates more slowly when relying on limited samples. The persistence gradually declines as the observed history lengthens.

When the agent learns from finitely many observations, beliefs and asset prices must be simulated (Appendix F.1). The analytical characterization of beliefs as maximizers of the memory-weighted likelihood (Equation (2)) relies on asymptotic convergence, which no longer holds when the agent's history is finite. I therefore simulate the model $50,000$ times for 312 quarters. For every quarter in each simulation, I draw the agent's recalled history according to the similarity-weighted memory function (Equation (17)). I focus on an

Table 3: Simulation results with baseline parameters and parameter uncertainty

Symbol	Subjective Total	Subjective Normal	Subjective Recession	Objective Total	Objective Normal	Objective Recession
Endowment growth						
Mean	1.467	1.918	−2.615	1.452	1.922	−2.801
Std ($\hat{\mu}_t$)	0.055	0.045	0.335	-	-	-
Volatility	2.025	2.024	2.034	2.018	–	–
corr($\hat{\sigma}_t, g_t$)	−0.082	−0.019	−0.270	-	-	-
Asset prices						
$\overline{er_t}$	18.949	19.099	17.590	-	-	-
Std (er_t)	6.086	6.135	5.370	-	-	-
corr (er_t, g_t)	0.050	0.051	0.066	-	-	-
$\overline{r_t^f}$	-	-	-	3.229	3.233	3.194
Std(r_t^f)	-	-	-	0.208	0.205	0.230
corr(r_t^f, g_t)	-	-	-	0.093	0.062	0.182
$\overline{rp_t}$	15.720	15.866	14.396	3.117	−0.430	35.212
Std(rp_t)	6.160	6.208	5.451	25.739	24.553	29.572
corr(rp_t, g_t)	0.050	0.050	0.064	−0.247	−0.137	−0.402

Table 3 reports the moments obtained from 50,000 simulations of the model for 312 quarters plus a 120 quarter burn-in period when the agent learns from a finite sample, such that the Bayesian posterior has a strictly positive variance around the parameter values. 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.

uninformative prior and include 120 quarters of log endowment growth as a burn-in period, so that the agent enters the market with 30 years of historical data.[30]

Table 3 reports model-implied moments based on 50,000 simulations, and Figure 5 plots one realization of the agent's beliefs. Relative to the $t \to \infty$ benchmark, learning from a finite sample generates more volatile expectations but similar long-run averages. Similarly, the properties of the subjective volatility discussed in Section 3.4 hold under parameter uncer-

[30]Pricing-relevant beliefs are restricted to be consistent with equilibrium under full-information rational expectations. The restriction binds in roughly 40% of simulated quarters, which primarily arises when the sample mean of the observed history exceeds the upper bound. As before, I report subjective beliefs and the subjectively expected return without the equilibrium restriction. Imposing the restriction on subjectively expected returns yields a biased estimate of the correlation between the subjectively expected return and the contemporaneous log endowment growth, which is attenuated and can even become negative. Relaxing the range of beliefs that support equilibrium also yields procyclical subjective return expectations (Table C.4).

Figure 5: Posterior beliefs with parameter uncertainty

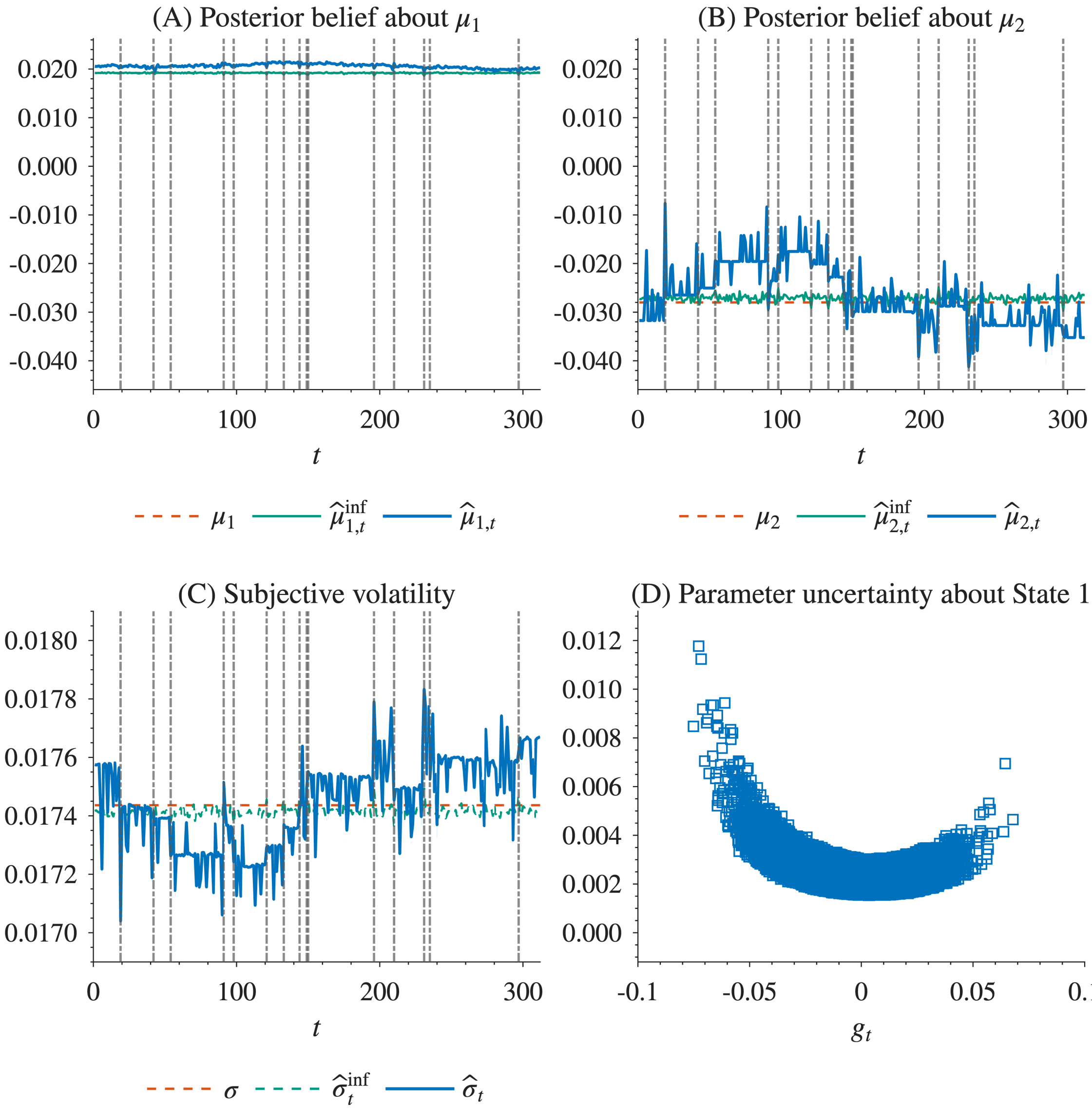

Figure 5 shows one realization of the agent's posterior beliefs when allowing for parameter uncertainty. Panel (A) plots the subjective mean for the good state ($\hat{\mu}_{1,t}$), panel (B) for the recession state ($\hat{\mu}_{2,t}$), panel (C) the subjective volatility, and panel (D) the parameter uncertainty. In Panels (A)-(C), the solid blue line shows the posterior belief with a finite number of observations, the green dashed line shows beliefs in the $t \to \infty$ limit and the orange dashed line marks the true parameter. Vertical gray lines mark recession periods. The parameters are as in Table 1, except for $p_1 = 0.96$ and a 360-quarter burn-in period. Panel (D) plots the uncertainty about the mean in state 1 against the contemporaneous endowment growth g_t across $1,000$ simulations. I increase the sensitivity of beliefs to the current context by setting $\kappa = 0.002$ for panel (D) only.

tainty with a lower, but still negative, correlation of subjective volatility and log endowment growth. The bottom row of Table C.1 shows that the agent's beliefs overreact more when learning from a limited sample, with a Coibion and Gorodnichenko (2015)-regression coefficient of $\hat{b}_{CG} = -0.470$. Moreover, Panel (D) of Figure 5 shows that parameter uncertainty—which is priced under Epstein and Zin (1989)-preferences—is convex in the current context. Economically, the agent is more uncertain about her beliefs in extremely good as well as bad times, generating a convex priced uncertainty that mirrors the convexity of the subjective volatility discussed in Section 3.2.

The invariance of the qualitative features of beliefs when moving from an infinite to a finite history implies that the qualitative features of asset prices discussed in Section 3.4 also hold under parameter uncertainty. Quantitatively, the model generates an average objective risk premium of 3.12%, roughly consistent with the results reported by Jin and Sui (2022) and Campbell and Cochrane (1999). Using long-term data by Jordà et al. (2019), I estimate a risk premium between 1.07% (Portugal) and 7.33% (Germany), with the US being at 4.46%. Moreover, I find an average risk premium of 5.63% in normal times and of 10.00% in recessions in the data by Nagel and Xu (2023) (see Table E.2). The model similarly produces a higher risk premium in recessions (35.21%) than in normal times (−0.43%), but the magnitude of the difference is larger than in the data.

Learning from a limited history also induces persistence in beliefs and valuation ratios. As shown in Table C.3, the model generates persistence in both beliefs and the price-dividend ratio, while excess returns exhibit fast mean reversion. The price-dividend ratio is somewhat more persistent than in the data (Campbell and Cochrane, 1999; Jin and Sui, 2022), but in a plausible range.

Learning from a finite-history under similarity-weighted memory improves the model's quantitative performance—raising the risk premium and introducing persistence—while preserving the qualitative features of beliefs and asset prices. The improvement comes at the cost of computational complexity and the loss of closed-form solutions. Over time, the economy with learning from a finite history converges to the infinite-history case analyzed in Section 3.4, such that the persistence gradually dissipates. I next extend the model to incorporate a persistent memory context, which gives a fundamental reason for persistence.

4.2 Persistent context

In this section, I extend the model by introducing a persistent memory context. Section 3 assumes that the context is given by today's log endowment growth g_t. I now generalize the memory function to

$$m^{\text{pers}}_{(c_t,s_t)}(g_\tau, s_\tau) = \exp\left[-\frac{(g_\tau - c_t)^2}{2\,\kappa}\right], \tag{38}$$

where context c_t evolves according to

$$c_t = \phi\, c_{t-1} + (1-\phi)\, g_t. \tag{39}$$

The agent's state-wise posterior mean is thus given by

$$\hat{\mu}_{s,t} = (1-\alpha_s)\mu_s + \alpha_s c_t, \tag{40}$$

so that persistence in c_t translates directly into persistence in beliefs and asset prices.

Figure 6 illustrates the resulting belief dynamics for the same realization of log endowment growth as in Figure 4. Relative to the i.i.d. benchmark, the posterior mean of the recession state $\hat{\mu}_{2,t}$ is more persistent and consistently above its true value μ_2. The higher persistence of both posterior means leads to a low fluctuation of the subjective volatility, which is generally below the true volatility σ_2.

Table 4 summarizes the results of $50,000$ simulations for 312 quarters under the parameters as in Table 1, but with a persistent memory context. I calibrate the persistence parameter $\phi = 0.85$ to match the empirical persistence of the annualized price-dividend ratio reported by Campbell and Cochrane (1999). The asset pricing moments are quantitatively similar to those obtained under the i.i.d. context (Table 2), but with somewhat lower volatil-

Figure 6: Posterior long-term beliefs with persistent context

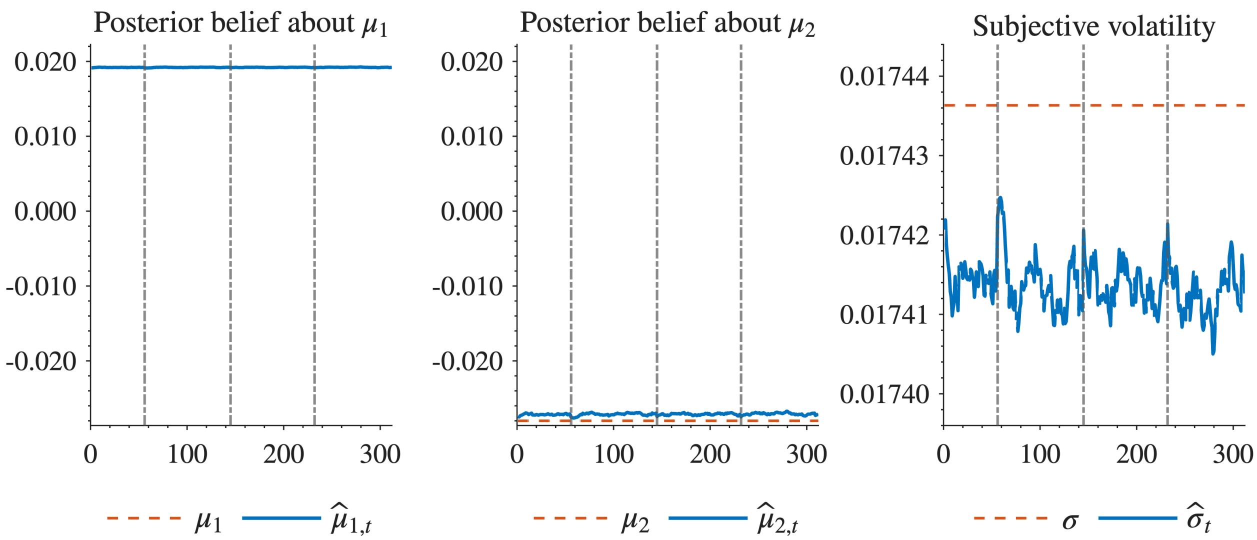

Figure 4 shows the annualized posterior beliefs of the agent for one realization of log endowment growth. The left (middle) panel shows the posterior mean of log endowment growth in the good state (bad state) and the right panel shows the subjective volatility. The dashed line in each panel plots the 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 $p_1 = 0.96$. Moreover, I use $c_t = \phi c_{t-1} + (1-\phi)g_t$ with $\phi = 0.85$.

ity and cyclicality in the risk-free rate, expected return, and both subjective and objective risk premia. Simulating the model with a persistent context and learning from a limited sample yields a realized risk premium of 3.12%, with a more reasonable variation across normal times (1.22%) and recessions (20.25%) than under an i.i.d. context.

The persistent memory context generates two additional patterns. First, it produces persistence in the annualized price-dividend ratio that closely matches the values documented by Campbell and Cochrane (1999) (see Table C.3). Second, it leads to slow mean reversion of the realized risk premium, implying that valuation ratios predict future returns even at long horizons. Regressing k-year returns on the current price-dividend ratio yields coefficients that gradually decline in magnitude: -0.024 at a one-year horizon ($R^2 = 0.84$), -0.021 at three years ($R^2 = 0.72$), and -0.010 at seven years ($R^2 = 0.18$). Empirically, the predictability of returns tends to be increasing with the forecast horizon.

The model's ability to accommodate a persistent context shows the flexibility and appli-

Table 4: Simulation results with persistent context and baseline parameters

Symbol	Subjective			Objective		
	Total	Normal	Recession	Total	Normal	Recession
Endowment growth						
Mean	1.457	1.919	−2.721	1.451	1.919	−2.794
Std ($\hat{\mu}_t$)	0.003	0.002	0.011	-	-	-
Volatility	2.019	2.019	2.019	2.017	–	–
corr($\hat{\sigma}_t, g_t$)	−0.513	−0.404	−0.727	-	-	-
Asset prices						
$\overline{er_t}$	4.611	4.611	4.610	-	-	-
Std (er_t)	0.001	0.001	0.001	-	-	-
corr (er_t, g_t)	0.513	0.405	0.728	-	-	-
$\overline{r_t^f}$	-	-	-	3.335	3.335	3.333
Std(r_t^f)	-	-	-	0.003	0.003	0.005
corr(r_t^f, g_t)	-	-	-	0.513	0.405	0.730
$\overline{rp_t}$	1.276	1.276	1.277	1.074	1.033	1.447
Std(rp_t)	0.001	0.001	0.001	0.342	0.321	0.448
corr(rp_t, g_t)	−0.513	−0.405	−0.727	−0.513	−0.404	−0.728

Table 4 shows the moments obtained from 50,000 simulations of the model for 312 quarters under the assumption that the agent has an infinite sample of observations and that the memory-context is persistent with $\phi = 0.85$. 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.

cability of the general framework developed in Section 2. From a behavioral perspective, a persistent context better captures how humans actually process and remember information than an i.i.d. context (Howard and Kahana, 1999, 2002). From an asset pricing perspective, a persistent context helps to reproduce the persistence in both cash-flow expectations and the price-dividend ratio.

