Investor beliefs and asset prices under selective memory

Maximilian Voigt*

November 5, 2023

Click here for the latest version.

— Job Market Paper —

Abstract

Evidence in economics and psychology emphasizes the central role of selective memory for belief formation. Motivated by these findings, I propose a consumption-based asset pricing model in which the representative agent forms Bayesian beliefs based on selectively recalled observations of fundamentals. I demonstrate that *similarity*—the selective recall of past fundamentals that are similar to today's fundamental—captures procyclical and overreacting expectations, countercyclical subjective volatility, and a low volatility of the subjective risk premium. The objective risk premium, on the contrary, is predictably countercyclical, volatile, and unrelated to objective risk measures. In addition, I show how my framework can be used to analyze further memory distortions by considering an extreme experience bias.

*I especially thank my advisors Sebastian Ebert and Francesco Sangiorgi for their invaluable guidance and support. I also thank Nick Barberis, Giovanni Burro, Adrian Buss, Paula Cocoma, Chukwuma Dim, Andreas Grunewald, Jonas Happel, Björn Höppner, Jonas Jensen, Yigitcan Karabulut, Emanuel Mönch, Cameron Peng, Francesco Sannino, Paul Schmidt-Engelbertz, Frederik Schwerter, Kaushik Vasudevan, as well as seminar participants at the Frankfurt-Bonn-Mannheim PhD conference, the Finance Theory Group, Frankfurt School, and Yale for their helpful suggestions and feedback. The author is from the Frankfurt School of Finance & Management. Part of this research was conducted while I was at Heidelberg University and at Yale. I am grateful for the hospitality of both institutions. Contact: m.voigt@fs.de.

1 Introduction

Understanding belief formation is key to understanding asset prices. In any equilibrium, the price of an asset reflects investor beliefs about future dividends and prices, discounted at the required return. The predominant modeling approach for beliefs—rational expectations—assumes that investors understand the temporal fluctuation of dividends and prices, but survey evidence differs markedly from rational expectations.[1] The positive correlation of subjective expectations between surveys of professionals and individuals (Greenwood and Shleifer 2014) and between asset classes (Nagel and Xu 2023) suggests a systematic source for the deviation of subjective beliefs from rational expectations, and evidence from economics and psychology highlights the central role of memory for belief formation and individual decision making. Although recent models that incorporate psychological theories of human memory account for micro-evidence, the effect of memory on subjectively expected and objectively realized asset prices is largely unexplored (Malmendier and Wachter 2021).

In this paper, I argue that the selective memory of past observations simultaneously explains important facts about belief formation, survey data, and realized asset prices. Consistent with evidence from psychology and economics, I model the Bayesian beliefs of a representative agent that is more likely to recall some observations than others and treats her recalled experiences as if they were all that ever occurred (naïvete).[2] Selective memory generates a persistent wedge between the agent's subjective beliefs and rational expectations, and I analyze the effect of this belief wedge on asset prices using an otherwise standard consumption-based asset pricing model in which the agent learns the parameters of the payoff process. The framework allows me to examine different forms of selective memory.

[1] See Adam and Nagel (2023) for a review.

[2] Established 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). Additionally, experiments in economics find that humans are (partially) unaware of their memory distortions (Enke et al. 2022, Gödker et al. 2022, Zimmermann 2020). Although my assumptions are consistent with this evidence, I do not attempt to model the encoding and retrieval process on a neural level, but focus on the effect of memory on beliefs, as is relevant for macro-finance.

I mostly focus on the implications of *similarity-weighted memory*—the selective recall of past observations that are similar to today's observation—on beliefs and asset prices. Recent evidence in economics and finance finds that similarity-weighted memory is a key mechanism of individual belief formation, and research in psychology describes similarity as a law of human memory (Kahana 2012, Bordalo et al. 2020b, Enke et al. 2022, Jiang et al. 2023). I show that similarity-weighted memory provides a microfoundation for subjective beliefs about fundamentals that are consistent with empirical results: The agent expects a high growth of the economy upon observing a high realization of the fundamental (procyclical fundamental expectations, Nagel and Xu 2022, Bordalo et al. 2023b), expectations overreact to news (Coibion and Gorodnichenko 2015, Bordalo et al. 2020a), and the subjective volatility of fundamentals, which is the agent's measure of aggregate risk, varies countercyclically (Lochstoer and Muir 2022).

Incorporating the agent's beliefs about fundamentals under similarity-weighted memory into a consumption-based asset pricing model explains empirically observed differences of subjectively expected and objectively realized returns regarding their cyclicality, predictability, and sensitivity to risk measures. Empirically, as well as in the model, subjective return expectations are procyclical (Amromin and Sharpe 2014, Greenwood and Shleifer 2014), are not predictable by aggregate valuation ratios, and are partially explained by subjective volatility (Nagel and Xu 2023). On the contrary, the objective risk premium is predictably countercyclical (Shiller 1981, Mehra and Prescott 1985, Campbell and Shiller 1988), and is not related to objective risk measures (Lettau and Ludvigson 2010). The model yields a realistically high objective risk-premium if the agent learns from a limited sample of past observations.

In the last section of the paper, I demonstrate the flexibility of my general framework by analyzing an agent who selectively recalls extreme observations as well as observations that are similar to today's realization, which is consistent with the higher memorability of extreme as well as recent events (Kahneman 2000) and with the long-lasting impact of extreme outcomes on individual beliefs (Malmendier and Nagel 2011, Malmendier and Shen 2018). An agent who overremembers extreme observations perceives the economy as very risky, and thus requires a high subjective risk-premium. The aforementioned effects continue to hold, because time-variation in the agent's beliefs is determined by implicit similarity-effect.

Incorporating selective memory into a canonical asset pricing model simultaneously explains stylized facts about subjective expectations, subjective volatility, and subjective as well as objective (realized) asset prices that were often treated separately. This paper highlights the role of selective memory as a belief formation mechanism that unifies empirical evidence on individual beliefs and aggregate asset prices.

Beliefs under selective memory. As a first step in my analysis, I show that selective memory systematically affects the posterior beliefs of the agent—even in the very long-term—and can microfound deviations from rational expectations while retaining Bayesian learning. Methodologically, I characterize the agent's subjective long-term beliefs under selective and stochastic memory as memory-weighted likelihood maximizers as in Fudenberg et al. (2023). The agent observes many draws from the fixed distribution of fundamentals. Without memory distortions, the agent recalls all observations and the histogram of recalled observations converges almost surely and uniformly to the true distribution. Under selective memory, the histogram instead 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, which are relevant for applications in finance, 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 realizations; while the agent's posterior volatility is higher (lower) than the true volatility if the agent is more (less) likely to recall extreme realizations. For tractability, I focus on the limit in which the agent has already observed infinitely many realizations of the fundamental (long-term beliefs) in most of this paper and relax this assumption in simulations.

Similarity-weighted memory. I use the characterization of the agent's long-term beliefs to incorporate similarity-weighted memory into an otherwise standard 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 mean of log endowment growth in each state from her recalled observations, which are distorted by similarity-weighted memory. Assets in the economy are (levered) claims on the aggregate endowment. I highlight the implications of similarity-weighted memory for the agent's beliefs about fundamentals, subjective, and objective returns.

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

The intuition for (i)—procyclicality and overreaction of the posterior mean—is as follows: Assume that today's endowment growth is high. The agent then overremembers similar high growth rates, and forms Bayesian beliefs based on the recalled growth rates. The agent's posterior mean is thus high upon observing today's high endowment growth and varies in the direction of today's endowment growth (procyclicality).[3] 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 thus more likely to be

[3]Moreover, the agent can only recall growth rates that she previously observed, such that her recalled experiences contain a kernel of truth. Because the mass of observations is centered at the true mean, the agent's posterior mean must be between the true mean and today's endowment growth. Proposition 2 establishes that the posterior mean is a convex combination of today's growth rate and the true mean.

above than below the fundamental mean. Moreover, the agent's posterior mean exceeds the fundamental mean if and only if today's endowment growth exceeds the fundamental mean. Consequently, the agent's posterior mean is more likely above than below the fundamental mean after an upward revision, implying a predictably negative forecast error.

The intuition for (ii)—countercyclical variation of subjective volatility—is more subtle. With two states, one capturing normal times and the other recessions, the volatility of log endowment growth depends on the difference between the mean log endowment growth in each state. Unconditionally, log endowment growth is more volatile if recessions are much worse than normal times, and less volatile if recessions are as good as normal times. I show that the agent's posterior mean about log endowment growth in a recession reacts more strongly to today's log endowment growth than the agent's posterior mean about log endowment growth in normal times. Therefore, the difference of the posterior means is small (large) if today's log endowment growth is high (low). The time-varying difference between the conditional posterior means about recessions and normal times implies that the agent perceives the economy as less volatile in normal times than in recessions.

In equilibrium, the agent's subjective beliefs about fundamentals affect subjectively expected as well as objectively realized returns, as all assets are (levered) claims to the aggregate endowment. Consistent with survey evidence, the expected return under similarity-weighted memory is procyclical. When today's log endowment growth is high, the agent becomes optimistic and expects high log endowment growth going forward. The expected return (and the risk-free rate) must then increase to induce investment in the risky asset.[4] The subjective risk premium—the difference between the subjectively expected return and the objectively realized risk-free rate—is lower than the objectively realized risk premium, mostly acyclical, and unrelated to aggregate valuation ratios. The agent's risk-aversion is constant, but the perceived riskiness of the economy (subjective volatility) is time-varying, which leads to

[4] As in Martin (2013), higher-order moments are generally non-zero if endowment growth is determined by a two-state Markov switching process and I consider higher-order effects on asset prices in the body of the paper. I focus on the first two moments here for simplicity.

time-variation of the subjective risk premium. However, I show that the variation of the subjective volatility is small under similarity-weighted memory,[5] such that the subjective risk premium is almost constant and acyclical. As the subjective risk-premium is almost constant, it is also not predictable by time-variation of aggregate valuation ratios, but it is related to the agent's subjective perception of risk. Overall, the agent's beliefs about fundamentals explain stylized empirical patterns of subjectively expected returns.

Next, I turn to the patterns of objective returns. The risk-free rate varies procyclically with the agent's fundamental expectation about endowment growth. Intuitively, the risk-free rate must be high if the agent expects high endowment growth to induce saving in the risk-free asset. Moreover, an econometrician 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. For example, if today's endowment growth is high, the agent becomes too optimistic about the fundamentals, pushes up the price of the risky asset too much, and next period's realization disappoints on average (Bordalo et al. 2023b). Objective excess returns therefore vary countercyclically and are predictable using aggregate valuation ratios. Moreover, because the agent updates her beliefs as if her recalled experiences are all that ever occurred, she perceives her (new) belief to be persistent, leading to volatile objective returns. The time-variation of objective returns is unrelated to objective risk or measures of risk-aversion as both are assumed constant.

I next calibrate the model to analyze the quantitative implications of similarity-weighted memory. First, I simulate the model assuming that the agent has observed infinitely many realizations of endowment growth as in the theoretical analysis. My simulations confirm the qualitative properties discussed above. However, the agent's beliefs vary only with contemporaneous endowment growth, which is too smooth to yield a realistically high risk premium (Mehra and Prescott 1985). Second, I consider the case in which the agent observed a realistic number of endowment growth realizations (30 years of data). With finitely many

[5] Proposition 4 shows that the agent's subjective volatility is bounded from above by $\sqrt{1.25}$ times the true volatility of endowment growth.

observations, the agent is subjectively uncertain about her posterior beliefs, and this uncertainty depends on the (stochastic) number of recalled endowment growth observations.[6] As in Collin-Dufresne et al. (2016), parameter uncertainty generates an additional source of risk for the agent and thus yields a realistically high risk premium while retaining the qualitative properties of beliefs and asset prices. The objective risk premium can even become negative if the agent is very optimistic, consistent with empirical findings showing negative excess returns in times of high sentiment (Greenwood and Hanson 2013, Cassella and Gulen 2018).

Peak-end memory. I briefly show that the framework proposed in this paper can be used to analyze further memory distortions. In Section 4 of this paper, I propose and analyze a memory function that captures the higher memorability of extreme observations as well as observations that are similar to the end of the current experience (today‘s observation), leading to a peak-end memory distortion (Kahneman 2000). The experience effects literature highlights the persistent influence of extreme experiences on risk taking (Malmendier and Nagel 2011), inflation expectations (Malmendier and Nagel 2016), managerial decisions (Malmendier et al. 2011), or real estate purchases (Happel et al. 2022). Relatedly, a literature in psychology argues that 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 typically more memorable than the beginning and middle of the experience (recency effect, Kahana 2012, Barberis 2018, Wachter and Kahana 2023).

Under the peak-end memory distortion, the agent's posterior variance of the economy is always higher than the fundamental variance. Being risk-averse, the agent thus requires a comparably high subjective risk premium. The time-variation of the agent's beliefs is determined by the higher memorability of observations similar to the end, such that the

[6] Parameter uncertainty is non-monotonic in realized endowment growth: If this period's endowment growth is extremely high or extremely low, the agent observed fewer similar observations in her limited sample, such that uncertainty is higher in extremely good and bad periods. Since the price of risk increases in parameter uncertainty (Collin-Dufresne et al. 2016), similarity-weighted memory with parameter uncertainty leads to a non-monotonic price of risk in the economy.

qualitative properties of beliefs and asset prices discussed before also hold under the peak-end memory distortion. The paper ends with a brief conclusion.