4.3 Peak-end memory

This section illustrates the flexibility of the general framework by analyzing the effects of a peak-end memory distortion on beliefs and asset prices. In contrast to the similarity-weighted recall mechanism thus far, the agent now disproportionately recalls both extreme past experiences and those similar to the current environment, in line with the peak-end rule of Kahneman (2000). The peak-end memory distortion, as modeled 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., 2023). 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, 2024).

Framework. Consider i.i.d. log-normal endowment growth

$$g_t = \mu + \sigma\, \epsilon_t, \qquad \epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}\left(0, 1\right). \tag{41}$$

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, which is distorted by a peak-end memory function

$$m^{\mathrm{PE}}\left(g_\tau, g_t\right) := \underbrace{\exp\left[-e^{-\frac{(g_\tau-\mu)^2}{2\,\sigma^2}}\right]}_{\text{Extreme experience bias}} \cdot \underbrace{\exp\left[-\frac{\left(g_\tau - g_t\right)^2}{2\,\kappa}\right]}_{\text{Similarity}}. \tag{42}$$

The first term gives more weight to extreme "peak" realizations, while the second term increases the recall of experiences similar to the "end."[31]

Subjective beliefs. The agent is more likely to recall extreme experiences and experiences that are similar to the end of the realized history H_t. The extreme experience bias does

[31] The exponential transformation of the normal pdf in the "extreme experience bias term" implies that experiences near μ are underweighted, while extreme realizations receive disproportionately high recall probabilities. A 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 F.2.1.

Figure 7: Posterior beliefs under peak-end memory

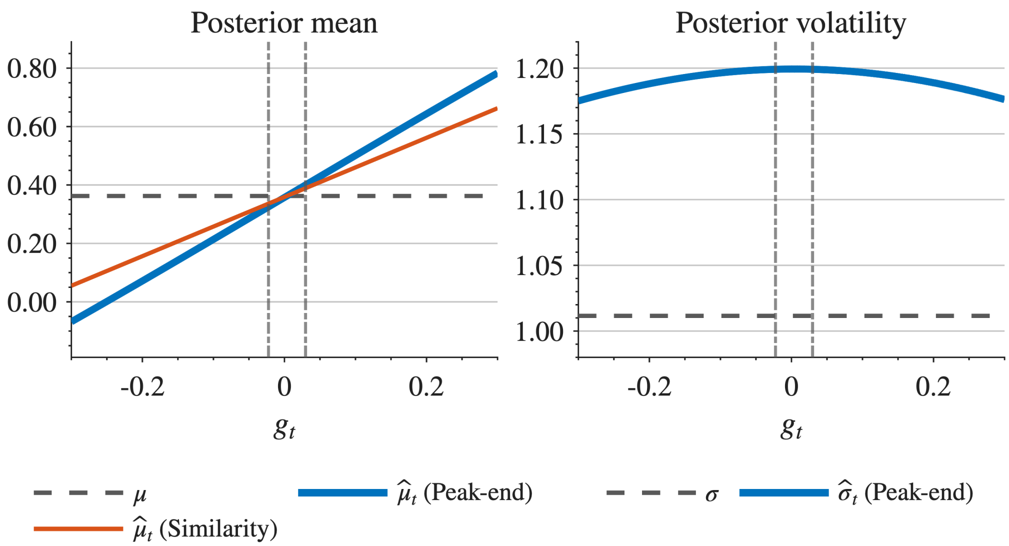

Figure 7 shows the posterior mean and variance of an agent with peak-end memory distortions as in Equation (42) for varying levels of contemporaneous endowment growth g_t. The parameters are $\mu = 0.35\%$, $\sigma = 1.03\%$, and $\kappa = 0.01$.

not systematically affect the posterior mean of the agent since both tails are overweighted symmetrically, but leads to an increased posterior variance. Appendix F.2.3 shows that a pure extreme experience bias yields $\hat{\mu}_t = \mu$ and $\hat{\sigma}_t^2 \approx 1.41 \cdot \sigma^2 > \sigma^2$.

When combined with the similarity term, the agent's posterior mean becomes procyclical, while the posterior variance is a concave function of g_t, peaking around the unconditional mean μ. At $g_t = \mu$, the memory function purely overweights extreme experiences, yielding $\hat{\sigma}_t^2 \approx 1.41 \cdot \sigma^2$. As g_t deviates from μ, similarity-weighting shifts recalled experiences toward one tail, reducing the posterior variance. Figure 7 illustrates these dynamics.

Asset pricing implications. The cumulant-generating function of log endowment growth under the agent's time-t belief is given by

$$\mathcal{K}_t^{PE}(k) = \log \tilde{\mathbb{E}}_t \left(e^{k\, g_{t+1}} \right) = k\, \hat{\mu}_t + \frac{1}{2} k^2\, \hat{\sigma}_t^2, \tag{43}$$

and I insert numerical estimates of $\hat{\mu}_t$ and $\hat{\sigma}_t^2$ to obtain subjective asset prices. I simulate the model $50,000$ times for 312 quarters and report average moments in Table 5. The parameters of the log endowment growth process are calibrated to the same mean and volatility as in

Table 5: Asset prices under the peak-end memory distortion

Symbol	Mean	Std.	Corr. with g_t
Endowment growth and subjective beliefs			
g_t	1.449	2.021	-
$\hat{\mu}_t$	1.450	0.015	1.000
$\hat{\sigma}_t$	2.397	< 0.001	< 0.001
Subjective asset prices			
er_t	4.947	0.009	1.000
rp_t	1.723	< 0.001	< 0.001
Objective asset prices			
r_t^f	3.224	0.017	1.000
rp_t	1.541	2.209	-1.000

Table 5 reports the model moments obtained from $50,000$ simulations of the model for 312 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 previous sections, with $\mu = 1.45\%$ and $\sigma = 2.02\%$. All other parameters are as in Table 1, and Table 5 summarizes the simulation results.

The agent's posterior mean is an unbiased estimate of the true mean ($\overline{\hat{\mu}}_t = 1.450\%$). Time-variation in the agent's beliefs is entirely due to the similarity component of the memory function in Equation (42), such that the agent's posterior mean is perfectly correlated with this period's endowment growth g_t, leading to procyclical beliefs. Table C.5 shows that a Coibion and Gorodnichenko (2015)-regression indicates overreaction to new information. The higher likelihood of recalling extreme experiences yields a higher posterior ($\overline{\hat{\sigma}}_t = 2.397\%$) than fundamental volatility ($\sigma = 2.021\%$), which is very stable and acyclical.

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 perceives the economy as fundamentally very risky, leading to a higher subjective risk premium than in Section 3. Moreover, the subjective risk premium is almost constant and not predictable using aggregate valuation ratios because of the acyclical posterior volatility (see

Table C.5).

The peak-end memory distortion analyzed here foremost demonstrates the generality of the framework developed in Section 3, which can accommodate empirically and psychologically motivated recall biases in a tractable asset pricing model. The extension links the propensity to recall emotionally salient experiences to the subjectively perceived riskiness of the economy and provides additional insights into the equilibrium asset pricing implications of experience effects that are complementary to cohort effects (Malmendier et al., 2020) or constant-gain learning (Nagel and Xu, 2022).

4.4 Discussion

This paper demonstrates how selective memory can be incorporated into in a standard general equilibrium asset-pricing framework to generate belief and price dynamics consistent with empirical evidence. For most of the analysis, I focus on similarity-weighted memory due to the growing empirical evidence that similarity shapes both individual and aggregate economic decision-making (Enke et al., 2024; Charles, 2025; Jiang et al., 2025). Beyond reproducing known regularities, the model generates new testable predictions.

First, the agent in the model can recall even temporally distant experiences if they are similar to the current context, implying that distant past episodes remain influential for belief formation and asset prices. The long-lasting effects of similar past episodes is consistent with the evidence in Charles and Sui (2025) and Chen et al. (2025). Both papers show that similarity is central in shaping aggregate beliefs and in predicting aggregate returns, which also motivates the focus on a representative-agent model, even though most micro-level evidence on similarity is from individual decisions.

Second, the model predicts a U-shaped pattern of subjective uncertainty and subjective risk premia—both are higher in extremely good as well as bad periods than in average times. I present preliminary empirical support using the data of Nagel and Xu (2023) in Tables E.2 and E.3. The model also has further predictions regarding higher moments of beliefs (Figure

B.1), which could, for example, be tested using survey measures of crash probabilities.

Finally, the model can be applied to analyze the general equilibrium asset pricing implications of further selective memory biases, such as a positive memory bias (Adler and Pansky, 2020) or a confirmatory memory bias (Esponda et al., 2024; Gödker et al., 2024), and extended to incorporate active learning, where agents' actions influence the data they observe and therewith their subsequent beliefs. An extension to active learning could help to link the model to evidence on return extrapolation (Barberis et al., 2018), where investors' expectations are shaped directly by realized returns.

5 Conclusion

This paper explores the implications of selective memory for asset prices and shows 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, the 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 current 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.

The model complements recent evidence from finance (Nagel and Xu, 2023; Bordalo et al., 2024; Charles and Sui, 2025; Jiang et al., 2025), experimental economics (Enke et al., 2024; Conlon, 2025), and cognitive psychology (Kahana, 2012) by offering a unified theoretical account of how selective—and especially similarity-weighted—memory distorts aggregate beliefs and prices. Without appealing to time-varying risk aversion, long-run risk, or disaster risk, the framework provides a psychologically grounded explanation for empirically observed belief and return dynamics in financial markets.

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.

Adler, O., & Pansky, A. (2020). A "rosy view" of the past: Positive memory biases. In *Cognitive biases in health and psychiatric disorders* (pp. 139–171). Elsevier.

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.

Andre, P., Schirmer, P., & Wohlfart, J. (2025). Mental models of the stock market. *Working paper.*

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

Ba, C., Bohren, J. A., & Imas, A. (2024). Over-and underreaction to information. *Working paper available at SSRN.*

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* (pp. 79–175). Amsterdam: Elsevier.

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. (2025). Expectations and learning from prices. *Review of Economic Studies*, *92*(3), 1341–1374.

Bastianello, F., & Imas, A. (2025). Biases in Belief Updating Within and Across Domains. *Working paper available at SSRN*.

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. (2023). 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. (2024). Belief overreaction and stock market puzzles. *Journal of Political Economy*, *132*(5).

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 SSRN.*

Charles, C. (2025). Memory Moves Markets. *Review of Financial Studies*, *38*(6), 1641–1686.

Charles, C., & Sui, P. (2025). Marketwide Memory. *Working paper available at SSRN.*

Chen, A., Hoberg, G., & Zhang, M. B. (2025). Haven't We Seen This Before? Return Predictions from 200 Years of News. *Working paper available at SSRN.*

Chen, Z., & Huang, J. (2025). Memory and beliefs in financial markets: A machine learning approach. *Working paper available at SSRN.*

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.

Conlon, J. J. (2025). Memory rehearsal and belief biases. *Working paper*.

Conlon, J. J., & Kwon, S. Y. (2025). Beliefs from Cues. *Working paper*.

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. (2024). Associative Memory, Beliefs and Market Interactions. *Journal of Financial Economics*, *157*.

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.

Esponda, I., Vespa, E., & Yuksel, S. (2024). Mental models and learning: The case of base-rate neglect. *American Economic Review*, *114*(3), 752–782.

Fudenberg, D., Lanzani, G., & Strack, P. (2024). Selective Memory Equilibrium. *Journal of Political Economy*, *132*(12).

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

Gödker, K., Jiao, P., & Smeets, P. (2025). Investor memory. *Review of Financial Studies*, *38*(6), 1595–1640.

Gödker, K., Odean, T., & Smeets, P. (2024). Disposed to be overconfident. *Working paper available at SSRN*.

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 SSRN*.

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.

Gruber, J. (2013). A tax-based estimate of the elasticity of intertemporal substitution. *Quarterly Journal of Finance*, *3*(01).

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. (2023). Shattered housing. *Journal of Financial Economics*, *156*.

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

Howard, M. W., & Kahana, M. J. (1999). Contextual variability and serial position effects in free recall. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*(4), 923–941.

Howard, M. W., & Kahana, M. J. (2002). A distributed representation of temporal context. *Journal of Mathematical Psychology*, *46*(3), 269–299.

Jiang, Z., Liu, H., Peng, C., & Yan, H. (2025). Investor Memory and Biased Beliefs: Evidence from the Field. *Quarterly Journal of Economics*, *140*(4), 2749–2804.

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.

Jordà, Ò., Knoll, K., Kuvshinov, D., Schularick, M., & Taylor, A. M. (2019). The Rate of Return on Everything, 1870–2015. *Quarterly Journal of Economics*, *134*(3), 1225–1298.

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. (2024). Laws of Human Memory. In M. J. Kahana & A. D. Wagner (Eds.), *The Oxford Handbook of Human Memory*. Oxford University Press.

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*, 354–381.

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.

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. (2024). Memory of Past Experiences and Economic Decisions. In M. J. Kahana & A. D. Wagner (Eds.), *The Oxford Handbook of Human Memory*. Oxford University Press.

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. (2024). Model Complexity, Expectations, and Asset Prices. *Review of Economic Studies*, *91*(4), 2462–2507.

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*, *35*(5), 2190–2245.

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*. Society for Industrial and Applied Mathematics.

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

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

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive psychology*, *5*(2), 207–232.

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. (2024). A retrieved-context theory of financial decisions. *Quarterly Journal of Economics*, *139*(2), 1095–1147.