Related literature. This paper contributes to a growing literature that analyzes the importance of memory on belief formation and on decision making. Empirical and experimental work in economics and finance shows that investors' information sets are systematically affected by selective memory (Zimmermann 2020, Charles 2021, Charles 2022, Enke et al. 2022, Gödker et al. 2022, Goetzmann et al. 2022, Graeber et al. 2022, Burro et al. 2023, Jiang et al. 2023). In line with a large literature in psychology (Tulving and Schacter 1990, Schacter 2008, Kahana 2012), a common finding is that recall of past observations is affected by the similarity of the past observation and the current context.[7] Using a representative survey of individual investors, Jiang et al. (2023) identify similarity-based recall as a key mechanism in the formation of investor beliefs. Based on evidence from psychology and economics, a theoretical literature analyzed the effect of memory mostly on individual decision making (Gilboa and Schmeidler 1995, Mullainathan 2002, Azeredo da Silveira and Woodford 2019, Bodoh-Creed 2020, Bordalo et al. 2023a, Fudenberg et al. 2023, Nagel and Xu 2022, Wachter and Kahana 2023). I contribute to this literature by analyzing the effect of selective memory—with a focus on similarity-weighted memory—on asset prices in an otherwise standard general equilibrium with a representative agent.

My paper also contributes to the growing theoretical literature on subjective beliefs in asset pricing (for an overview, see Adam and Nagel 2023). A common empirical finding is that individual investor expectations are procyclical: Investors expect asset prices to rise further after high returns and to continue to fall after low returns (Vissing-Jorgensen 2004, Bacchetta et al. 2009, Amromin and Sharpe 2014, Greenwood and Shleifer 2014, Kuchler and Zafar 2019, Da et al. 2021). Theoretical research modeled procyclical expectations by assuming

[7] The memory literature identified three main regularities of selective recall: Similarity (a higher likelihood of recalling observations that are similar to today's context), recency (a higher likelihood of recalling recent rather than past observations), and contiguity (a higher likelihood of recalling observations that co-occur temporally). Recency is closely related to extrapolative expectations (Nagel and Xu 2022) is also incorporated in my setup of the peak-end rule (Kahneman 2000).

that the agent overweights recent observations when forming her beliefs (over-extrapolation, see Barberis et al. 2015, Adam et al. 2017, Barberis et al. 2018, Jin and Sui 2022), diagnostic expectations (Bordalo et al. 2018, Bordalo et al. 2019), partial-equilibrium thinking (Bastianello and Fontanier 2022), or overlapping generations that learn from personal experiences (Ehling et al. 2018, Malmendier et al. 2020). Li and Liu (2023) show theoretically that procyclical fundamental expectations, but not procyclical return expectations, lead to a volatile equity premium; and Bordalo et al. (2023b) and Nagel and Xu (2022) empirically confirm that procyclical fundamental expectations explain the predictably countercyclical and high equity premium. I show that procyclical (fundamental) expectations arise naturally under similarity-weighted memory, since the agent is more likely to recall past observations that are similar to the current realization of endowment growth. As a consequence of these procyclical fundamental expectations, the model also generates procyclical expected returns and a positive correlation between the past return and the current expected return. In addition to procyclical expectations (first moment), my model also generates variation in the subjective volatility (second moment), consistent with survey evidence (Lochstoer and Muir 2022, Nagel and Xu 2023). Similarity-weighted memory parsimoniously generates empirically plausible patterns of subjective beliefs, and predicts a time-variation in the subjective risk premium that is explained by time-variation in the subjective volatility.

Based on evidence that lifetime experiences shape macroeconomic expectations (Malmendier and Nagel 2011, Malmendier and Nagel 2016, Malmendier and Wachter 2022, Happel et al. 2022), Nagel and Xu (2022) theoretically analyze an economy in which a representative agent learns with fading memory about mean endowment growth, leading to procyclical fundamental expectations. In their model, the objective risk premium varies countercyclically due to the predictability of forecast errors, whereas the subjective risk premium is constant because the agent knows the constant fundamental volatility and always recalls the same number of observations (constant parameter uncertainty). In my model, I also find a countercyclical objective risk premium due to the predictability of forecast errors. In addition, my model leads to time-variation in the subjective risk premium of the economy.[8] Psychologically, selective (similarity-weighted) memory as considered in this paper differs from fading memory: Consistent with the evidence on the long-term effect of personal experiences on economic decisions and the psychology literature (see Wachter and Kahana 2023), experiences have a long-lasting effect on the agent's beliefs and are never truly forgotten under selective memory. Under fading memory, instead, the effect of past experiences on the agent's beliefs gradually fades with the passage of time, and past experiences will eventually be forgotten.[9]

My model also builds on previous work that analyzes the asset-pricing implications of learning (Timmermann 1993, Lewellen and Shanken 2002, Weitzman 2007). Most closely related, Collin-Dufresne et al. (2016) analyze the asset pricing implications of parameter learning without memory distortions. In their model, a representative agent with Epstein-Zin preferences prices the parameter uncertainty that emerges from gradual Bayesian learning. I abstract from parameter uncertainty in my main analysis for tractability. Using the methods developed in Collin-Dufresne et al. (2016) and Johnson (2007),[10] I numerically show that parameter uncertainty emerging from selective memory leads to a realistically high risk premium.

My work applies results from the statistics and economics literature on misspecified learning (Berk 1966, Esponda and Pouzo 2016, Molavi 2019, Heidhues et al. 2021, Molavi et al. 2023) to asset pricing. Closely related, Fudenberg et al. (2023) propose the concept

[8]Most closely related, I confirm in simulations that the number of recalled observations is time-varying under selective memory, leading to time-variation in parameter uncertainty which is absent in Nagel and Xu (2022). Moreover, I obtain a countercyclical variation of the perceived volatility of the economy in the Markov-switching framework because of time-variation in the difference between the posterior means. An agent with fading memory would also experience time-variation in the difference between posterior means due to time-variation in the recalled experiences from each state, but this time-variation is not systematically related to macroeconomic conditions.

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

[10]I thank the authors of both papers for sharing code and details on the implementation of the algorithm with me.

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

2 Asset prices under a general selective memory distortion

In this section, I characterize the agent's long-term beliefs under selective memory and describe the asset pricing framework used throughout this paper. Selective memory is my only departure from a rational expectations model, and all asset pricing effects are driven by the agent's subjective long-term beliefs.

I specify the learning environment and model of selective memory in Section 2.1. The agent has observed infinitely many draws from a fixed but unknown distribution. The agent's recall of past observations is distorted by a memory function that specifies the probability of recalling a given past observation. Conditional on her recalled observations, the agent forms Bayesian beliefs about the unknown distribution. I simplify the exposition by focusing on discrete distributions as in Fudenberg et al. (2023) and provide an extension to continuous distributions in Online Appendix OA.1. My Proposition 1 in Section 2.2 highlights the effect of selective memory on the agent's subjective beliefs for normal distributions, as is relevant for applications in finance. Similarity-weighted memory (Section 3) and the peak-end rule (Section 4) are special cases of the treatment below.

Furthermore, I describe the asset pricing framework in Section 2.3. I consider a representative agent endowment economy with Epstein and Zin (1989)-preferences. The assets are (levered) claims to the endowment, and the distribution of endowment growth is fixed, such that I use the canonical model by Martin (2013) which nests the standard consumption-based asset pricing model (Mehra and Prescott 1985). Readers who are interested in the applications may prefer to go directly to the specific section.

2.1 Learning framework

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

Economy. I study a representative agent endowment economy in discrete time. In every period t, the agent observes the state of the world s_t and log endowment growth $g_t = \log C_t/C_{t-1}$. The assets in the economy are levered claims on the endowment stream, and I will introduce them more carefully in Section 2.3. For now, assume that the state $s_t = s$ is drawn from a finite set $S \subseteq \mathbb{N}$ according to the fixed, i.i.d. and full support distribution $\Xi \in \Delta(S)$. The state $s_t = s$ induces a fixed and i.i.d. distribution $q_s^* \in \Delta(G)$ over the finite set of possible endowment growth realizations G, that is $q_s^*(g) = Pr\left(g_t = g | s_t = s\right)$.[11] I assume that q_s^* belongs to the family of parametric distributions, $q_s^* \in \{q_\theta : \theta \in \Theta\}$, $\forall s \in S$, with $\Theta \subseteq \mathbb{R}^k, k \in \mathbb{N}$, closed and convex.

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

Memory. The agent observes an infinite history of log endowment growth and state re-

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

alizations, $H_t = \{(g_\tau, s_\tau)\}_{\tau=-\infty}^{t}$, where I call the tuple (g_τ, s_τ) an *experience*. $t_s = \sum_\tau \mathbb{1}_{\{s_\tau = s\}}$ denotes the number of experiences with $s_\tau = s, \tau \leq t$. In any period t, the agent recalls a subset of past experiences. She always 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 (g_t, s_t). The *recalled periods* r_t are a subset of $\{-\infty, ..., t\}$ and the *recalled history* $H_t^R \subseteq H_t$ is the collection of recalled experiences $\{(g_\tau, s_\tau)\}_{\tau \in r_t}$ with $|H_t^R|$ past experiences. Similarly, $r_{t,s}$ denotes the recalled periods in state s and the *recalled history of state* s $H_{s,t}^R \subseteq H_t^R$ is the collection of recalled experiences with $s_\tau = s$.

Beliefs. The agent forms Bayesian beliefs as if her recalled history H_t^R is all that occurred (naïvety).[12] Her posterior belief in period t is

$$b_t(C|H_t^R) = \frac{\int_{q\in C} \prod_{\tau \in r_t} q_{s_\tau}(g_\tau)\, db_0(q)}{\int_{q\in Q} \prod_{\tau \in r_t} q_{s_\tau}(g_\tau)\, db_0(q)}\ \forall C \subseteq Q.(\textbf{??}) \tag{2.1}$$

2.2 Subjective long-term beliefs

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

$$\mathrm{LM}(g_t, s_t) = \operatorname*{argmax}_{q\in Q} \Big(\sum_{s\in S} \psi(s) \sum_{g \in G} m_{(g_t,s_t)}(g,s)\, q_s^*(g)\ \log q_s(g) \Big). \tag{2.2}$$

The memory-weighted likelihood maximizer is the element of the agent's prior support that maximizes the likelihood of the recalled history H_t^R, and Fudenberg et al. (2023) show

[12] 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 distortions. If the agent correctly perceives her memory selectivity, she will perfectly undo any memory distortions and learn q^*. Alternatively, if she believes that her recalled experiences are representative for the experiences she does not recall, then her belief b_t is not affected by partial naïvety. An analysis of intermediate assumptions about the agent's perception of her memory selectivity is provided in Fudenberg et al. (2023).

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, such that 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,[13] which implies that the agent's beliefs concentrate on the memory-weighted likelihood maximizer.[14]

I now illustrate the effect of selective memory on the agent's subjective beliefs if log endowment growth is normally distributed conditional on the state, as is relevant for asset pricing applications. Assume that log endowment growth is drawn from a normal distribution, $q_s^* \in \{q_\theta = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(g-\mu)^2}{2\,\sigma^2}\right) | \theta = (\mu;\sigma^2) \in \Theta, \mu \in \mathbb{R}, \sigma^2 \in \mathbb{R}_{\geq 0}, g \in \mathbb{R}\} =: \Theta_{\mathcal{N}}$, where Θ is closed and convex.[15] I associate each q_s^* with a parameter vector $\theta_s = (\mu_s, \sigma_s^2)$ whenever no confusion arises. Assumption 1 holds for the remainder of this paper, and Proposition 1 shows how the agent's state-wise posterior belief depends on selective memory. The agent observes the state realizations and thus performs state-wise inference.

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

Proposition 1 For each state $s \in S$, the agent's belief $b_{s,t}$ is almost surely given by the

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

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

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

unique normal distribution with $\hat{\theta}_{s,t} := (\hat{\mu}_{s,t}, \hat{\sigma}^2_{s,t})$, and

$$\hat{\mu}_{s,t} = \mu_s + \underbrace{\mathbb{E}\Big[\frac{t_s}{|H^R_{s,t}|}\Big]}_{\text{Forgetfulness}} \cdot \underbrace{\mathrm{Cov}\Big[g, \mathbb{1}_{\{g \in H^R_{s,t}\}}\Big]}_{\text{Selectivity}}, \text{ and} \tag{2.3}$$

$$\hat{\sigma}^2_{s,t} = \sigma^2_s + \big(\hat{\mu}_{s,t} - \mu_s\big)^2 + \underbrace{\mathbb{E}\Big[\frac{t_s}{|H^R_{s,t}|}\Big]}_{\text{Forgetfulness}} \cdot \underbrace{\mathrm{Cov}\Big[(g - \hat{\mu}_{s,t})^2, \mathbb{1}_{\{g \in H^R_{s,t}\}}\Big]}_{\text{Selectivity}}, \tag{2.4}$$

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

Equation 2.3 shows that the agent's posterior mean of log endowment growth in state s depends on two elements: (i) the true fundamental mean μ_s, and (ii) an adjustment term that arises from selective memory. The "forgetfulness" term in the adjustment, $\mathbb{E}\Big[\frac{t_s}{|H^R_{s,t}|}\Big]$, is the expected size of the realized history relative to the recalled history of state s. If the agent recalls almost all past observations, $m_{(g_t,s_t)}(g_\tau, s_\tau) \approx 1$, the recalled history will be as long as the realized history, $|H^R_{s,t}| \approx t_s$. The recalled history will be "shorter" than the realized history if the agent, instead, barely recalls past growth rates.[16] The second term in the adjustment, $\mathrm{Cov}\Big[g, \mathbb{1}_{\{g \in H^R_{s,t}\}}\Big]$, captures the selectivity. The posterior mean will be higher than the true mean if the agent is more likely to recall high growth rates from state s, as measured by the covariance between g and the propensity of recalling $g_\tau = g$.[17] On the contrary, if the agent is as likely to recall high as low log endowment growth rates, such that $\mathrm{Cov}\Big[g, \mathbb{1}_{\{g \in H^R_{s,t}\}}\Big] \approx 0$, the agent almost surely learns μ_s.