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

A Proofs

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

$$M = \sum_{s \in S} \psi(s) \int_{-\infty}^{\infty} m_{(g_s,t_t)}(g,s)\, q_s^*(g)\, dg. \tag{A.1}$$

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

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

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

$$\mu_{LM} = \tilde{\mu}_m, \tag{A.3}$$

$$\sigma^2_{LM} = \tilde{\sigma}^2_m. \tag{A.4}$$

The parameter space Θ is closed and convex, 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 $\tau_0 \to -\infty$

$$\begin{aligned}
\hat{\mu}_t =& \mathbb{E}\Big[\frac{1}{|H_t^R|}\sum_{\tau=\tau_0}^{t} g_\tau \mathbb{1}_{\{g_\tau \in H_t^R\}}\Big] \\
=& \mathbb{E}\Big[\frac{1}{|H_t^R|}\Big]\,\mathbb{E}\Big[\sum_{\tau=\tau_0}^{t} g_\tau \mathbb{1}_{\{g_\tau \in H_t^R\}}\Big] + \underbrace{\mathrm{Cov}\Big[\frac{1}{|H_t^R|}, \sum_{\tau=\tau_0}^{t} g_\tau \mathbb{1}_{\{g_\tau \in H_t^R\}}\Big]}_{0 \text{ for } \tau_0 \to -\infty} \\
=& \mathbb{E}\Big[\frac{1}{|H_t^R|}\Big]\sum_{\tau=\tau_0}^{t} \mu \cdot m_{(g_t,s_t)}(g_\tau, s_\tau) + \mathbb{E}\Big[\frac{1}{|H_t^R|}\Big]\sum_{\tau=\tau_0}^{t} \mathrm{Cov}\Big[g, \mathbb{1}_{\{g \in H_t^R\}}\Big] \\
=& \mu \cdot \mathbb{E}\Big[\frac{1}{\sum_{\tau=\tau_0}^{t} \mathbb{1}_{\{g_\tau \in H_t^R\}}}\Big] \cdot \sum_{\tau=\tau_0}^{t} \mathbb{E}\left(\mathbb{1}_{\{g_\tau \in H_t^R\}}\right) + \mathbb{E}\Big[\frac{1}{|H_t^R|}\Big]\sum_{\tau=\tau_0}^{t} \mathrm{Cov}\Big[g, \mathbb{1}_{\{g \in H_t^R\}}\Big] \\
=& \mu \cdot \mathbb{E}\Big[\frac{1}{\sum_{\tau=\tau_0}^{t} \mathbb{1}_{\{g_\tau \in H_t^R\}}}\Big] \cdot \mathbb{E}\left(\sum_{\tau=\tau_0}^{t} \mathbb{1}_{\{g_\tau \in H_t^R\}}\right) + \mathbb{E}\Big[\frac{1}{|H_t^R|}\Big]\sum_{\tau=\tau_0}^{t} \mathrm{Cov}\Big[g, \mathbb{1}_{\{g \in H_t^R\}}\Big] \\
=& \mu + \mathbb{E}\Big[\frac{t}{|H_t^R|}\Big] \cdot \mathrm{Cov}\Big[g, \mathbb{1}_{\{g \in H_t^R\}}\Big].
\end{aligned} \tag{A.5}$$

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, $\hat{\sigma}^2_t = \frac{1}{|H_t^R|}\sum_{\tau=1}^{t}(g_\tau \mathbb{1}_{\{g_\tau \in H_t^R\}} - \hat{\mu}_t)^2$, which almost surely equals it's expected value for $\tau_0 \to -\infty$

$$\begin{aligned}
\hat{\sigma}^2_t =& \mathbb{E}\Big[\frac{1}{|H_t^R|}\sum_{\tau=\tau_0}^{t} \mathbb{1}_{\{g_\tau \in H_t^R\}}(g_\tau - \hat{\mu}_t)^2\Big] \\
=& \mathbb{E}\Big[\frac{1}{|H_t^R|}\Big]\,\mathbb{E}\Big[\sum_{\tau=\tau_0}^{t} \mathbb{1}_{\{g_\tau \in H_t^R\}}\Big(g_\tau - \hat{\mu}_t\Big)^2\Big] \\
=& \mathbb{E}\Big[\frac{1}{|H_t^R|}\Big]\sum_{\tau=\tau_0}^{t} \mathbb{E}\Big[\mathbb{1}_{\{g_\tau \in H_t^R\}}\Big]\,\mathbb{E}\Big[(g_\tau - \hat{\mu}_t)^2\Big] + \mathrm{Cov}\Big(\mathbb{1}_{\{g \in H_t^R\}}, (g - \hat{\mu}_t)^2\Big).
\end{aligned} \tag{A.6}$$

We now solve for

$$\mathbb{E}\Big[(g_\tau - \hat{\mu}_t)^2\Big] = \underbrace{\mathbb{E}\Big[(g_\tau - \mu)^2\Big]}_{=\sigma^2} - 2\,\underbrace{\mathbb{E}\Big[(g_\tau - \mu)(\hat{\mu}_t - \mu)\Big]}_{(a)} + \underbrace{\mathbb{E}\Big[(\hat{\mu}_t - \mu)^2\Big]}_{(b)}, \tag{A.7}$$

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

$$\begin{aligned}(a)\ \ \mathbb{E}\Big[(g_\tau - \mu)(\hat{\mu}_t - \mu)\Big] &= \underbrace{\mathbb{E}\Big[(g_\tau - \mu)\Big]}_{=0}\,\mathbb{E}\Big[(\hat{\mu}_t - \mu)\Big] + \mathrm{Cov}\Big((g_\tau - \mu), (\hat{\mu}_t - \mu)\Big) \\ &= \mathrm{Cov}\Big(g_\tau, \frac{1}{|H_t^R|}\sum_{j=\tau_0}^{t} g_j \mathbb{1}_{\{g_j \in H_t^R\}}\Big) \\ &= \underbrace{\frac{1}{|H_t^R|}}_{=\frac{1}{\infty}\text{ for }\tau_0\to-\infty}\ \underbrace{\mathrm{Cov}\Big(g_\tau, g_\tau \mathbb{1}_{\{g_\tau \in H_t^R\}}\Big)}_{<\sigma^2<\infty} = 0. \end{aligned} \tag{A.8}$$

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

$$\begin{aligned}(b)\ \ \mathbb{E}\Big[(\hat{\mu}_t - \mu)^2\Big] = \mathbb{E}\Big[(\mathbb{E}(\hat{\mu}_t) - \mu)^2\Big] &= \mathbb{E}\left[\Big(\mathbb{E}\Big[\frac{t}{|H_t^R|}\Big] \cdot \mathrm{Cov}\Big[g, \mathbb{1}_{\{g \in H_t^R\}}\Big]\Big)^2\right] \\ &= \Big(\mathbb{E}\Big[\frac{t}{|H_t^R|}\Big] \cdot \mathrm{Cov}\Big[g, \mathbb{1}_{\{g \in H_t^R\}}\Big]\Big)^2 = (\hat{\mu}_t - \mu)^2\,. \end{aligned} \tag{A.9}$$

Overall, we thus have

$$\mathbb{E}\Big[(g_\tau - \hat{\mu}_t)^2\Big] = \sigma^2 + (\hat{\mu}_t - \mu)^2\,, \tag{A.10}$$

and inserting into Equation (A.6) yields the claim.

A.2 Proposition 2

I derive the posterior belief of an agent with similarity-weighted memory as given by Equation (17). The state $s \in \{1, 2\}$ follows an observable Markov chain, such that the agent can perform state-wise inference. We can write the probability density function (pdf) of the mixture distribution as

$$\pi_1 \cdot q(g|s_\tau = 1) + (1 - \pi_1) \cdot q(g|s_\tau = 2), \tag{A.11}$$

where $q(g|s_\tau = s)$ denotes the pdf of a normal distribution with mean μ_s and standard deviation σ_s. The memory-weighted pdf is then found as

$$\pi_1 \cdot q(g|s_\tau = 1) \cdot m^{\text{sim}}_{(g_t,s_t)}(g, s) + (1 - \pi_1) \cdot q(g|s_\tau = 2) \cdot m^{\text{sim}}_{(g_t,s_t)}(g, s), \tag{A.12}$$

where $m^{\text{sim}}_{(g_t,s_t)}(g, s)$ is given in Equation (17). We first solve for the memory-weighted state-wise pdfs, $q(g|s_\tau = s) \cdot m^{\text{sim}}_{(g_t,s_t)}(g, s)$, and then rescale the mixture pdf to find the subjective probability of each state. The process is psychologically reminiscent of the two-state process in Conlon and Kwon (2025).

It is

$$\begin{aligned} q(g|s_\tau = s) \cdot m^{\text{sim}}_{(g_t,s_t)}(g, s) &= \frac{1}{\sqrt{2\,\pi\,\sigma_s^2}} \exp\left[-\frac{(g - g_t)^2}{2\,\kappa}\right] \exp\left[-\frac{(g - \mu_s)^2}{2\,\sigma_s^2}\right] \\ &= \frac{1}{\sqrt{2\,\pi\,\sigma_s^2}} \exp\left[-\frac{g^2 - 2\,\frac{\kappa\,\mu_s + \sigma_s^2\,g_t}{\kappa+\sigma_s^2}\,g + \frac{\kappa\,\mu_s^2 + \sigma_s^2\,g_t^2}{\kappa+\sigma_s^2}}{2\frac{\kappa\,\sigma_s^2}{\kappa+\sigma_s^2}}\right]. \end{aligned} \tag{A.13}$$

The exponential term is Gaussian with

$$\hat{\mu}_{s,t} = \frac{\kappa\,\mu_s + \sigma_s^2\,g_t}{\kappa + \sigma_s^2} = \frac{\kappa}{\kappa + \sigma_s^2}\,\mu_s + \frac{\sigma_s^2}{\kappa + \sigma_s^2}\,g_t = (1 - \alpha_s)\mu_s + \alpha_s\,g_t, \text{ and} \tag{A.14}$$

$$\hat{\sigma}^2_{s,t} = \frac{\kappa\,\sigma_s^2}{\kappa + \sigma_s^2} = (1 - \alpha_s)\,\sigma_s^2, \tag{A.15}$$

with $\alpha_s := \frac{\sigma_s^2}{\kappa+\sigma_s^2}$.

Next, we need to rescale the expression in Equation (A.12) to be a valid pdf. It is

$$\begin{aligned}
&\int_{-\infty}^{\infty} \pi_1 \cdot q(g|s_\tau = 1) \cdot m^{\text{sim}}_{(g_t,s_t)}(g,s) + (1-\pi_1) \cdot q(g|s_\tau = 2) \cdot m^{\text{sim}}_{(g_t,s_t)}(g,s)\; dg \\
=&\pi_1 \int_{-\infty}^{\infty} q(g|s_\tau = 1) \cdot m^{\text{sim}}_{(g_t,s_t)}(g,s) + (1-\pi_1) \int_{-\infty}^{\infty} q(g|s_\tau = 2) \cdot m^{\text{sim}}_{(g_t,s_t)}(g,s)\; dg. \qquad \text{(A.16)}
\end{aligned}$$

We can use the following known formula to obtain the area under the memory-weighted statewise pdfs, that is, the mass of recalled observations from each state:

$$\int_{-\infty}^{\infty} e^{-(ax^2+bx+c)}\, dx = \sqrt{\frac{\pi}{a}} e^{\frac{b^2}{4a}-c} \quad (a > 0), \qquad \text{(A.17)}$$

which yields

$$\begin{aligned}
&\frac{1}{\sqrt{2\,\pi\,\sigma_s^2}} \int_{-\infty}^{\infty} \exp\left[-\frac{g^2\,(\kappa+\sigma_s^2) - 2\,g\,(\mu_s\,\kappa + g_t\,\sigma_s^2) + \kappa\,\mu_s^2 + g_t^2\,\sigma_s^2}{2\,\sigma_s^2\,\kappa}\right] dg \\
=&\sqrt{(1-\alpha_s)}\, e^{-\frac{(\mu_s - g_t)^2}{2\,(\kappa+\sigma_s^2)}} := \mathcal{M}_{s,t}. \qquad \text{(A.18)}
\end{aligned}$$

The total mass of recalled observations is thus

$$\pi_1\,\mathcal{M}_{1,t} + (1-\pi_1)\,\mathcal{M}_{2,t}, \qquad \text{(A.19)}$$

and rescaling Equation (A.12) to integrate to one gives

$$\hat{\pi}_{s,t} = \frac{\pi_s\,\mathcal{M}_{s,t}}{\pi_1\,\mathcal{M}_{1,t} + (1-\pi_1)\,\mathcal{M}_{2,t}}, \text{ for } s \in \{1,2\}. \qquad \text{(A.20)}$$

The prior support of the agent contains all mixtures of normal distributions. Therefore, the unique maximizer of the memory-weighted likelihood given in Equation (2) is the mixture derived above. An alternative, but longer, proof that explicitly uses Proposition 1 is available.

A.3 Proposition 3

It is

$$\mathbb{E}\left[\hat{\mu}_{s,t+1}|s_{t+1}=s\right]=\mathbb{E}\left[\left(1-\alpha_{s_{t+1}}\right)\mu_{s_{t+1}}+\alpha_{s_{t+1}}\,g_{t+1}|s_{t+1}=s\right]$$
$$=\left(1-\alpha_{s_{t+1}}\right)\mu_{s}+\alpha_{s_{t+1}}\,\mathbb{E}[g_{t+1}|s_{t+1}=s]=\mu_{s}, \tag{A.21}$$

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

$$\mathbb{E}\left[\hat{\mu}_{s',t+1}|s_{t+1}=s\right]=\mathbb{E}\left[\left(1-\alpha_{s'}\right)\mu_{s_-}+\alpha_{s'}\,g_{t+1}|s_{t+1}=s\right]$$
$$=\left(1-\alpha_{s'}\right)\mu_{s'}+\alpha_{s'}\,\mathbb{E}[g_{t+1}|s_{t+1}=s]=\mu_{s'}+\alpha_{s'}\,\left(\mu_{s}-\mu_{s'}\right). \tag{A.22}$$

Combining both expressions, it is

$$\mathbb{E}\left[\hat{\mu}_{s,t+1}\right]=\pi_s\,\mathbb{E}\left[\hat{\mu}_{s,t+1}|s_{t+1}=s\right]+\left(1-\pi_s\right)\mathbb{E}\left[\hat{\mu}_{s,t+1}|s_{t+1}=s'\right], \tag{A.23}$$

and inserting yields the claim in Proposition 3.

A.4 Proposition 4

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

$$\mathbb{E}\left[\mathrm{Var}_t\left(g_{t+1}\right)\right]=\pi_1\,\sigma_1^2+\pi_2\,\sigma_2^2+\pi_1\,\pi_2\,\mathbb{E}\left[\left(\hat{\mu}_{1,t}-\hat{\mu}_{2,t}\right)^2\right], \tag{A.24}$$

and we have

$$\mathbb{E}\left(\hat{\mu}_{1,t}-\hat{\mu}_{2,t}\right)=\left(\mu_1-\mu_2\right)\left[1-\left(\alpha_1\,\pi_2+\alpha_2\,\pi_1\right)\right] \tag{A.25}$$

$$\mathrm{Var}\left(\hat{\mu}_{1,t}-\hat{\mu}_{2,t}\right)=\mathrm{Var}\left[(1-\alpha_1)\mu_1-(1-\alpha_2)\mu_2+(\alpha_1-\alpha_2)g_t\right]=(\alpha_1-\alpha_2)^2\ \mathrm{Var}(g_t). \tag{A.26}$$

By the i.i.d. assumption and the definition of variance, we can write

$$
\begin{aligned}
\mathbb{E}\left[\operatorname{Var}_t\left(g_{t+1}\right)\right] =& \pi_1\,\sigma_1^2+\pi_2\,\sigma_2^2+\pi_1\,\pi_2\,\left[\left(\alpha_1-\alpha_2\right)^2\,\operatorname{Var}(g_{t+1})+\right. \\
& \left.\left(\mu_1-\mu_2\right)^2\left[1-\left(\alpha_1\,\pi_2+\alpha_2\,\pi_1\right)\right]^2\right] \qquad \text{(A.27)} \\
=& \left(\pi_1\,\sigma_1^2+\pi_2\,\sigma_2^2\right)\left[1+\pi_1\,\pi_2\,\left(\alpha_1-\alpha_2\right)^2\right] \\
& +\left(\mu_1-\mu_2\right)^2\,\pi_1\,\pi_2\,\left[\pi_1\,\pi_2\left(\alpha_1-\alpha_2\right)^2+\left[1-\left(\alpha_1\,\pi_2+\alpha_2\,\pi_1\right)\right]^2\right]. \qquad \text{(A.28)}
\end{aligned}
$$

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

$$
\begin{aligned}
& \mathbb{E}\left[\operatorname{Var}_t\left(g_{t+1}\right)\right]\geq \operatorname{Var}\left(g_{t+1}\right) \\
\iff\ & \mathbb{E}\left[\left(\hat{\mu}_{1,t}-\hat{\mu}_{2,t}\right)^2\right]\geq\left(\mu_1-\mu_2\right)^2 \qquad \text{(A.29)} \\
\iff\ & \left(\alpha_1-\alpha_2\right)^2\,\operatorname{Var}(g_{t+1})\geq\left(\mu_1-\mu_2\right)^2\left(1-\left[1-\left(\alpha_1\,\pi_2+\alpha_2\,\pi_1\right)\right]^2\right) \qquad \text{(A.30)} \\
\iff\ & \frac{\left(\alpha_1-\alpha_2\right)^2\,\left(\pi_1\,\sigma_1^2+\pi_2\,\sigma_2^2\right)}{2\,\left(\pi_2\,\alpha_1+\pi_1\,\alpha_2\right)-\left(\pi_2\,\alpha_1^2+\pi_1\,\alpha_2^2\right)}\geq\left(\mu_1-\mu_2\right)^2, \qquad \text{(A.31)}
\end{aligned}
$$