The agent's posterior variance of log endowment growth in state s (Equation 2.4) is anchored at the true underlying variance σ^2_s and learned by an agent without memory dis-

[16] Using Jensen's Inequality, we know that for a positive random variable X, it is $\mathbb{E}\left(\frac{1}{X}\right) \geq \frac{1}{\mathbb{E}(X)}$. Therefore, we must have $\mathbb{E}\Big[\frac{t}{|H^R_t|}\Big] \geq \frac{t}{\mathbb{E}\big[|H^R_t|\big]}$.

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

tortions. The second term in Equation 2.4, $(\hat{\mu}_{s,t}-\mu_s)^2$, is the usual adjustment for the usage of a biased mean estimate. The last term in Equation 2.4, $\mathbb{E}\left[\frac{t_s}{|H^R_{s,t}|}\right] \cdot \mathrm{Cov}\left[(g-\hat{\mu}_{s,t})^2, \mathbb{1}_{\{g \in H^R_{s,t}\}}\right]$, captures the direct effect of selective memory on the agent's posterior variance. Selective memory increases the posterior variance whenever the agent is more likely to recall more spread-out log endowment growth rates; while selectivity decreases the posterior variance if the agent tends to recall endowment growth rates that are close to the posterior mean.

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

2.3 Asset pricing framework

In this subsection, I incorporate the agent's subjective beliefs that arise from selective memory into a standard consumption-based asset pricing model (Lucas 1978, Mehra and Prescott 1985). For analytical tractability, I follow the framework of Martin (2013).[18]

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

$$U_t = \left\{(1-\beta)\, C_t^{\frac{1-\gamma}{\eta}} + \beta \left(\tilde{\mathbb{E}}_t \left[U_{t+1}^{1-\gamma}\right]\right)^{\frac{1}{\eta}}\right\}^{\frac{\eta}{1-\gamma}}, \tag{2.5}$$

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

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

conditional on her recalled history H_t^R.

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

Consider an asset that pays a dividend stream $\{D_{t+k}\}_{k\geq 0}$ with $D_{t+k} = C_{t+k}^{\lambda}$ 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 $\tilde{er}_t = \log\left(\tilde{\mathbb{E}}_t\ R_{t+1}\right)$. Similarly, the *log risk-free rate* is the log (subjective) expected return on the riskless bond,[19] and the *subjective risk premium* on the λ-asset is the difference between the log subjective expected return and the log risk-free rate. With Epstein and Zin (1989)-preferences, Martin (2013) shows that[20]

$$r_t^f = -\log(\beta) - \mathcal{K}_t(-\gamma) + \left(1 - \frac{1}{\eta}\right)\mathcal{K}_t(1-\gamma), \tag{2.6}$$

$$dp_t = -\log(\beta) - \mathcal{K}_t(\lambda-\gamma) + \left(1 - \frac{1}{\eta}\right)\mathcal{K}_t(1-\gamma), \tag{2.7}$$

$$\tilde{er}_t = dp_t + \mathcal{K}_t(\lambda), \tag{2.8}$$

$$\tilde{rp}_t = \tilde{er}_t - r_t^f = \mathcal{K}_t(\lambda) + \mathcal{K}_t(-\gamma) - \mathcal{K}_t(\lambda-\gamma), \tag{2.9}$$

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

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

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

under the agent's subjective beliefs are defined as

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

Both the moment-generating function and the cumulant-generating function provide expressions for the moments of log endowment growth under the agent's posterior belief b_t. Expanding the cumulant-generating function $\mathcal{K}_t(k)$ as a power series yields

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

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

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

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

Both the risk-free rate and the subjective expected return are increasing in the posterior mean of the agent, $\hat{\mu}_t$, while the dividend-price ratio is decreasing in $\hat{\mu}_t$ if $\lambda > \gamma$. Intuitively, if the agent becomes more optimistic (higher $\hat{\mu}_t$), she consumes more today. The risk-free rate and subjective expected return must then increase to induce saving/investment. The dividend yield dp_t decreases in $\hat{\mu}_t$ because the price—which reflects the discounted sum of all future dividends—increases in $\hat{\mu}_t$ and leverage λ exceeds the agent's discounting due to γ.[21] 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 $\gamma\,\hat{\mu}_t$.

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

To obtain further insights into the effect of selective memory on asset prices, I impose further assumptions on the agent's memory and the log endowment growth process in the next Section.

[21] 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.

3 Asset prices under a similarity-weighted memory distortion

A central principle of memory is that humans are more likely to recall past experiences that are similar to the current experience (Mullainathan 2002, Kahana 2012, Bordalo et al. 2023b, Enke et al. 2022, Jiang et al. 2023, Wachter and Kahana 2023). Although similarity can be measured with respect to different features of the current experience, I focus on similarity with respect to this period's endowment growth, which is the most ubiquitous feature in an endowment economy. Consider the following example: The economy is either in a normal state or in a recession. Let endowment growth always be high if the economy is in a normal state, while endowment growth can be high or low if the economy is in a recession (recessions have higher fundamental uncertainty). The agent observes this period's endowment growth and recalls past experiences with similar endowment growth, regardless of the state that generated the past endowment growth. Since endowment growth is always high in normal times, the agent can always only recall high endowment growth from normal times and her posterior belief about normal times does not react to current endowment growth. In contrast, her posterior belief about recessions reacts to the cue provided by this period's endowment growth: If this period's endowment growth is high (low), the agent recalls more high (low) endowment growth experiences from the recession than fundamentally occurred, and her posterior mean will be high (low). Therefore, the unconditional posterior mean of the agent positively covaries with this period's endowment growth, and the difference between the posterior about the mean endowment growth in normal times and in a recession varies over time.[22] In line with this motivating example, Enke et al. (2022) conducted a series of laboratory experiments and document that similarity-weighted memory leads to an

[22]The example here is an extension of the thirsty traveler example in Bordalo et al. (2020b). Consider a thirsty traveler who retrieves water prices from her memory. At the airport, water prices are always high and thus, she can only retrieve high water prices from her airport experiences regardless of the current cue. In contrast, water prices downtown are sometimes low (at the corner store) and sometimes high (at a luxury hotel), such that similarity to the current context systematically affects her norm for water prices downtown.

overreaction of experimental market prices to news. Jiang et al. (2023) provide evidence of similarity-weighted memory as a key mechanism of belief formation in financial markets.

In this section, I assume that the agent's memory is distorted by a similarity-weighted memory function and impose further structure on log endowment growth, which I describe in Section 3.1. The assumptions and results from Section 2 continue to hold. Using Proposition 1, I analyze the agent's long-term beliefs under similarity-weighted memory in Section 3.2, and discuss the asset pricing implications in Section 3.3. Finally, I use standard data to estimate the parameters of log endowment growth under my structural assumptions and simulate asset prices in Section 3.4. I show that similarity-weighted memory qualitatively explains survey evidence and objective returns, and can generate a realistically high risk premium.

3.1 Structural assumptions

The assumptions in Section 2.1 continue to hold, but I impose additional structure on the economy. Let $S = \{1, 2\}$. $s_t = s$ follows a two-state observable Markov chain with constant transition matrix Π. The elements of Π are $\pi_{ij} = \Pr[s_t = j | s_{t-1} = i]$, and I restrict $\pi_{11} = \pi_{21} =: \pi_1$ and $\pi_{12} = \pi_{22} = 1 - \pi_1 =: \pi_2$ to ensure that the process is i.i.d. Conditional on state $s_t = s$, log endowment growth is normally distributed,

$$g_t = \mu_{s_t} + \sigma_{s_t} \epsilon_t, \quad \epsilon_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, 1). \tag{3.1}$$

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

Markov-switching models have been widely used to model the aggregate endowment dynamics due to their flexibility and tractability (Mehra and Prescott 1985, Rietz 1988, Barro 2006, Johannes et al. 2016). I use the two-state structure to analyze how similarity-weighted memory affects the agent's perception of recessions in comparison to normal times.

All results that relate to the agent's posterior mean hold if endowment growth is log-normally distributed, but the perceived riskiness of the economy is then constant (see Online Appendix OA.2.1).

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

$$m^{\text{sim}}_{(g_t, s_t)}\left(g_\tau, s_\tau\right) = \exp\left[-\frac{\left(g_\tau - g_t\right)^2}{2\,\kappa}\right], \tag{3.2}$$

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

3.2 Long-term beliefs

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

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

Proposition 2 Under similarity-weighted memory as in Equation 3.2, almost surely,

$$\hat{\mu}_{s,t} = \frac{\kappa}{\kappa + \sigma_s^2}\,\mu_s + \frac{\sigma_s^2}{\kappa + \sigma_s^2}\,g_t = \left(1 - \alpha_s\right)\mu_s + \alpha_s\,g_t, \text{ and} \tag{3.3}$$

$$\hat{\sigma}_{s,t}^2 = \left(1 - \alpha_s\right)\sigma_s^2, \tag{3.4}$$

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

Equation 3.3 shows that the state-dependent posterior mean of the agent is a convex combination of the true state-dependent mean μ_s and this period's log endowment growth g_t. If today's endowment growth is high, the agent is more likely to recall past experiences with a high endowment growth than with a low endowment growth due to similarity. The agent will therefore be more (less) optimistic if this period's endowment growth is high (low). Similarity-weighted memory distortions provide a microfoundation for extrapolative beliefs in that the agent's posterior is formed as if she overweights contemporaneous endowment growth when forming her beliefs (for an overview on extrapolative beliefs, see Barberis 2018).

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

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

The agent's state-dependent posterior variance, as given by Equation 3.4, is independent of today's endowment growth and always smaller than the fundamental state-dependent variance σ_s^2 since $\alpha_s \in (0, 1)$. The agent's similarity-weighted memory distortion symmetrically overweights observations that are close to today's consumption growth, while she tends to forget experiences that are less similar to today's experience. Therefore, the scale of the posterior distribution is smaller than the scale of the fundamental distribution.[23] Going forward, I focus on the case in which the agent knows the state-dependent variances, σ_1^2 and σ_2^2, but learns about the mean endowment growth μ_1 and μ_2.[24]

Subsequently, I focus on the agent's unconditional time-t beliefs about log endowment growth. Note that unconditionally, log endowment growth generally does not follow a normal distribution. A characterization of the agent's unconditional posterior distribution of log endowment growth can be derived from the cumulant-generating functions under the agent's posterior beliefs introduced in Section 2.3:

$$\mathcal{K}_t(k) = \log\left[\mathcal{M}_t(k)\right] = \log\left[\pi_1\, e^{k\,\hat{\mu}_{1,t}+\frac{1}{2}\,k^2\,\sigma_1^2} + \pi_2\, e^{k\,\hat{\mu}_{2,t}+\frac{1}{2}\,k^2\,\sigma_2^2}\right]. \tag{3.5}$$

We can find the n'th (non-central) moment by taking the n'th derivative of the moment-generating function with respect to k and evaluating the derivative at $k = 0$. The unconditional expected log endowment growth under the agent's posterior belief is[25,26]

$$\tilde{\mathbb{E}}_{\mathrm{t}}\left(g_{t+1}\right) = \pi_1\,\hat{\mu}_{1,t} + \pi_2\,\hat{\mu}_{2,t}. \tag{3.6}$$

Equation 3.6 shows that agent's expected log endowment growth is the probability-weighted average of the state-wise posterior means. The expected log endowment growth is thus increasing in this period's endowment growth (procyclical). The sensitivity of the expected log endowment growth to the contemporaneous endowment growth depends on a weighted

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

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

[25] Figure D.1 in the appendix shows a plot of the dependence of the first four moments on the contemporaneous endowment growth g_t. However, since the effect of higher-order moments on asset prices is small under similarity-weighted memory, I do not discuss higher-order moments in the main body of the paper.

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

average of the state-wise variances and on the scrutiny κ. The following Proposition 3 characterizes the agent's state-wise and unconditional posterior mean.