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

$$
\begin{aligned}
\frac{\mathbb{E}\left[\operatorname{Var}_t\left(g_{t+1}\right)\right]}{\operatorname{Var}(g_{t+1})} =& \frac{\operatorname{Var}(g_{t+1})+\pi_1\,\pi_2\left(\alpha_1-\alpha_2\right)^2\,\operatorname{Var}(g_{t+1})}{\operatorname{Var}(g_{t+1})}+ \\
& \frac{\pi_1\,\pi_2\,\left(\mu_1-\mu_2\right)^2\,\left[-2\,\left(\alpha_1\,\pi_2+\alpha_2\,\pi_1\right)+\left(\alpha_1\,\pi_2+\alpha_2\,\pi_1\right)^2\right]}{\operatorname{Var}(g_{t+1})} \qquad \text{(A.32)} \\
=& 1+\underbrace{\pi_1\,\pi_2}_{\leq 0.25}\,\underbrace{\left(\alpha_1-\alpha_2\right)^2}_{\leq 1}+ \\
& \underbrace{\left(\alpha_1\,\pi_2+\alpha_2\,\pi_1-2\right)}_{\leq -1}\,\underbrace{\frac{\pi_1\,\pi_2\,\left(\mu_1-\mu_2\right)^2\,\left(\alpha_1\,\pi_2+\alpha_2\,\pi_1\right)}{\operatorname{Var}(g_{t+1})}}_{>0} \qquad \text{(A.33)} \\
\leq& 1.25. \qquad \text{(A.34)}
\end{aligned}
$$

A.5 Proposition 5

To find the coefficients of the Coibion and Gorodnichenko (2015)-regression, note that the forecast revision is

$$\tilde{\mathbb{E}}_{\mathrm{t}}(g_{t+h}) - \tilde{\mathbb{E}}_{t-1}(g_{t+h}) = (\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\;(g_t - g_{t-1})\,, \tag{A.35}$$

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

$$\beta_{CG} = \frac{\mathrm{Cov}\left[g_{t+h} - \tilde{\mathbb{E}}_{\mathrm{t}}(g_{t+h}), \tilde{\mathbb{E}}_{\mathrm{t}}(g_{t+h}) - \tilde{\mathbb{E}}_{t-1}(g_{t+h})\right]}{\mathrm{Var}\left[\tilde{\mathbb{E}}_{\mathrm{t}}(g_{t+h}) - \tilde{\mathbb{E}}_{t-1}(g_{t+h})\right]} \tag{A.36}$$

$$= \frac{(\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\;\mathbb{E}\left[(g_{t+h} - (\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\;g_t - \mu + (\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\;\mu)\cdot(g_t - g_{t-1})\right]}{2\;(\pi_1\,\alpha_1 + \pi_2\,\alpha_2)^2\,\mathrm{Var}(g)} \tag{A.37}$$

$$= \frac{\mathbb{E}\left[g_{t+h}\,g_t - g_{t+h}\,g_{t-1} - (\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\;g_t^2 + (\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\;g_t\,g_{t-1}\right]}{2\;(\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\,\mathrm{Var}(g)} \tag{A.38}$$

$$= \frac{-\,(\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\;\left[\mathbb{E}(g_t^2) - \mu^2\right]}{2\;(\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\,\mathrm{Var}(g)} = -\frac{1}{2}. \tag{A.39}$$

Finally, using Equation (26) and noting that $\mathbb{E}\left(\tilde{\mathbb{E}}_{\mathrm{t}}(g_{t+h}) - \tilde{\mathbb{E}}_{t-1}(g_{t+h})\right) = 0$, it is $a_{CG} = -\pi_1\,\pi_2\,(\mu_2 - \mu_1)\;(\alpha_1 - \alpha_2)$.

In addition, we can also solve for the coefficients of a Mincer and Zarnowitz (1969)-regression, given by

$$g_{t+h} = a_{MZ} + \beta_{MZ}\;\tilde{\mathbb{E}}_{\mathrm{t}}\,(g_{t+h}) + u_{t+h}. \tag{A.40}$$

Write

$$\tilde{\mathbb{E}}_{\mathrm{t}}(g_{t+h}) = \pi_1\;((1-\alpha_1)\,\mu_1 + \alpha_1\,g_t) + \pi_2\;((1-\alpha_2)\,\mu_2 + \alpha_2\,g_t) = \tilde{\mu} + (\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\;g_t, \tag{A.41}$$

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

$$\beta_{MZ} = \frac{\text{Cov}(g_{t+h}, \tilde{\mathbb{E}}_t(g_{t+h}))}{\text{Var}\left(\tilde{\mathbb{E}}_t(g_{t+h})\right)} = \frac{\text{Cov}(g_{t+h}, (\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\;g_t)}{\text{Var}\left((\pi_1\,\alpha_1 + \pi_2\,\alpha_2)\;g_t\right)} = 0, \tag{A.42}$$

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

A.6 Proposition 6

The results follow from inserting the cumulant-generating function given in Equation (21) into Result 1 in Martin (2013), as shown in Section 2.4. Under Epstein and Zin (1989)-preferences, it is

$$\begin{aligned} dp_t =& -\log(\beta) - \log\left(\pi_1\,e^{(\lambda-\gamma)\,\hat{\mu}_{1,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_1^2} + \pi_2\,e^{(\lambda-\gamma)\,\hat{\mu}_{2,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_2^2}\right) \\ &+ \left(1-\frac{1}{\eta}\right)\,\log\left(\pi_1\,e^{(1-\gamma)\,\hat{\mu}_{1,t}+\frac{1}{2}\,(1-\gamma)^2\,\sigma_1^2} + \pi_2\,e^{(1-\gamma)\,\hat{\mu}_{2,t}+\frac{1}{2}\,(1-\gamma)^2\,\sigma_2^2}\right), \end{aligned} \tag{A.43}$$

$$\begin{aligned} r_t^f =& -\log(\beta) - \log\left(\pi_1\,e^{-\gamma\,\hat{\mu}_{1,t}+\frac{1}{2}\gamma^2\,\sigma_1^2} + \pi_2\,e^{-\gamma\,\hat{\mu}_{2,t}+\frac{1}{2}\,\gamma^2\,\sigma_2^2}\right) \\ &+ \left(1-\frac{1}{\eta}\right)\,\log\left(\pi_1\,e^{(1-\gamma)\,\hat{\mu}_{1,t}+\frac{1}{2}\,(1-\gamma)^2\,\sigma_1^2} + \pi_2\,e^{(1-\gamma)\,\hat{\mu}_{2,t}+\frac{1}{2}\,(1-\gamma)^2\,\sigma_2^2}\right), \end{aligned} \tag{A.44}$$

$$er_t = dp_t + \log\left(\pi_1\,e^{\lambda\,\hat{\mu}_{1,t}+\frac{1}{2}\,\lambda^2\,\sigma_1^2} + \pi_2\,e^{\lambda\,\hat{\mu}_{2,t}+\frac{1}{2}\,\lambda^2\,\sigma_2^2}\right), \tag{A.45}$$

$$\begin{aligned} rp_t =& \log\left(\pi_1\,e^{-\gamma\,\hat{\mu}_{1,t}+\frac{1}{2}\gamma^2\,\sigma_1^2} + \pi_2\,e^{-\gamma\,\hat{\mu}_{2,t}+\frac{1}{2}\,\gamma^2\,\sigma_2^2}\right) + \log\left(\pi_1\,e^{\lambda\,\hat{\mu}_{1,t}+\frac{1}{2}\,\lambda^2\,\sigma_1^2} + \pi_2\,e^{\lambda\,\hat{\mu}_{2,t}+\frac{1}{2}\,\lambda^2\,\sigma_2^2}\right) \\ &- \log\left(\pi_1\,e^{(\lambda-\gamma)\,\hat{\mu}_{1,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_1^2} + \pi_2\,e^{(\lambda-\gamma)\,\hat{\mu}_{2,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_2^2}\right). \end{aligned} \tag{A.46}$$

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

B Additional figures

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

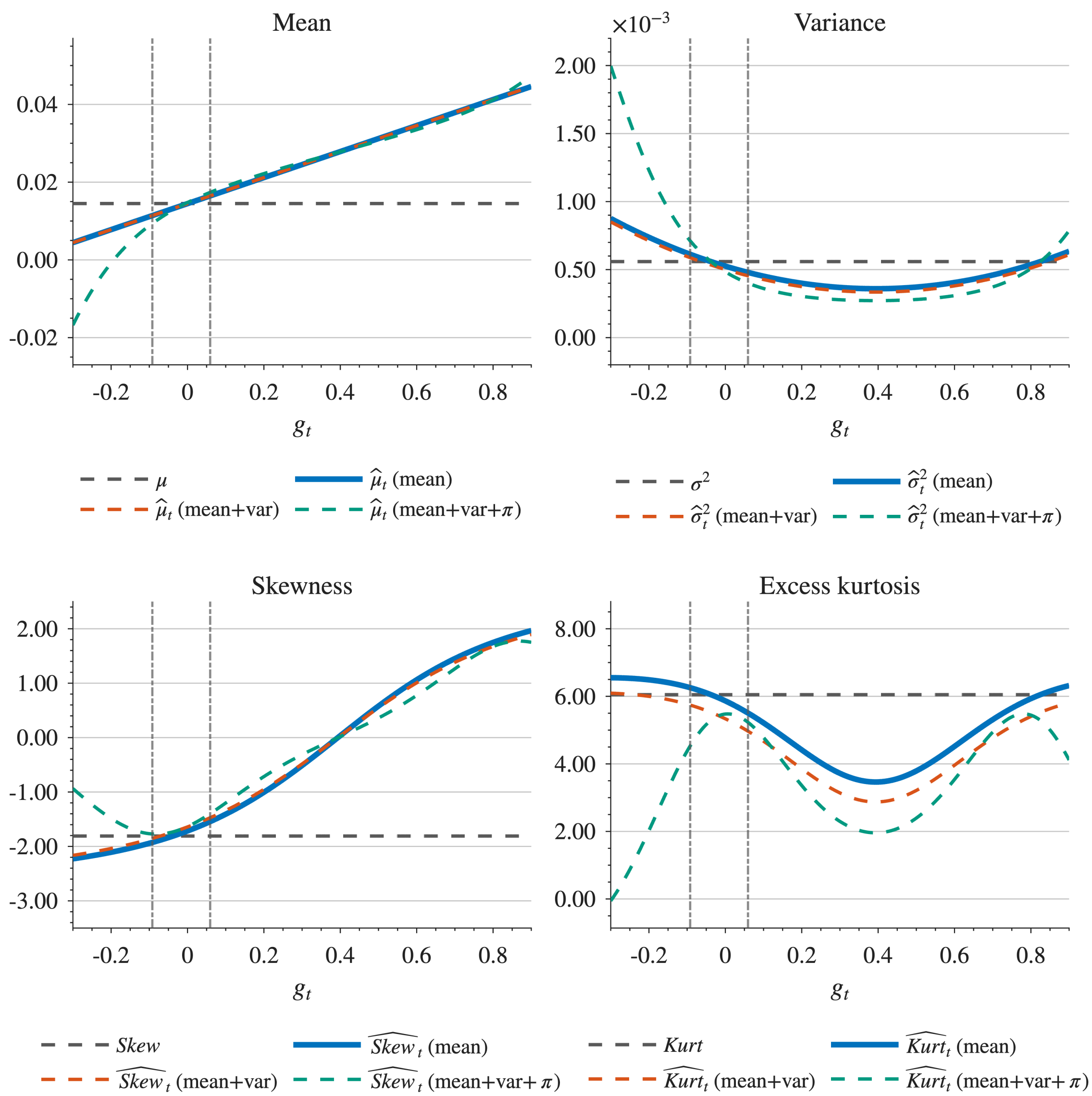

Figure B.1 shows the first four moments of endowment growth under the agent's subjective beliefs under similarity-weighted memory. In each panel, the dashed grey line shows the true underlying values, the solid blue line shows the agent's subjective beliefs when only learning about the state-dependent means, and the dashed orange line shows the agent's subjective beliefs when learning about the state-dependent means and variances, and the dashed teal line shows the agent's subjective beliefs when learning about the state-dependent means, variances and the transition probability. The vertical dashed lines show the 0.5% and 99.5% quantiles of the log endowment growth process. Log endowment growth is distributed as in Equation (16), and the similarity-weighted memory function is as in Equation (17). The parameters are as in Table 1, except for $\kappa = 0.01$.

C Additional tables

Table C.1: Predictability and Coibion and Gorodnichenko (2015)-regressions

	Baseline			Parameter uncertainty			Persistent context		
	rp_t	rp_t^O	$\hat{b}_{CG}$	rp_t	rp_t^O	$\hat{b}_{CG}$	rp_t	rp_t^O	$\hat{b}_{CG}$
dp_t	0.001	0.603		0.803	2.110		0.000	0.167	
$(\tilde{\mathbb{E}}_t - \tilde{\mathbb{E}}_{t-1})\, g_{t+1}$			-0.252			-0.470			-0.549

Table C.1 reports the median estimates from regressions for 50,000 simulations of the model for 312 quarters plus a 120 quarters burn-in period in the parameter-uncertainty case. The first row shows the median coefficients when regressing subjectively expected and objectively realized risk premia on the log dividend-price ratio, as in Nagel and Xu (2023). The log dividend-price ratio is rescaled to unit standard deviation. The second row shows the median estimates from Coibion and Gorodnichenko (2015)-regressions of the forecast error on the forecast revision. The left (middle, right) panel shows the results obtained without parameter uncertainty (with parameter uncertainty, with a persistent context).

Table C.2: Average asset pricing outcomes per quintile of log endowment growth

	Baseline				Parameter uncertainty			
Quintile	$\overline{g}_t$	$\overline{\hat{\mu}}_t$	$\overline{rp}_t$	$\overline{rp}_t^O$	$\overline{g}_t$	$\overline{\hat{\mu}}_t$	$\overline{rp}_t$	$\overline{rp}_t^O$
1	-4.258	1.432	1.283	4.363	-4.260	1.419	13.433	23.520
2	-0.047	1.450	1.278	1.936	-0.045	1.455	14.672	4.664
3	1.733	1.458	1.276	0.911	1.734	1.470	15.540	-0.156
4	3.453	1.466	1.274	-0.079	3.455	1.484	16.543	-4.067
5	6.373	1.479	1.272	-1.750	6.375	1.509	18.410	-8.375

Table C.2 reports average endowment 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 quintile of endowment growth. The moments are obtained from 50,000 simulations of the model for 312 quarters. For the parameter uncertainty simulations, I use 120 quarters as burn-in period. All averages are annualized by multiplying quarterly means by four.

Table C.3: Persistence of beliefs and asset prices

	Parameter uncertainty			Persistent context		
Lags	$\hat{\mu}_t$	$\log(P_t/D_t)$	rp_t^O	$\hat{\mu}_t$	$\log(P_t/D_t)$	rp_t^O
1	0.794	0.892	-0.386	0.837	0.799	-0.059
2	0.774	0.869	-0.001	0.699	0.641	-0.051
3	0.754	0.849	0.000	0.583	0.520	-0.043
5	0.715	0.823	0.000	0.404	0.358	-0.031
7	0.678	0.797	0.000	0.276	0.245	-0.022

Table C.3 reports the autocorrelations of the posterior mean, log price-dividend ratios and log excess returns over various lags. The values are calculated using $50,000$ simulations of 312 quarterly observations from the model. I compound the price-dividend ratio by summing the last four quarterly dividends for comparability with empirical results, but excess returns and posterior means are quarterly.

Table C.4: Robustness results with baseline parameters and parameter uncertainty

	Subjective			Objective		
Symbol	Total	Normal	Recession	Total	Normal	Recession
Endowment growth						
Mean	1.467	1.918	-2.615	1.452	1.922	-2.801
	(1.465)	(1.917)	(-2.634)			
Std ($\hat{\mu}_t$)	0.055	0.045	0.355	-	-	-
	(0.054)	(0.044)	(0.342)			
Volatility	2.025	2.024	2.034	2.018	-	-
	(2.025)	(2.024)	2.034			
corr($\hat{\sigma}_t, g_t$)	-0.082	-0.019	-0.270			
	(-0.082)	(-0.019)	-0.270	-	-	-
Asset prices						
$\overline{er_t}$	9.768	9.782	9.646	-	-	-
	(9.720)	(9.734)	(9.598)			
Std (er_t)	1.246	1.244	1.231	-	-	-
	(1.180)	(1.181)	(1.160)			
corr (er_t, g_t)	0.052	0.044	0.085	-	-	-
	(0.055)	(0.475)	(0.095)			
$\overline{r_t^f}$	-	-	-	5.090	5.094	5.050
Std(r_t^f)	-	-	-	0.134	0.130	0.151
corr(r_t^f, g_t)	-	-	-	0.137	0.130	0.274
$\overline{rp_t}$	4.678	4.687	4.596	1.044	-1.372	22.907
	(4.631)	(4.640)	(4.548)			
Std(rp_t)	1.275	1.273	1.264	22.499	21.911	24.243
	(1.209)	(1.210)	(1.193)			
corr(rp_t, g_t)	0.041	0.036	0.064	-0.260	-0.180	-0.410
	(0.043)	(0.039)	(0.068)			

Table C.4 reports the moments obtained from 50,000 simulations of the model for 312 quarters plus a 120 quarter burn-in period when the agent learns from a finite sample. Parameters are as in Table 1, except for $\beta = 0.99$ and $\psi = 1.1$ to ensure equilibrium existence for a wide range of subjective beliefs. Values in parentheses report results when subjective beliefs are truncated to the equilibrium-consistent range. 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.

Table C.5: Predictability and Coibion and Gorodnichenko (2015)-regressions under peak-end rule

	RP_{Subj}	RP_{Obj}	$\hat{b}_{CG}$
dp_t	-0.00002	1.104	
$\left(\tilde{\mathbb{E}}_t - \tilde{\mathbb{E}}_{t-1}\right) g_{t+1}$			-0.411

Table C.5 reports the median estimates from regressions for 50,000 simulations of the model for 312 quarters. The first row shows the median estimates when regressing subjectively expected and objectively realized 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 median 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 (42).

D Asset pricing model

In this appendix, I derive the asset-pricing results in Section 2.4 following Martin (2013). Consider the objective function in Equation (5) with $\psi \neq 1$[32] The stochastic discount factor is

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

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

$$R_{w,t+1} = \frac{C_{t+1} + W_{t+1}}{W_t} = \frac{C_{t+1}}{C_t}\left(\frac{C_t}{W_t} + \frac{C_t}{W_t}\frac{W_{t+1}}{C_{t+1}}\right) = \frac{C_{t+1}}{C_t}\left(CW + 1\right), \tag{D.2}$$

and the last equality holds if the consumption-wealth ratio CW is constant under the agent's beliefs. I verify the 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

$$\begin{aligned} \frac{P_t}{D_t} &= \mathbb{E}_t\left[\sum_{j=1}^{\infty} \beta^{j\,\eta} \left(\frac{C_{t+j}}{C_t}\right)^{-\frac{\eta}{\psi}} \left(\frac{C_{t+j}}{C_t}\right)^{\lambda} \left(\frac{C_{t+j}}{C_t}\right)^{\eta-1} (1+CW)^{j\,(\eta-1)}\right] \\ &= \sum_{j=1}^{\infty} \beta^{j\,\eta}\, \mathbb{E}_t\left[e^{(\lambda-\gamma)\,g_{t+1}}\right]^j \; e^{j\,(\eta-1)\,cw} = \frac{1}{e^{-\eta\,\log(\beta)+(1-\eta)\,cw-\mathcal{K}_t(\lambda-\gamma)} - 1}, \end{aligned} \tag{D.3}$$

if $-\eta\,\log(\beta) + (1-\eta)\,cw - \mathcal{K}_t(\lambda-\gamma) > 0$ and where $cw = \log\left(1+\frac{C}{W}\right)$. 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

$$cw = -\log(\beta) - \frac{1}{\eta}\mathcal{K}_t(1-\gamma), \tag{D.4}$$

[32] 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 $\eta \to \infty$.

which is constant under the agent's time-t beliefs 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

$$dp_t = -\log(\beta) + \left(1 - \frac{1}{\eta}\right) \mathcal{K}_t(1-\gamma) - \mathcal{K}_t(\lambda - \gamma). \tag{D.5}$$

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

$$\tilde{\mathbb{E}}_{\mathrm{t}}\left[R_{t+1}\right] = \tilde{\mathbb{E}}_{\mathrm{t}}\left[\frac{D_{t+1}}{D_t}\right]\left(1 + \frac{D_t}{P_t}\right) = \tilde{\mathbb{E}}_{\mathrm{t}}\left[e^{\lambda\, g_{t+1}}\right] e^{dp_t} \tag{D.6}$$

and the log of the expected return is

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

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

$$r_t^f = -\log(\beta) + \left(1 - \frac{1}{\eta}\right) \mathcal{K}_t(1-\gamma) - \mathcal{K}_t(-\gamma), \tag{D.8}$$

and the risk premium on any asset is

$$rp_t = \mathcal{K}_t(\lambda) + \mathcal{K}_t(-\gamma) - \mathcal{K}_t(\lambda - \gamma). \tag{D.9}$$

Note that, because the agent perceives the economy as i.i.d. beliefs, the risk premium is independent of the elasticity of intertemporal substitution.

Next, consider the expected return under the econometrician's filtration,

$$\mathbb{E}\left(R_{t+1}\right) = \frac{D_t}{P_t}\, \mathbb{E}\left[\frac{D_{t+1}}{D_t}\right]\left(1 + \mathbb{E}\left[\frac{P_{t+1}}{D_{t+1}}\right]\right), \tag{D.10}$$

where I can no longer use the observation that the dividend-price ratio is subjectively constant. 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

$$r_{t+1} - \tilde{\mathbb{E}}_t(r_{t+1}) = \lambda \left(\mathbb{E}_{t+1} - \mathbb{E}_t\right) \sum_{j=0}^{\infty} \bar{p}^j \, g_{t+1+j} - \left(\mathbb{E}_{t+1} - \mathbb{E}_t\right) \sum_{j=1}^{\infty} \bar{p}^j r_{t+1+j}, \tag{D.11}$$

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

$$r_{t+1} - \tilde{\mathbb{E}}_t(r_{t+1}) = \lambda \left(g_{t+1} - \tilde{\mathbb{E}}_t(g_{t+1})\right) - \frac{\bar{p}}{1-\bar{p}} \left(dp_{t+1} - dp_t\right). \tag{D.12}$$

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

$$r_{t+1} - r_t^f = \lambda g_{t+1} + \mathcal{K}_t(-\gamma) - \mathcal{K}_t(\lambda-\gamma) - \frac{\bar{p}}{1-\bar{p}} \left(dp_{t+1} - dp_t\right). \tag{D.13}$$

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

$$\mathbb{E}\left(r_{t+1}\right) - r_t^f = \lambda \mathbb{E}(g_{t+1}) + \mathcal{K}_t(-\gamma) - \mathcal{K}_t(\lambda-\gamma) - \frac{\bar{p}}{1-\bar{p}} \left(\mathbb{E}\left(dp_{t+1}\right) - dp_t\right), \tag{D.14}$$

with

$$\mathbb{E}(dp_{t+1}) = -\log(\beta) + \left(1 - \frac{1}{\eta}\right) \mathbb{E}\left(\mathcal{K}_{t+1}\left(1-\gamma\right)\right) - \mathbb{E}\left(\mathcal{K}_{t+1}\left(\lambda-\gamma\right)\right). \tag{D.15}$$

In general, we cannot obtain the expectation of the subjective cumulant-generating function in closed-form. A second-order Taylor approximation around the long-term subjective moment-generating function, $\tilde{\mathcal{M}}(k) := \mathbb{E}\left(\mathcal{M}_{t+1}(k)\right)$, gives

$$\mathbb{E}\left(\mathcal{K}_{t+1}(k)\right) = \mathbb{E}\left(\log \mathcal{M}_{t+1}(k)\right) \approx \log\left(\tilde{\mathcal{M}}(k)\right) + \frac{1}{2}\left(\frac{\mathbb{E}\left(\mathcal{M}_{t+1}(k)^2\right)}{\tilde{\mathcal{M}}(k)^2} - 1\right). \tag{D.16}$$

I next show how the approximation can be applied to the similarity-weighted memory discussed in Section 3.

Case 1 (Log-normal endowment endowment growth) First, consider

$$g_t = \mu + \sigma\,\epsilon_t, \tag{D.17}$$

where the agent learns about the mean under similarity-weighted memory. Her posterior mean is $\mu_t = (1-\alpha)\,\mu + \alpha\, g_t$ with $\alpha = \sigma^2/(\kappa+\sigma^2)$.

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

$$\mathcal{M}_t(k) = \tilde{\mathbb{E}}_t\left(e^{k\,g_t}\right) = e^{k\mu_t + \frac{1}{2}\,k^2\,\sigma^2} \tag{D.18}$$

$$\mathcal{K}_t(k) = = \log\left(\mathcal{M}_t(k)\right) = k\mu_t + \frac{1}{2}\,k^2\,\sigma^2. \tag{D.19}$$

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

$$\mathbb{E}\left(\mathcal{K}_t(k)\right) = k\,\mu + \frac{1}{2}\,k^2\,\sigma^2 = \mathcal{K}^*(k). \tag{D.20}$$

Inserting into the previous equations gives

$$r_t^f = -\log(\beta) + \frac{1}{\psi}\,\mu_t - \frac{1}{2}\,\sigma^2\left(\gamma - \frac{1-\gamma}{\psi}\right), \tag{D.21}$$

$$dp_t = -\log(\beta) + \left(\frac{1}{\psi} - \lambda\right)\mu_t - \frac{1}{2}\,\sigma^2\left(\gamma - \frac{1-\gamma}{\psi} + \lambda\,(\lambda - 2\,\gamma)\right), \tag{D.22}$$

$$rp_t = \lambda\,\gamma\,\sigma^2. \tag{D.23}$$

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

$$\mathbb{E}\left(r_{t+1}\right) - r_t^f = \left(\frac{1}{1-\bar{p}}\,\lambda - \frac{\bar{p}}{1-\bar{p}}\frac{1}{\psi}\right)(\mu - \mu_t) + rp_t - \frac{1}{2}\,\lambda^2\,\sigma^2. \tag{D.24}$$

Case 2 (Two-state Markov process) Consider endowment growth as in the main text,

$$g_t = \mu_s + \sigma_s \, \epsilon_t, \tag{D.25}$$

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 (21) in the main text gives the cumulant-generating function under the agent's beliefs. The moment-generating function under the agent's beliefs follows from $\mathcal{M}_t(k) = \exp\left[\mathcal{K}_t(k)\right]$, and Equation (32) gives the expected moment-generating function under the econometrician's beliefs. The asset-pricing quantities under Epstein-Zin preferences are given above.

I now use the second-order approximation of the expected cumulant-generating function to approximate the objective risk premium. It is

$$\mathcal{M}_{t+1}(k)^2 = \pi_1^2 \, e^{2\,k\,\hat{\mu}_{1,t+1} + k^2\,\sigma_1^2} + \pi_2^2 \, e^{2\,k\,\hat{\mu}_{2,t+1} + k^2\,\sigma_2^2} + 2\,\pi_1\,\pi_2\, e^{k\,(\hat{\mu}_{1,t+1} + \hat{\mu}_{2,t+1}) + \frac{1}{2} k^2 \left(\sigma_1^2 + \sigma_2^2\right)}, \tag{D.26}$$

and the objectively expected squared moment-generating function is

$$\mathbb{E}\left(\mathcal{M}_{t+1}(k)^2\right) = \pi_1 \, \mathbb{E}\left[\mathcal{M}_{t+1}(k)^2 | s_{t+1} = 1\right] + \pi_2 \, \mathbb{E}\left[\mathcal{M}_{t+1}(k)^2 | s_{t+1} = 2\right], \tag{D.27}$$

with

$$\begin{aligned} \mathbb{E}\left[\mathcal{M}_{t+1}(k)^2 | s_{t+1} = 1\right] = & \pi_1^2 \, \left[e^{2\,k\,\mu_1 + k^2\,\sigma_1^2\,(1 + 2\,\alpha_1^2)}\right] + \pi_2^2 \, \left[e^{2\,k\,[(1-\alpha_2)\,\mu_2 + \alpha_2\,\mu_1] + k^2\,\sigma_2^2 + 2\,k^2\,\alpha_2^2 \sigma_1^2}\right] \\ & + 2\,\pi_1\,\pi_2 \left[e^{k\,[(1+\alpha_2)\,\mu_1 + (1-\alpha_2)\,\mu_2] + \frac{1}{2} k^2 \left(\sigma_1^2 + \sigma_2^2\right) + \frac{1}{2}\,k^2\,(\alpha_1 + \alpha_2)^2\,\sigma_1^2}\right], \end{aligned} \tag{D.28}$$

$$\begin{aligned} \mathbb{E}\left[\mathcal{M}_{t+1}(k)^2 | s_{t+1} = 2\right] = & \pi_1^2 \, \left[e^{2\,k\,[(1-\alpha_1)\,\mu_1 + \alpha_1\,\mu_2] + k^2\,\sigma_1^2 + 2\,k^2\,\alpha_1^2 \sigma_2^2}\right] + \pi_2^2 \, \left[e^{2\,k\,\mu_2 + k^2\,\sigma_2^2\,(1 + 2\,\alpha_2^2)}\right] \\ & + 2\,\pi_1\,\pi_2 \left[e^{k\,[(1-\alpha_1)\,\mu_1 + (1+\alpha_1)\,\mu_2] + \frac{1}{2} k^2 \left(\sigma_1^2 + \sigma_2^2\right) + \frac{1}{2}\,k^2\,(\alpha_1 + \alpha_2)^2\,\sigma_2^2}\right]. \end{aligned} \tag{D.29}$$

We can insert the expressions into the objectively expected risk premium to derive numerical approximations of the expected risk premium. The approximation is used for Figure 3.

E Data and estimation

This appendix describes the data and estimation procedure used to obtain the parameters for the simulations, as well as to test for (non-)linear relationships between consumption growth and asset prices.

E.1 Data

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 Q2 2025. 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 from December 1929 until December 2024. I deflate dividends using the Consumer Price Index (CPI) series from Shiller's website. The average annualized quarterly endowment growth is 1.80% (4.14% for dividend growth), and the annualized quarterly volatility of endowment growth is 1.80% (5.12% for dividend growth).

Most of the asset pricing data, including the future 1-year excess returns on the CRSP index (December 1946 until December 2019), the subjectively expected excess return (June 1972 until March 2021), and the subjectively perceived volatility (October 2001 until May 2021), is from Nagel and Xu (2023). In addition, I use the 3-month T-bill rate from FRED (December 1964 until December 2024) to measure the risk-free rate, and use the data from Jordà et al. (2019) to report long-term averages.

E.2 Estimation of endowment growth process parameters

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

$$p(\mu_i, \sigma_i^2) \sim \mathcal{NIG}(a_i, A_i, b_i/2, B_i/2) \tag{E.1}$$

$$p(\mu_i | \sigma_i^2) \sim \mathcal{N}(a_i, A_i\, \sigma_i^2) \tag{E.2}$$

$$p(\sigma_i^2) \sim \mathcal{IG}(b_i/2, B_i/2), \tag{E.3}$$

and set the parameters of these distributions as

$$E(\mu_i) = a_i \tag{E.4}$$

$$\text{Var}(\mu_i) = A_i\, \frac{B_i}{b_i - 2} = A_i\, E(\sigma_i^2), \tag{E.5}$$

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

$$E(\sigma_i^2) = \frac{B_i/2}{b_i/2 - 1} \tag{E.6}$$

$$\text{Var}(\sigma_i^2) = \frac{B_i/2}{(b_i/2 - 1)^2\, (b_i/2 - 2)} = E(\sigma_i^2)^2\, \frac{1}{b_i/2 - 2}. \tag{E.7}$$

Thus, I find the parameters as follows:

$$a_i = E(\mu_i) \tag{E.8}$$

$$A_i = \frac{\text{Var}(\mu_i)}{E(\sigma_i^2)} \tag{E.9}$$

$$b_i = 2\, \frac{E(\sigma_i^2)^2}{\text{Var}(\sigma_i^2)} + 4 \tag{E.10}$$

$$B_i = E(\sigma_i^2)\, (b_i - 2)\,. \tag{E.11}$$

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

$$E(\pi_1) = \frac{c_1}{c_1 + C_1} \tag{E.12}$$

$$\text{Var}(\pi_1) = \frac{c_1\, C_1}{(c_1 + C_1)^2\,(C_1 + c_1 + 1)} \tag{E.13}$$

$$= E(\pi_1)\ (1 - E(\pi_1))\ \frac{1}{C_1 + c_1 + 1} \tag{E.14}$$

$$= E(\pi_1)\ (1 - E(\pi_1))\ \frac{1}{\frac{1}{E(\pi_1)}\, c_1 + 1} \tag{E.15}$$

such that I find

$$c_1 = \frac{E(\pi_1)^2\ (1 - E(\pi_1))}{\text{Var}(\pi_1)} - E(\pi_1) \tag{E.16}$$

$$C_1 = c_1 \left(\frac{1 - E(\pi_1)}{E(\pi_1)} \right). \tag{E.17}$$

Table E.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. I use a Markov-Chain-Monte-Carlo (MCMC) and the prior parameters given in Table E.1 to estimate the parameters of endowment growth.