Proposition 3 The *average state-dependent posterior mean conditional on the current state is*

$$\mathbb{E}\left(\hat{\mu}_{1,t+1} | s_{t+1} = 1\right) = \mu_1 \tag{3.7}$$

$$\mathbb{E}\left(\hat{\mu}_{1,t+1} | s_{t+1} = 2\right) = \mu_1 + \alpha_1 \left(\mu_2 - \mu_1\right), \tag{3.8}$$

and the average state-dependent posterior mean is

$$\mathbb{E}\left(\hat{\mu}_{1,t+1}\right) = \mu_1 + \alpha_1 \, \pi_2 \left(\mu_2 - \mu_1\right). \tag{3.9}$$

Equations 3.7—3.9 hold for $\hat{\mu}_{2,t+1}$ with the respective change of indices. Moreover, the average unconditional posterior mean of log endowment growth is

$$\mathbb{E}\left(\pi_1 \, \hat{\mu}_{1,t+1} + \pi_2 \, \hat{\mu}_{2,t+1}\right) = \pi_1 \, \mu_1 + \pi_2 \, \mu_2 + \pi_1 \, \pi_2 \left(\mu_2 - \mu_1\right) \left(\alpha_1 - \alpha_2\right). \tag{3.10}$$

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

The average unconditional posterior mean given in Equation 3.10 is biased upward, since $\mu_1 > \mu_2$ and $\alpha_1 < \alpha_2$ by assumption. In normal times, $s_t = 1$, endowment growth tends to be high and the posterior mean of state 2 is biased upwards. Similarly, the posterior mean of state 1 is biased downwards in a recession, $s_t = 2$, but the downward bias of the posterior mean of state 1 is smaller than the upward bias of the posterior mean of state 2 because $\alpha_1 < \alpha_2$. The net effect is an upward bias of the unconditional posterior mean, such that the agent is, on average, too optimistic about endowment growth.[27]

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

$$\mathrm{Var}_t\left(g_{t+1}\right) = \pi_1\, \sigma_1^2 + \pi_2\, \sigma_2^2 + \pi_1\, \pi_2\, \left(\hat{\mu}_{1,t} - \hat{\mu}_{2,t}\right)^2. \tag{3.11}$$

Equation 3.11 highlights that the perceived riskiness of the economy, $\mathrm{Var}_t(g_{t+1})$, depends not only on the state-wise variances (σ_1^2 and σ_2^2), but also on the squared distance between the state-wise posterior means $(\hat{\mu}_{1,t} - \hat{\mu}_{2,t})^2$. Intuitively, the agent perceives the economy as more risky if recessions are severe compared to normal times ($\hat{\mu}_{2,t} << \hat{\mu}_{1,t}$). Proposition 4 characterizes the average posterior variance.

Proposition 4 The average posterior variance of the agent is

$$\begin{aligned}\mathbb{E}\left[\mathrm{Var}_t\left(g_{t+1}\right)\right] = &\left(\pi_1\, \sigma_1^2 + \pi_2\, \sigma_2^2\right) \left[1 + \pi_1\, \pi_2\, \left(\alpha_1 - \alpha_2\right)^2\right] \\ &+ \left(\mu_1 - \mu_2\right)^2\, \pi_1\, \pi_2\, \left[\pi_1\, \pi_2 \left(\alpha_1 - \alpha_2\right)^2 + \left[1 - \left(\alpha_1\, \pi_2 + \alpha_2\, \pi_1\right)\right]^2\right],\end{aligned} \tag{3.12}$$

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

$$\frac{\left(\alpha_1 - \alpha_2\right)^2\, \left(\pi_1\, \sigma_1^2 + \pi_2\, \sigma_2^2\right)}{2\, \left(\pi_2\, \alpha_1 + \pi_1\, \alpha_2\right) - \left(\pi_2\, \alpha_1^2 + \pi_1\, \alpha_2^2\right)} \geq \left(\mu_1 - \mu_2\right)^2, \tag{3.13}$$

and bounded by

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

[27] It follows that the posterior mean under similarity-weighted memory is unbiased if endowment growth is log-normally distributed. The effects of similarity-weighted memory discussed in Proposition 4 do not arise if endowment growth is log-normally distributed.

Condition 3.13 allows us to characterize situations in which the average perceived riskiness of the economy exceeds the fundamental riskiness. The left-hand side of condition 3.13 is always positive, so that Condition 3.13 holds for $(\mu_1 - \mu_2) \to 0$. Similarity-weighted memory systematically increase the perceived variance if the mean in both states is approximately equal because the sensitivity of the posterior means to the contemporaneous endowment growth differs. Additionally, the left-hand side of Condition 3.13 increases in the difference of the state-dependent variances σ_1^2 and σ_2^2. The sensitivity of the posterior mean of state s depends on the state-dependent variance σ_s^2. In situations in which $\sigma_2^2 >> \sigma_1^2$, the posterior mean of state 2 reacts stronger to this period's endowment growth g_t than the posterior mean of state 1, such that the squared difference of the posterior means is, on average, higher if $\sigma_2^2 >> \sigma_1^2$ than if $\sigma_2^2 \approx \sigma_1^2$.

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

Forecast rationality. I now compare the expectations of an agent under similarity-weighted memory to the expectations of an agent without memory distortions.[28] Intuitively, realizations of a stochastic process should move one-for-one with a rational forecast, such that forecast rationality implies $a_{MZ} = 0$ and $\beta_{MZ} = 1$ in the following Mincer and Zarnowitz (1969)-regression:

$$g_{t+h} = a_{MZ} + \beta_{MZ}\, \tilde{\mathbb{E}}_{\mathrm{t}}\left(g_{t+h}\right) + u_{t+h} \tag{3.14}$$

Another property of rational forecasts is that they should neither overreact nor underreact

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

to new information, and Coibion and Gorodnichenko (2015) propose the following regression to test over- or underreaction of forecasts:

$$g_{t+h} - \tilde{\mathbb{E}}_{\mathrm{t}}\left(g_{t+h}\right) = a_{CG} + \beta_{CG}\left[\tilde{\mathbb{E}}_{\mathrm{t}}\left(g_{t+h}\right) - \mathbb{E}_{t-1}\left(g_{t+h}\right)\right] + u_{t+h} \tag{3.15}$$

Forecast rationality implies $a_{CG} = 0$ and $\beta_{CG} = 0$ in Equation 3.15, because the forecast revision $\tilde{\mathbb{E}}_{\mathrm{t}}\left(g_{t+h}\right) - \mathbb{E}_{t-1}\left(g_{t+h}\right)$ is known to the agent at time t and should thus not predict forecast errors (otherwise, the agent would adjust her forecast). If $\beta_{CG} < 0$, the forecaster overreacts since the revision is too strong on average, while $\beta_{CG} > 0$ captures an underreaction of the forecast. The following Proposition 5 shows that we can reject forecast rationality in Mincer and Zarnowitz (1969)-regressions and find overreaction in Coibion and Gorodnichenko (2015)-regressions for an agent with similarity-weighted memory.