Table E.1: Prior parameters for estimation

Parameter	Mean	Std.
μ_1	0.50%	0.15%
μ_2	-2.00%	0.55%
σ_1^2	$(0.79\%)^2$	$(0.14\%)^2$
σ_2^2	$(2.10\%)^2$	$(1.21\%)^2$
π_1	84.00%	3.80%

Table E.1 reports parameters of the priors used to estimate the properties of an i.i.d. two state Markov-switching process for endowment growth.

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. I assume that, for any asset, $D_t = C_t^{\lambda}$, such that

$$\log \frac{D_{t+1}}{D_t} = \lambda \, g_{t+1}, \tag{E.18}$$

and I accordingly run the regression

$$\log \frac{D_{t+1}}{D_t} = b \, g_{t+1} + \epsilon_{t+1}. \tag{E.19}$$

Empirically, I find $\hat{b} = 3.31$, which is close to the parameters used in the literature. Collin-Dufresne et al. (2016) and Nagel and Xu (2022) use $\lambda = 3$ under a different dividend-growth process, such that I choose $\lambda = 3$ in simulations for comparability.

E.3 Relationship between consumption growth and asset prices

Table E.2 reports empirical summary statistics of the main asset pricing variables that are analyzed in the simulations (Section 3.4). I construct a recession-indicator that is consistent with the notion of the recession state in the model in Section 3.4. Under the maintained assumption of a two-state i.i.d. endowment growth process and the parameters in Table 1, I obtain, for each observed quarterly log endowment growth, the posterior probability that the observation comes from the recession state. I set the indicator equal to one if the posterior probability of being in the recession state is above 0.5. The recession-indicator has a high overlap with the NBER recession indicator, but assigns also quarters with a very high log endowment growth to the recession state, due to the higher fundamental volatility of endowment growth in the recession state.

Table E.2: Empirical summary statistics of main asset pricing variables

	Mean	Std.	N	Recession		g_t quintile				
				Yes	No	Q1	Q2	Q3	Q4	Q5
Objective asset pricing quantities										
Consumption growth g_t	1.80	3.62	311	-0.00	2.05	-1.75	1.07	1.86	2.73	5.17
Risk-free rate r_t^f	3.80	2.90	313	2.72	3.98	3.69	3.58	3.78	4.47	3.61
Risk premium rp_t	6.07	15.29	293	10.00	5.63	4.42	5.91	10.19	3.73	6.22
Subjective asset pricing quantities										
Expected return $\hat{er}_t$	8.64	1.66	143	7.68	8.68	7.79	8.26	8.46	9.50	9.94
Risk premium $\hat{rp}_t$	5.63	1.91	143	6.46	5.59	5.46	5.80	5.29	5.52	6.34
Volatility $\hat{\sigma}_t$	12.98	2.20	75	15.57	12.83	13.97	12.40	12.56	12.26	15.29

Table E.2 reports summary statistics for the main asset pricing variables discussed in Section 3.4. The data is quarterly from 1946Q4 to 2024Q4, but all quantities are annualized. Consumption data is taken from the BEA, the objective and subjectiverisk premium as well as the perceived volatility are taken from Nagel and Xu (2023), and the risk-free rate is the 3-month T-bill rate from FRED. The recession indicator is computed from consumption data under the maintained assumption of a two-state i.i.d. Markov process with the same parameters as in Table 1.

The risk-free rate is volatile (std. of 2.90) and procyclical, averaging 3.98% in normal times and 2.72% in recessions. The objective risk premium is countercyclical, with an average of 10.00% in recessions and 5.63% in normal times, and seems to vary nonlinearly across quintiles of consumption growth. Subjectively expected returns are procyclical. The subjective risk premium is relatively stable (std. of 1.91)—especially compared to the objective risk premium—but higher in recessions (6.46%) than in normal times (5.59%). Across consumption-growth quintiles, the subjective risk premium as well as the perceived volatility both suggest a U-shaped pattern, with higher values in the lowest and highest quintiles—consistent with a priced nonlinear relation between perceived risk and economic conditions.

Table E.3 reports the results of regressions that test for a linear pro- or countercyclical as well as a non-linear quadratic relationship between consumption growth and asset prices. Specifically, for each outcome variable y, I estimate

$$y_t = a_1 + b_1\, g_t + \epsilon_t, \quad \text{and} \quad y_t = a_2 + b_2\, g_t + c_2\, g_t^2 + \epsilon_t,$$

Table E.3: Regression results testing for a non-linear relation between consumption growth and asset pricing variables

Panel A: Objective asset pricing quantities

	Risk-free rate r_t^f		Risk premium rp_t	
	(1)	(2)	(3)	(4)
Constant	0.038	0.039	0.069	0.066
	(0.004)	(0.004)	(0.022)	(0.022)
Cons. growth g_t	0.038	0.010	-0.411	-0.637
	(0.051)	(0.057)	(0.559)	(0.590)
Sq. cons. growth g_t^2	–	-0.285	–	8.447
	–	(0.107)	–	(12.238)
Observations	311	311	291	291
R^2	0.002	0.012	0.003	0.005

Panel B: Subjective asset pricing quantities

	Expected return er_t		Perceived volatility $\hat{\sigma}_t$		Risk premium $\hat{rp}_t$	
	(5)	(6)	(7)	(8)	(9)	(10)
Constant	0.086	0.085	0.130	0.132	0.056	0.055
	(0.003)	(0.003)	(0.006)	(0.006)	(0.003)	(0.003)
Cons. growth g_t	0.055	0.072	0.020	-0.371	-0.002	0.039
	(0.047)	(0.058)	(0.242)	(0.369)	(0.031)	(0.018)
Sq. cons. growth g_t^2	–	0.113	–	4.751	–	0.275
	–	(0.101)	–	(2.837)	–	(0.035)
Observations	143	143	75	75	143	143
R^2	0.023	0.032	0.000	0.073	0.000	0.041

Table E.3 reports regression results testing for a linear and quadratic relation between log consumption growth and the indicated asset pricing variables. Panel A shows results for objective quantities (realized risk premium and risk-free rate); Panel B shows results for subjective quantities (expected return, subjective volatility and the subjective risk premium). Robust standard errors (Newey–West, 4 lags) are in parentheses. The data is quarterly from 1946Q4 to 2024Q4. Consumption data is taken from the BEA, the objective and subjectiverisk premium as well as the perceived volatility are taken from Nagel and Xu (2023), and the risk-free rate is the 3-month T-bill rate from FRED.

where g_t denotes log consumption growth.

Panel A reports results for objective asset prices. Columns (1)–(2) show that the risk-free rate exhibits an insignificant procyclical but significantly concave relation with consumption growth: the rate is higher in good times ($\hat{b}_2 > 0$) but the effect diminishes at extreme values ($\hat{c}_2 < 0$). Columns (3)–(4) show no statistically significant link between consumption growth and the realized risk premium—consistent with the equity premium puzzle—, although the negative coefficient on g_t is consistent with a countercyclical objective risk premium.

Panel B reports results for subjective asset prices. Subjectively expected returns (Columns 5–6) do not significantly vary with consumption growth, though the coefficient signs suggest a procyclical and convex pattern as in the model. Focusing on Column (8), subjective volatility seems to be insignficantly countercyclical, and marginally significantly convex ($\hat{c}_2 > 0$) in consumption growth, consistent with the theoretical prediction of U-shaped perceived risk (Section 3.2). Finally, the subjective risk premium (Column 10) is significantly procyclical ($\hat{b}_2 > 0$) and convex ($\hat{c}_2 > 0$) in consumption growth, in line with the simulation results in Section 3.4.

F Extensions

F.1 Limited history

In this section, I highlight how parameter uncertainty emerges and how it affects asset prices. Since closed-form solutions exist for log-normal endowment growth, I analyze this case first to highlight the mechanism. Thereafter, I describe how I simulate asset prices for the two-state Markov-switching process analyzed in Section 3.

Case 1 (Log-normal endowment endowment growth) First, consider

$$g_t = \mu + \sigma\,\epsilon_t, \tag{F.1}$$

where the agent learns about the mean under similarity-weighted memory. 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 subjective informativeness of the prior. The Bayesian posterior of the agent is

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

with

$$z_t^{-1} = k_t + \nu \tag{F.3}$$

$$\hat{\mu}_t = \frac{1}{z_t^{-1}}\left(\nu\,\mu_0 + \sum_{\tau\in r_t} g_\tau\right). \tag{F.4}$$

The formulation of the agent's posterior mean, $\hat{\mu}_t$ is identical to that in Proposition 1, but we now need to account for the prior belief due to the finite sample. The agent is naïve with respect to her memory distortions and believes that she will surely recall $k_t + 1$ observations next period. The perceived belief and endowment growth dynamics are thus standard under Bayesian learning with

$$g_{t+1} = \hat{\mu}_t + \sqrt{1+z_t}\,\sigma\,\tilde{\epsilon}_{t+1} \tag{F.5}$$

$$\epsilon_{t+1} = \frac{g_{t+1} - \hat{\mu}_t}{\sqrt{1+z_t}\,\sigma} \tag{F.6}$$

$$z_{t+1}^{-1} = z_t^{-1} + 1 \tag{F.7}$$

$$\hat{\mu}_{t+1} = \hat{\mu}_t + \frac{z_t}{\sqrt{1+z_t}}\,\sigma\,\tilde{\epsilon}_{t+1}. \tag{F.8}$$

Asset prices exist in closed form for $\psi = 1$. Define $vc_t = \log(V_t/C_t)$ and use the value function iteration as in Hansen et al. (2008) to find

$$vc_t = \frac{\beta}{1-\gamma}\,\log\left(\tilde{\mathbb{E}}_t\, e^{(1-\gamma)(vc_{t+1}+g_{t+1})}\right). \tag{F.9}$$

I conjecture that the solution is linear in the state variable, $vc_t = a_t + B\,\hat{\mu}_t$, which yields

$$vc_t = \frac{\beta}{1-\gamma}\,\log\left(\tilde{\mathbb{E}}_t\, e^{(1-\gamma)(a_{t+1}+B\,\hat{\mu}_{t+1}+g_{t+1})}\right) = \beta\left(a_{t+1} + (B+1)\,\hat{\mu}_t + \frac{1}{2}(1-\gamma)\,\frac{\left((B+1)\,z_t+1\right)^2}{1+z_t}\sigma^2\right),$$

which we can solve for

$$B = \frac{\beta}{1-\beta}, \tag{F.10}$$

$$\begin{aligned} a_t &= \beta a_{t+1} + \frac{1}{2}\beta(1-\gamma)\ \left((B+1)z_t+1\right)^2 \frac{1}{1+z_t}\,\sigma^2 \\ &= \frac{1}{2}\,\beta\,(1-\gamma)\,\sigma^2\,\sum_{j=0}^{\infty}\beta^j \frac{\left(1+\frac{1}{1-\beta}\,z_{t+j}\right)^2}{1+z_{t+j}}. \end{aligned} \tag{F.11}$$

The geometric series defining a_t converges due to $0 < \beta < 1$ and because z_{t+j} decreases deterministically in j under the agent's subjective beliefs.

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

$$m_{t+1} = \log\left(\beta\left(\frac{C_t}{C_{t+1}}\right)\frac{V_{t+1}^{1-\gamma}}{\tilde{\mathbb{E}}_{\mathrm{t}}\left[V_{t+1}^{1-\gamma}\right]}\right) \tag{F.12}$$

$$= \log\left(\beta e^{-g_{t+1}}\,\frac{e^{((1-\gamma)\,(vc_{t+1}+g_{t+1}))}}{\tilde{\mathbb{E}}_{\mathrm{t}}\left(e^{((1-\gamma)\,(vc_{t+1}+g_{t+1}))}\right)}\right) \tag{F.13}$$

$$= \mu_{m,t} - \hat{\mu}_t - \zeta_t\,\sigma\,\tilde{\epsilon}_{t+1}, \tag{F.14}$$

with

$$\mu_{m,t} = \log(\beta) - \frac{1}{2}\,(1-\gamma)^2\frac{\left((B+1)\,z_t+1\right)^2}{1+z_t}\,\sigma^2, \tag{F.15}$$

$$\zeta_t = \left[\gamma + (\gamma-1)\,B\,\frac{z_t}{1+z_t}\right]\,\sqrt{1+z_t}. \tag{F.16}$$

Shocks to the log SDF are thus

$$m_{t+1} - \tilde{\mathbb{E}}_{\mathrm{t}}(m_{t+1}) = -\zeta_t\,\sigma\,\tilde{\epsilon}_t. \tag{F.17}$$

All assets in the economy are (levered) claims to the endowment growth. I can thus use the joint log-normality of endowment growth and the SDF to find the log return r_{t+1} of any asset as

$$0 = \tilde{\mathbb{E}}_{\mathrm{t}}(m_{t+1}) + \tilde{\mathbb{E}}_{\mathrm{t}}(r_{t+1}) + \frac{1}{2}\,\mathrm{Var}_t(m_{t+1}) + \frac{1}{2}\mathrm{Var}_t(r_{t+1}) + \mathrm{Cov}_t(m_{t+1}, r_{t+1}). \tag{F.18}$$

Next, $\psi = 1$ yields a constant wealth-consumption ratio. Writing the return on the consumption claim as $R_{c,t+1} = \frac{W_{t+1}+C_{t+1}}{W_t}$, we can use the Euler-equation to find $WC = \frac{\beta}{1-\beta}$. The expected log-return on the consumption-claim is

$$\tilde{\mathbb{E}}_{\mathrm{t}}(r_{c,t+1}) = \hat{\mu}_t - \log(\beta), \tag{F.19}$$

and the log risk-free rate is

$$r_t^f = -\,\tilde{\mathbb{E}}_{\mathrm{t}}(m_{t+1}) - \frac{1}{2}\,\mathrm{Var}_t(m_{t+1}) = \hat{\mu}_t - \log(\beta) - \frac{1}{2}\,(1+z_t)\,\sigma^2 + (1-\gamma)\,[(B+1)z_t + 1]\;\sigma^2. \tag{F.20}$$

The subjective risk premium on the consumption-claim is then

$$\begin{aligned} \tilde{\mathbb{E}}_{\mathrm{t}}(r_{c,t+1}) - r_t^f &= \frac{1}{2}\,(1+z_t)\,\sigma^2 - (1-\gamma)\,[(B+1)z_t + 1]\;\sigma^2 \\ &= \gamma\,\sigma^2\,(1+z_t) - \frac{1}{2}\,(1+z_t)\,\sigma^2 + (\gamma-1)\,B\,z_t\,\sigma^2. \end{aligned} \tag{F.21}$$

The subjective risk-premium contains a Jensen's term since we consider the expected log return. Instead, in the main text, we consider the log of the expected return, and we need to adjust for the subjective risk-premium by adding a Jensen's correction to find

$$\tilde{\mathbb{E}}_{\mathrm{t}}(r_{c,t+1}) - r_t^f + \frac{1}{2}Var_t(r_{c,t+1}) = \log(\tilde{\mathbb{E}}_{\mathrm{t}}(R_{c,t+1}) - r_t^f = \gamma\,\sigma^2\,(1+z_t) + (\gamma-1)\,B\,z_t\,\sigma^2.$$

For $\gamma > 1$, the risk-premium is increasing in parameter uncertainty.

Next, let us derive the objective risk premium, given by $\log(\mathbb{E}(R_{c,t+1}) - r_t^f$. The econometrician knows the parameters of the economy, such that $\mathbb{E}(R_{c,t+1}) = \frac{1}{\beta}\,\mathbb{E}(e^{g_{t+1}}) = \frac{1}{\beta}e^{\mu+\frac{1}{2}\sigma^2}$.

The objective risk premium is thus

$$\log(\mathbb{E}(R_c)) - r_t^f = \underbrace{(\mu - \hat{\mu}_t)}_{\text{Belief wedge}} + \underbrace{\gamma\,\sigma^2\,(1+z_t) + (\gamma-1)\,B\,z_t\,\sigma^2}_{\text{Subjective risk premium}} \underbrace{-\frac{1}{2}\,z_t\,\sigma^2}_{\text{Jensen's inequality}} \tag{F.22}$$

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 that arises because the agent and the econometrician perceive a different volatility of the consumption claim. We can equivalently derive the objective risk premium on the dividend-paying asset.

Case 2 (Two-state Markov process) Consider endowment growth as in the main text,

$$g_t = \mu_s + \sigma_s\,\epsilon_t, \tag{F.23}$$

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. Closed-form solutions for asset prices with parameter uncertainty cease to exist, such that I detail the numerical procedure to obtain asset pricing equations, following 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 $H^R_{s,t}$ is

$$\mu \sim \mathcal{N}\left(\hat{\mu}_{s,t}, z_{s,t}\, \sigma_s^2\right), \tag{F.24}$$

with

$$z_{s,t} = \left(k_{s,t} + \nu_s\right)^{-1}, \tag{F.25}$$

$$\hat{\mu}_{s,t} = z_{s,t} \left(\nu_s\, \hat{\mu}_{s,0} + \sum_{\tau \in H^R_{s,t}} g_\tau\right). \tag{F.26}$$

Due to the agent's naïvete, her perceived dynamics of endowment and beliefs are the same as for a rational Bayesian agent, with

$$z^{-1}_{t+1,s} = z^{-1}_{s,t} + \mathbb{1}_{s_{t+1}=s} \tag{F.27}$$

$$\hat{\mu}_{t+1,s} = \mu_{s,t} + \mathbb{1}_{s_{t+1}=s} \frac{z_{s,t}}{1+z_{s,t}} \left(g_{t+1} - \mu_{s,t}\right), \tag{F.28}$$

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

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

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

$$\tilde{\mathbb{E}}_t \left[\beta^\eta \left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\eta}{\psi}} R^{\eta}_{w,t+1}\right] = 1, \tag{F.29}$$

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

$$\left(\frac{W_t}{C_t}\right)^{\eta} = \beta^{\eta}\, \tilde{\mathbb{E}}_{\mathrm{t}} \left[e^{(1-\gamma)\, g_{t+1}} \left(\frac{W_{t+1}}{C_{t+1}} + 1\right)^{\eta}\right]. \tag{F.30}$$

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

$$\begin{aligned} WC(X_t, s_t)^{\eta} = \beta^{\eta}\, \pi_1\, \tilde{\mathbb{E}}_{\mathrm{t}} \left(e^{(1-\gamma)\, g_{t+1}}\, (WC(X_t, s_{t+1}, g_{t+1}) + 1)^{\eta}\, | s_{t+1} = 1, s_t, X_t\right) + \\ \beta^{\eta}\, \pi_2\, \tilde{\mathbb{E}}_{\mathrm{t}} \left(e^{(1-\gamma)\, g_{t+1}}\, (WC(X_t, s_{t+1}, g_{t+1}) + 1)^{\eta}\, | s_{t+1} = 2, s_t, X_t\right), \end{aligned} \tag{F.31}$$

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:

$$WC_{\infty} = \frac{\beta\, e^{\frac{1}{\eta}\, \mathcal{K}(1-\gamma)}}{1 - \beta\, e^{\frac{1}{\eta}\, \mathcal{K}(1-\gamma)}}, \tag{F.32}$$

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 expect to learn when state 1 realizes. The wealth-consumption ratio is then

$$WC(X_t, s_t)^\eta = \beta^\eta \, \pi_1 \, e^{(1-\gamma)\hat{\mu}_{1,\infty} + \frac{1}{2}(1-\gamma)^2 \, \sigma_1^2} \, (WC(X_t, s_t) + 1)^\eta + \beta^\eta \, \pi_2 \, \tilde{\mathbb{E}}_{\mathrm{t}} \left(e^{(1-\gamma)\, g_{t+1}} \, (WC(X_t, s_{t+1}, g_{t+1}) + 1)^\eta \, | s_{t+1} = 2, s_t, X_t \right). \tag{F.33}$$

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 the true mean μ_2. I use a Gauss-Hermite quadrature to approximate the expectation. The numerical approximation for the expectation is

$$\tilde{\mathbb{E}}_{\mathrm{t}} \left(e^{(1-\gamma)\, g_{t+1}} \, (WC(X_t, s_{t+1}, g_{t+1}) + 1)^\eta \, | s_{t+1}, s_t, X_t \right) \tag{F.34}$$

$$\approx \sum_{j=1}^{J} \omega_\epsilon(j) \sum_{k=1}^{K} \omega_{\mu_{2,t}}(k) \left(e^{(1-\gamma)\, g_{t+1}} \, (WC(X_t, s_{t+1}, g_{t+1}) + 1)^\eta \, | s_{t+1} = 2, s_t, X_t \right), \tag{F.35}$$

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

$$g_{t+1}(k, j) = n_{\mu_{2,t}}(k) + \sigma_s \, n_\epsilon(j), \tag{F.36}$$

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,

$$R_{t+1} = \frac{P_{t+1} + D_{t+1}}{P_t} = \frac{D_t}{P_t}\,\frac{D_{t+1}}{D_t}\,\left(\frac{P_{t+1}}{D_{t+1}} + 1\right) = e^{\lambda\, g_{t+1}}\,\frac{PD(X_t, s_{t+1}, g_{t+1}) + 1}{PD(X_t, s_t)}, \tag{F.37}$$

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

$$1 = \tilde{\mathbb{E}}_{\mathrm{t}}\left[M_{t+1}\,R_{t+1}\right] = \tilde{\mathbb{E}}_{\mathrm{t}}\left[\beta^{\eta}\,\left(\frac{C_{t+1}}{C_t}\right)^{-\frac{\eta}{\psi}}\,R_{w,t+1}^{\eta-1}\,R_{t+1}\right]. \tag{F.38}$$

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

$$PD(X_t, s_t) = \beta^{\eta}\,\tilde{\mathbb{E}}_{\mathrm{t}}\left[e^{(\lambda-\gamma)\,g_{t+1}}\,\left(\frac{WC(X_{t+1}, s_{t+1}) + 1}{WC(X_t, s_t)}\right)^{\eta-1}\,\left(PD(X_t, s_{t+1}, g_{t+1}) + 1\right)\right] \tag{F.39}$$

$$\begin{aligned} &= \beta^{\eta}\,\pi_1\,\tilde{\mathbb{E}}_{\mathrm{t}}\left[e^{(\lambda-\gamma)\,g_{t+1}}\,\left(\frac{WC(X_{t+1}, s_{t+1}) + 1}{WC(X_t, s_t)}\right)^{\eta-1}\,\left(PD(X_t, s_{t+1}, g_{t+1}) + 1\right)|s_{t+1} = 1\right] + \\ &\quad \beta^{\eta}\,(1-\pi_1)\,\tilde{\mathbb{E}}_{\mathrm{t}}\left[e^{(\lambda-\gamma)\,g_{t+1}}\,\left(\frac{WC(X_{t+1}, s_{t+1}) + 1}{WC(X_t, s_t)}\right)^{\eta-1}\,\left(PD(X_t, s_{t+1}, g_{t+1}) + 1\right)|s_{t+1} = 2\right]. \end{aligned} \tag{F.40}$$

We can thus solve for the price-dividend ratio of any asset exactly 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:

$$er_t = \log\left(\tilde{\mathbb{E}}_{\mathrm{t}}\ R_{t+1}\right) \tag{F.41}$$

$$r_t^f = \log\left(\tilde{\mathbb{E}}_{\mathrm{t}}\ R_{t+1}^f\right) \tag{F.42}$$

$$rp_t = er_t - r_t^f, \tag{F.43}$$

as well as the following objective quantity

$$rp_t^o = \log\left(\mathbb{E}\, R_{t+1}\right) - r_t^f. \tag{F.44}$$

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

$$r_t^f = \log\left[\tilde{\mathbb{E}}_t\left(\frac{PD(X_{t+1}, s_{t+1}|\lambda = 0) + 1}{PD(X_t, s_t|\lambda = 0)}\right)\right], \tag{F.45}$$

$$er_t = \log\left[\tilde{\mathbb{E}}_t\left(e^{\lambda\, g_{t+1}}\,\frac{PD(X_{t+1}, s_{t+1}) + 1}{PD(X_t, s_t)}\right)\right], \tag{F.46}$$

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 $R_{t+1} = \frac{P_{t+1}+D_{t+1}}{P_t}$.

F.2 Peak-end rule

In this appendix, I derive numerical approximations for the agent's beliefs under the peak-end rule. I focus on the effect of the extreme experience bias on the agent's posterior beliefs in this appendix. Define the extreme experience bias in Equation (42) as

$$m^P(g_\tau) := \exp\left[-e^{-\frac{(g_\tau-\mu)^2}{2\,\sigma^2}}\right]. \tag{F.47}$$

F.2.1 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 $X^*_n = \max_{1\leq\tau\leq n} X_\tau$. 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

$$\mathbb{P}(X^*_n \leq x) = \mathbb{P}\left(X_1 \leq x, X_2 \leq x, ..., X_n \leq x\right) = \prod_{j=1}^{n} P(X_j \leq n) = \Phi^n(x). \tag{F.48}$$

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

$$\Phi(-x) = \int_x^\infty \phi(z)\, dz \tag{F.49}$$

$$\leq \int_x^\infty \frac{z}{x}\phi(z)\, dz = \frac{1}{x\sqrt{2\pi}} \int_x^\infty z\, e^{-\frac{z^2}{2}}\, dz = \frac{1}{x}\phi(x). \tag{F.50}$$

We can tighten the bound using Gordon's Inequality as

$$1 \leq \frac{\phi(x)}{x\,\Phi(-x)} \leq 1 + \frac{1}{x^2}. \tag{F.51}$$

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

$$\log\Phi(-a_n - b_n z) = \log\Phi(-a_n) - b_n\, z\,\frac{\phi(-a_n)}{\Phi(-a_n)} + o(b_n\, z) = \log\frac{1}{n} - z + o(b_n\, z). \tag{F.52}$$

We can then find

$$\Pr\left(X_1 \le a_n + b_n z\right) = \Phi(a_n + b_n z) = 1 - \Phi(-a_n - b_n z) \approx 1 - \frac{1}{n}e^{-z}, \tag{F.53}$$

and

$$\Pr\left(X_n^* \le a_n + b_n z\right) \approx \left[1 - \frac{1}{n}e^{-z}\right]^n \approx \exp\left(-e^{-z}\right) := G(z), \tag{F.54}$$

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. □

The CDF of the Gumbel-distribution for the maximum is

$$G(z;\mu,\sigma^2) = \exp\left(-e^{-\frac{z-\mu}{2\sigma^2}}\right), \tag{F.55}$$

and I obtain $m^P(g_\tau)$ by squaring the distance in the double exponential to get a symmetric overweighting of the tails.

F.2.2 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

$$\int_{-\infty}^{\infty} \tilde{m}^P(g)\, q^*(g)\, dg = \int_{\mathbb{R}} e^{-\frac{(g-\mu)^2}{2\,\sigma^2}}\; e^{-e^{-\frac{(g-a)^2}{2\,b}}}\; dg \tag{F.56}$$

$$= \int_{\mathbb{R}} e^{-\frac{(g-\mu)^2}{2\,\sigma^2}} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!}\, e^{-k\,\frac{(g-a)^2}{2\,b}}\, dg \tag{F.57}$$

$$= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{\mathbb{R}} e^{-\frac{g^2\left(\sigma^2 k + b\right) - 2g\left(\sigma^2\, k\, a + \mu\, b\right) + \left(\sigma^2\, k\, a^2 + \mu^2\, b\right)}{2\sigma^2\, b}}\, dg \tag{F.58}$$

$$= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \sqrt{\frac{2\,\pi\,\sigma^2\, b}{\sigma^2\, k + b}}\; e^{-\frac{k}{2}\frac{(\mu - a)^2}{\sigma^2\, k + b}} = A^{-1}, \tag{F.59}$$

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.

F.2.3 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 $b = \sigma^2$.[33] 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

$$\hat{\mu}_t = A \int_{\mathbb{R}} g\; e^{-\frac{(g-\mu)^2}{2\sigma^2}}\; e^{-e^{-\frac{(g-\mu)^2}{2\,\sigma^2}}}\; dg \tag{F.60}$$

$$= A \int_{\mathbb{R}} (x + \mu)\, e^{-\frac{x^2}{2\sigma^2}}\; e^{-e^{-\frac{x^2}{2\sigma^2}}}\; dx \tag{F.61}$$

[33]Similar results are available for more general versions with either $a \neq \mu$ or $b \neq \sigma^2$.

$$= A \int_{\mathbb{R}} x\, e^{-\frac{x^2}{2\sigma^2}}\, e^{-e^{-\frac{x^2}{2\sigma^2}}}\, dx + \mu\, A \int_{\mathbb{R}} e^{-\frac{x^2}{2\sigma^2}}\, e^{-e^{-\frac{x^2}{2\sigma^2}}}\, dx \tag{F.62}$$

$$= A \int_{\mathbb{R}} x\, e^{-\frac{x^2}{2\sigma^2}}\, e^{-e^{-\frac{x^2}{2\sigma^2}}}\, dx + \mu, \tag{F.63}$$

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

$$\int_{\mathbb{R}} x\, e^{-\frac{x^2}{2\sigma^2}}\, e^{-e^{-\frac{x^2}{2\sigma^2}}}\, dx = \int_{-\infty}^{0} x\, e^{-\frac{x^2}{2\sigma^2}}\, e^{-e^{-\frac{x^2}{2\sigma^2}}}\, dx + \int_{0}^{\infty} x\, e^{-\frac{x^2}{2\sigma^2}}\, e^{-e^{-\frac{x^2}{2\sigma^2}}}\, dx \tag{F.64}$$

$$= \int_0^1 -\sigma^2\, e^{-y}\, dy + \int_1^0 -\sigma^2\, e^{-y}\, dy \tag{F.65}$$

$$= -\sigma^2 \left(\int_0^1 e^{-y}\, dy - \int_0^1 e^{-y}\, dy \right) = 0, \tag{F.66}$$

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

$$\hat{\sigma}_t^2 = A \int_{\mathbb{R}} (g-\mu)^2\; e^{-\frac{(g-\mu)^2}{2\,\sigma^2}}\; e^{-e^{-\frac{(g-\mu)^2}{2\,\sigma^2}}}\, dg = A\, 2\sqrt{2}\,\sigma^3 \int_{\mathbb{R}} u^2\, e^{-u^2}\, e^{-e^{-u^2}}\, du. \tag{F.67}$$

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

$$\int_{\mathbb{R}} u^2\, e^{-u^2}\, e^{-e^{-u^2}}\, du = 2 \int_0^{\infty} u^2\, e^{-u^2}\, e^{-e^{-u^2}}\, du \tag{F.68}$$

$$= \int_0^1 \sqrt{-\ln(y)}\; e^{-y}\, dy = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_0^1 \sqrt{-\ln(y)}\; y^k\, dy. \tag{F.69}$$

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

$$\int_0^1 \sqrt{-\ln(y)}\; y^k\, dy = \int_{\infty}^{0} \sqrt{v}\, e^{-k\,v} \left(\frac{dy}{dv}\right) dv = \int_0^{\infty} \sqrt{v} e^{-(k+1)v}\, dv \tag{F.70}$$

$$= \frac{\Gamma(1.5)}{(k+1)^{3/2}} = \frac{\sqrt{\pi}}{2\,(k+1)^{3/2}}, \tag{F.71}$$

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

Putting terms together, it is

$$\hat{\sigma}_t^2 = A\,2\,\sqrt{2}\,\sigma^3 \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{\sqrt{\pi}}{2\,(k+1)^{3/2}} = A\,\sqrt{2\,\pi}\,\sigma^3 \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{1}{(k+1)^{3/2}} \tag{F.72}$$

$$= \sigma^2 \frac{\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{1}{(k+1)^{3/2}}}{\sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \frac{1}{\sqrt{m+1}}} \approx \sigma^2 \cdot 1.4108 > \sigma^2. \tag{F.73}$$

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 beliefs.

F.2.4 Numerical approximation of the subjective mean under the peak-end rule

Let us consider the entire memory function defined in Equation (42). First, we need to show that the memory-weighted probability distribution exists by showing that the integration constant $\mathcal{V}$ exists:

$$\int_{-\infty}^{\infty} m^{PE}(g, g_t)\, q^*(g)\, dg = \int_{\mathbb{R}} e^{-\frac{(g-\mu)^2}{2\,\sigma^2}}\; e^{-\frac{(g-g_t)^2}{2\,\kappa}}\; e^{-e^{-\frac{(g-\mu)^2}{2\,\sigma^2}}}\, dg \tag{F.74}$$

$$= \int_{\mathbb{R}} e^{-\left[\frac{(g-\mu)^2}{2\,\sigma^2} + \frac{(g-g_t)^2}{2\,\kappa}\right]} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!}\, e^{-k\,\frac{(g-\mu)^2}{2\,\sigma^2}}\, dg$$

$$= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{\mathbb{R}} e^{-\left[(1+k)\,\frac{(g-\mu)^2}{2\,\sigma^2} + \frac{(g-g_t)^2}{2\,\kappa}\right]}\, dg$$

$$= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{\mathbb{R}} e^{-\frac{g^2\left((1+k)\,\kappa+\sigma^2\right)-2g\left((1+k)\,\kappa\,\mu+\sigma^2\,g_t\right)+\left((1+k)\,\kappa\,\mu^2+\sigma^2\,g_t^2\right)}{2\sigma^2\,\kappa}}\, dg$$

$$= \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \sqrt{\frac{2\,\pi\,\sigma^2\,\kappa}{\sigma^2+(1+k)\,\kappa}}\; e^{-\frac{(1+k)}{2}\,\frac{(\mu-g_t)^2}{(1+k)\,\kappa+\sigma^2}} = \mathcal{V}^{-1}, \tag{F.75}$$

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

$$\hat{\mu}_t^S = \frac{\kappa\,\mu + \sigma^2\,g_t}{\kappa + \sigma^2} = (1-\alpha)\,\mu + \alpha\,g_t \tag{F.76}$$

$$(\hat{\sigma}_t^S)^2 = \frac{\kappa\sigma^2}{\kappa + \sigma^2}, \tag{F.77}$$

with $\alpha = \frac{\sigma^2}{\kappa+\sigma^2}$. We can then rewrite the product of the peak-end memory function with the probability density function of a normal distribution as

$$e^{-\frac{(g_T-\mu)^2}{2\,\sigma^2}}\; e^{-\frac{(g_T-g_t)^2}{2\,\kappa}}\; e^{-e^{-\frac{(g_T-\mu)^2}{2\,\sigma^2}}} = e^{-\frac{(\mu-g_t)^2}{2\,(\kappa+\sigma^2)}}\; e^{-\frac{(g_T-\hat{\mu}_t^S)^2}{2\,(\hat{\sigma}_t^S)^2}}\; e^{-e^{-\frac{(g_T-\mu)^2}{2\,\sigma^2}}}. \tag{F.78}$$

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

$$\hat{\mu}_t = \mathcal{V}\,e^{-\frac{(\mu-g_t)^2}{2\,(\kappa+\sigma^2)}} \int_{\mathbb{R}} g\,e^{-\frac{(g-\hat{\mu}_t^S)^2}{2\,(\hat{\sigma}_t^S)^2}}\; e^{-e^{-\frac{(g-\mu)^2}{2\,\sigma^2}}}\,dg \tag{F.79}$$

$$= \mathcal{V}\,e^{-\frac{(\mu-g_t)^2}{2\,(\kappa+\sigma^2)}} \int_{\mathbb{R}} \left(x + \mu + \hat{\mu}_t^S\right) e^{-\frac{(x+\mu)^2}{2\,(\hat{\sigma}_t^S)^2}}\; e^{-e^{-\frac{(x+\hat{\mu}_t^S)^2}{2\,\sigma^2}}}\,dx \tag{F.80}$$

$$= \mathcal{V}\,e^{-\frac{(\mu-g_t)^2}{2\,(\kappa+\sigma^2)}} \int_{\mathbb{R}} x\,e^{-\frac{(x+\mu)^2}{2\,(\hat{\sigma}_t^S)^2}}\; e^{-e^{-\frac{(x+\hat{\mu}_t^S)^2}{2\,\sigma^2}}}\,dx + \mu + \hat{\mu}_t^S. \tag{F.81}$$

We cannot simplify the integral using the same steps as in Appendix F.2.3, 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

$$\int_{\mathbb{R}} x\,e^{-\frac{(x+\mu)^2}{2\,(\hat{\sigma}_t^S)^2}}\; e^{-e^{-\frac{(x+\hat{\mu}_t^S)^2}{2\,\sigma^2}}} = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \int_{\mathbb{R}} x\,e^{-\frac{(x+\mu)^2}{2\,(\hat{\sigma}_t^S)^2}}\; e^{-k\,\frac{(x+\hat{\mu}_t^S)^2}{2\,\sigma^2}}\,dx \tag{F.82}$$

$$= - \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{\sqrt{2\,\pi}\,\sigma\,\hat{\sigma}_t^S}{(\sigma^2 + k\,(\hat{\sigma}_t^S)^2)^{3/2}} \cdot \left(\mu\,\sigma^2 + k\,\hat{\mu}_t^S\,(\hat{\sigma}_t^S)^2\right)\, e^{-\frac{k}{2}\cdot\frac{(\mu-\hat{\mu}_t^S)^2}{\sigma^2+k\,(\hat{\sigma}_t^S)^2}}. \tag{F.83}$$

Putting terms together, it is

$$\begin{aligned} \hat{\mu}_t &= \mu + \hat{\mu}_t^S - \mathcal{V}\, e^{-\frac{(\mu-g_t)^2}{2\,(\kappa+\sigma^2)}} \left(\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{\sqrt{2\,\pi}\,\sigma\,\hat{\sigma}_t^S}{(\sigma^2 + k\,(\hat{\sigma}_t^S)^2)^{3/2}} \left(\mu\,\sigma^2 + k\,\hat{\mu}_t^S\,(\hat{\sigma}_t^S)^2\right) e^{-\frac{k}{2}\cdot\frac{(\mu-\hat{\mu}_t^S)^2}{\sigma^2+k\,(\hat{\sigma}_t^S)^2}} \right) \\ &= \mu + \hat{\mu}_t^S - e^{-\frac{(\mu-g_t)^2}{2\,(\kappa+\sigma^2)}} \frac{\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{\kappa+\sigma^2}{((1+k)\,\kappa+\sigma^2)^{3/2}} \left(\mu + k\,\frac{\kappa}{\kappa+\sigma^2}\hat{\mu}_t^S\right) e^{-\frac{k}{2}\cdot\frac{\alpha^2\,(\mu-g_t)^2}{\sigma^2+k\,(\hat{\sigma}_t^S)^2}}}{\sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \sqrt{\frac{1}{\sigma^2+(1+k)\,\kappa}}\, e^{-\frac{(1+k)}{2}\,\frac{(\mu-g_t)^2}{(1+k)\,\kappa+\sigma^2}}}, \end{aligned} \tag{F.84}$$

which is fast to evaluate numerically.

F.3 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

$$g_t = \mu + \sigma\epsilon_t, \quad \epsilon_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1). \tag{F.85}$$

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 (17).

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

$$\hat{\mu}_t = (1-\alpha)\,\mu + \alpha\,g_t, \text{ and} \tag{F.86}$$

$$\hat{\sigma}_t^2 = (1-\alpha)\,\sigma^2, \tag{F.87}$$

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 fundamental 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

$$\mathcal{K}_t^{SL}(k) = \log \tilde{\mathbb{E}}_t \left(e^{k\,g_{t+1}}\right) = k\,\hat{\mu}_t + \frac{1}{2}\,k^2\,\hat{\sigma}_t^2 = \underbrace{\alpha\,k\,g_t}_{\text{Time-varying}} + \underbrace{(1-\alpha)\,\left[k\,\mu + \frac{1}{2}\,k^2\,\sigma^2\right]}_{\text{Constant}}. \tag{F.88}$$

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 $50,000$ times for 312 quarters and report average moments in Table 5. All parameters are as in Section 4.3.

The simulation results in Table F.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.

Table F.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	1.449	2.021	-
$\hat{\mu}_t$	1.450	0.010	1.000
$\hat{\sigma}_t$	2.018	< 0.001	< 0.001
Subjective asset prices			
er_t	4.573	0.006	1.000
rp_t	1.223	0.002	0.671
Objective asset prices			
r_t^f	3.350	0.010	0.956
rp_t	1.428	1.282	-0.956

Table F.1 reports the model moments obtained from $50,000$ simulations of the model for 312 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.

F.4 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

$$m^S_{(g_t,s_t)}\left(g_\tau, s_\tau\right) = \exp\left[-\frac{\left(g_\tau - g_t\right)^2}{\left(2 - |s_t - s_\tau|\right)\kappa}\right]. \tag{F.89}$$

If $s_t = s_\tau$, the memory function is as in Equation (17). 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 (F.89), 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

$$\hat{\mu}_{1,t} = \mu_1 + \begin{cases} \frac{\sigma_1^2}{\sigma_1^2+\kappa}\,(g_t - \mu_1) & \text{if } s_t = 1 \\ \frac{2\,\sigma_1^2}{2\,\sigma_1^2+\kappa}\,(g_t - \mu_1) & \text{if } s_t = 2 \end{cases}, \text{ and} \tag{F.90}$$

$$\hat{\mu}_{2,t} = \mu_2 + \begin{cases} \frac{2\,\sigma_2^2}{2\,\sigma_2^2+\kappa}\,(g_t - \mu_2) & \text{if } s_t = 1 \\ \frac{\sigma_2^2}{\sigma_2^2+\kappa}\,(g_t - \mu_2) & \text{if } s_t = 2 \end{cases}. \tag{F.91}$$

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

$$\hat{\sigma}_{1,t}^2 = \sigma_1^2 \cdot \begin{cases} \frac{\kappa}{\kappa+\sigma_1^2} & \text{if } s_t = 1 \\ \frac{\kappa}{\kappa+2\,\sigma_1^2} & \text{if } s_t = 2 \end{cases}, \tag{F.92}$$

$$\hat{\sigma}_{2,t}^2 = \sigma_2^2 \cdot \begin{cases} \frac{\kappa}{\kappa+2\,\sigma_2^2} & \text{if } s_t = 1 \\ \frac{\kappa}{\kappa+\sigma_2^2} & \text{if } s_t = 2 \end{cases}. \tag{F.93}$$

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 shrunk more than her posterior variance of endowment growth in the current state.

G Selective memory for continuous distributions

I now extend the learning framework from Section 2 to continuous outcome distributions.

Economy. Let us assume that the realized state $s_t = s$ induces a fixed and i.i.d. density q_s^* of log endowment growth g, such that $Pr\,(g \in [a,b]) = \int_a^b q_s^*(g)\,dg$ conditional on $s_t = s$.

The density $q_s^* \in \mathcal{D}$, where $\mathcal{D}$ is the set of densities over $\mathbb{R}$.[34] I maintain the assumption that q_s^* belongs to the family of parametric 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 state 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.[35]

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 experience as

$$\mathrm{LM}^c(g_t, s_t) = \operatorname*{argmax}_{q \in Q} \sum_{s \in S} \psi(s) \int_{-\infty}^{\infty} m_{(g_t,s_t)}(g, s)\, q_s^*(g)\, \log q_s(g)\, dg. \tag{G.1}$$

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. (2024) and proceeds

[34] 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 $\upsilon : \Omega \mapsto \mathbb{R}$ that are non-negative almost everywhere and satisfy $\int_\Omega \upsilon(x)\, dx = 1$.

[35] 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.

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 the claim.

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

$$\hat{F}_t(g, s) = \frac{1}{|H_t^R|} \sum_{\tau \in r_t} \mathbb{1}\{g_\tau \leq g, s_\tau \leq s\}, \tag{G.2}$$

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[36] ensures uniform almost sure convergence of the empirical joint distribution, $\hat{F}_t(g, s)$, to the true distribution, $F(g, s)$ as $t \to \infty$:

$$\sup_{g \in \mathbb{R}, s \in S} \left| \hat{F}_t(g, s) - F(g, s) \right| \overset{\text{a.s.}}{\to} 0. \tag{G.3}$$

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—up to a normalization constant—to $m_{(g_t,s_t)}(g, s) \cdot F(g, s)$, that is

$$\hat{F}_t(g, s) - m_{(g_t,s_t)}(g, s) \cdot F(g, s) \overset{\text{a.s.}}{\to} 0. \tag{G.4}$$

[36]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).

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

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

Consequently,

$$\sup_{x \in \mathbb{R} \times S} \left| \hat{F}_t(x) - F_{m,t}(x) \right| \leq \max_{k \in \{1,...,m\}} \left| \hat{F}_t(x_k) - F_{m,t}(x_k) \right| + \frac{1}{m}. \tag{G.5}$$

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 $t \geq \text{T}$ 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

$$\hat{f}_t(g) = \sum_{s \in S} \frac{\hat{F}_t(d_{k+1}, s) - \hat{F}_t(d_k, s)}{h}, \tag{G.6}$$

for $g \in [d_{k+1}, d_k]$. Note that the marginal memory-weighted distribution of endowment growth, $f_{m,t}(g)$, 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 $\text{Var}\left(\hat{f}_t(g)\right) = \frac{\Pr(d_k \leq g \leq d_{k+1})\,(1-\Pr[d_k \leq g \leq d_{k+1}])}{|H_t^R|\,h^2}$. These properties of the empirical histogram of recalled growth rates imply that

$$\hat{f}_t(g) \xrightarrow{p} f_{m,t}(g). \tag{G.7}$$

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

$$\begin{aligned}
\log\left(\prod_{\tau \in r_t} q_{s_\tau}(g_\tau)\right) &= \sum_{s \in S} \psi(s) \int_{-\infty}^{\infty} |H_t^R|\, \hat{f}_t(g)\, \log q_s\,(g)\; dg &\text{(G.8)}\\
&\xrightarrow{p} |H_t^R| \sum_{s \in S} \psi(s) \int_{-\infty}^{\infty} f_{m,t}(g)\, \log q_s\,(g)\; dg &\text{(G.9)}\\
&= |H_t^R| \sum_{s \in S} \psi(s) \int_{-\infty}^{\infty} m_{(g_t,s_t)}(g,s)\, q_s^*(g)\, \log q_s\,(g)\; dg &\text{(G.10)}\\
&= |H_t^R|\, L(q, H_t^R), &\text{(G.11)}
\end{aligned}$$

where I used the convergence of the empirical density $\hat{f}_t(g)$ to the memory-weighted true density from Step 1 to go from the second to the third line, 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

$$\frac{\prod_{\tau \in r_t} q_{s_\tau}(g_\tau)\, b_0(q)}{\prod_{\tau \in r_t} q'_{s_\tau}(g_\tau)\, b_0(q')} = \rho \frac{\exp\left[\log \prod_{\tau \in r_t} q_{s_\tau}(g_\tau)\right]}{\exp\left[\log \prod_{\tau \in r_t} q'_{s_\tau}(g_\tau)\right]}$$
$$= \rho \exp\left[|H_t^R| \left(L(q, H_t^R) - L(q', H_t^R)\right)\right]. \tag{G.12}$$

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: Finally, I must argue that indeed $|H_t^R| \to \infty$ for $t \to \infty$, which follows directly from claim 1 in Fudenberg et al. (2024) and the proof therein.