Proposition 5 Under similarity-weighted memory, we can reject rationality of the agent's forecast measured by Mincer and Zarnowitz (1969)-regressions as in Equation 3.14, since

$$\beta_{MZ} = 0 < 1, \qquad \text{and} \qquad a_{MZ} = \pi_1\,\mu_1 + \pi_2\,\mu_2 \neq 0. \tag{3.16}$$

The long-term beliefs of an agent with similarity-weighted memory overreact as measured by Coibion and Gorodnichenko (2015)-regressions in Equation 3.15, since

$$\beta_{CG} = -\frac{1}{2} < 0, \qquad \text{and} \qquad a_{GC} = \pi_1\,\pi_2\left(\mu_2 - \mu_1\right)\,\left(\alpha_1 - \alpha_2\right). \tag{3.17}$$

The results summarized in Proposition 5 show that the agent's forecast is uninformative for the realization of log endowment growth ($\beta_{MZ} = 0$). The best forecast of future realizations of log endowment growth is the long-term mean $\pi_1\mu_1 + \pi_2\mu_2$, because log endowment growth is i.i.d. The agent's expectation, however, covaries with the current realization of endowment endowment growth g_t due to her similarity-weighted memory, but the current realization of an i.i.d. process is not predictive for future realizations, yielding $\beta_{MZ} = 0$.

Similarity-weighted memory also leads to an overreaction of the agent's forecast ($\beta_{GC} < 0$). For simplicity, focus on $\hat{\mu}_{1,t} = (1-\alpha_1)\,\mu_1 + \alpha_1\, g_t$. The agent revises the posterior mean of state 1 upward if and only if tomorrow's log endowment growth exceeds the contemporaneous log endowment growth, $g_{t+1} > g_t$. Thus, conditional on upward revision, it is $\Pr(g_{t+1} \geq \mu_1 | g_{t+1} > g_t) > 0.5$, since g_{t+1} must exceed g_t. In addition, we also know that $\hat{\mu}_{1,t+1}$ exceeds the true mean μ_1 if and only if $g_{t+1} > \mu_1$. Consequently, the agent's posterior mean is more likely above than below the fundamental mean after an upward revision, implying a predictably negative forecast error.

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

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

$$\begin{aligned}\tilde{\mathcal{M}}(m) := \mathbb{E}\left[\mathcal{M}_{t+1}(m)\right] =& \pi_1\, \mathbb{E}\left[\, e^{m\,\alpha_1\, g_{t+1}}\right]\; e^{m\,(1-\alpha_1)\,\mu_1 + \frac{1}{2}\, m^2\, \sigma_1^2} + \\ & \pi_2\, \mathbb{E}\left[\, e^{m\,\alpha_2\, g_{t+1}}\right]\; e^{m\,(1-\alpha_2)\,\mu_2 + \frac{1}{2}\, m^2\, \sigma_2^2} \\ =& \pi_1\, e^{\mathcal{K}^*(m\,\alpha_1) + m\,(1-\alpha_1)\,\mu_1 + \frac{1}{2}\, m^2\, \sigma_1^2} \\ & + \pi_2\, e^{\mathcal{K}^*(m\,\alpha_2) + m\,(1-\alpha_2)\,\mu_2 + \frac{1}{2}\, m^2\, \sigma_2^2},\end{aligned} \tag{3.18}$$

where $\mathcal{K}^*(k) = \log\left(\pi_1\, e^{k\,\mu_1 + \frac{1}{2}\, k^2\, \sigma_1^2} + \pi_2\, e^{k\,\mu_2 + \frac{1}{2}\, k^2\, \sigma_2^2}\right)$ denotes the true cumulant-generating function of endowment growth. Equation 3.18 shows that the expected belief of an agent with similarity-weighted memory is constant over time as a result of the i.i.d.-structure of the economy. A constant expectation of the agent's posterior beliefs implies that belief revisions are predictable and mean-reverting. If this period's log endowment growth is very high, such that the relevant moments of the agent's posterior beliefs are inflated, then we expect to observe a downward revision of the agent's beliefs in the next period.

Summary. I find that the agent's posterior belief under similarity-weighted memory is time-varying, although the agent has access to infinite data and also forms beliefs using infinite data. The focus on long-run implications of similarity-weighted memory on the agent's beliefs allowed me to derive parsimonious expressions for the agent's posterior beliefs. Consistent with empirical evidence, I find that the agent's subjective beliefs are such that (i) the posterior mean varies procyclically and overreacts to new information; and that (ii) the agent's subjective volatility of fundamentals varies countercyclically. Moreover, belief revisions are predictable mean-reverting. The next section highlights the asset pricing consequences of the agent's subjective long-term beliefs under similarity-weighted memory.

3.3 Asset pricing implications

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

Subjective asset prices. Proposition 6 characterizes the equilibrium asset prices under similarity-weighted memory using the subjective cumulant-generating function in Equation 3.5 and the results from Section 2.3. I focus on power utility for simplicity and give the results for Epstein and Zin (1989)-preferences in Appendix B.5.

Proposition 6 Let us focus on power utility ($\psi = 1/\gamma$). Under similarity-weighted memory as in Equation 3.2 and an i.i.d. two-state Markov-switching process for log endowment growth, we have

$$r_t^f = -\log(\beta) - \log\left(\pi_1\, e^{-\gamma\,\hat{\mu}_{1,t}+\frac{1}{2}\gamma^2\,\sigma_1^2} + \pi_2\, e^{-\gamma\,\hat{\mu}_{2,t}+\frac{1}{2}\,\gamma^2\,\sigma_2^2}\right), \tag{3.19}$$

$$dp_t = -\log(\beta) - \log\left(\pi_1\, e^{(\lambda-\gamma)\,\hat{\mu}_{1,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_1^2} + \pi_2\, e^{(\lambda-\gamma)\,\hat{\mu}_{2,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_2^2}\right), \tag{3.20}$$

$$\tilde{er}_t = dp_t + \log\left(\pi_1\, e^{\lambda\,\hat{\mu}_{1,t}+\frac{1}{2}\,\lambda^2\,\sigma_1^2} + \pi_2\, e^{\lambda\,\hat{\mu}_{2,t}+\frac{1}{2}\,\lambda^2\,\sigma_2^2}\right), \tag{3.21}$$

$$\begin{aligned}\tilde{rp}_t =& \log\left(\pi_1\, e^{-\gamma\,\hat{\mu}_{1,t}+\frac{1}{2}\gamma^2\,\sigma_1^2} + \pi_2\, e^{-\gamma\,\hat{\mu}_{2,t}+\frac{1}{2}\,\gamma^2\,\sigma_2^2}\right) + \log\left(\pi_1\, e^{\lambda\,\hat{\mu}_{1,t}+\frac{1}{2}\,\lambda^2\,\sigma_1^2} + \pi_2\, e^{\lambda\,\hat{\mu}_{2,t}+\frac{1}{2}\,\lambda^2\,\sigma_2^2}\right) \\ &- \log\left(\pi_1\, e^{(\lambda-\gamma)\,\hat{\mu}_{1,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_1^2} + \pi_2\, e^{(\lambda-\gamma)\,\hat{\mu}_{2,t}+\frac{1}{2}\,(\lambda-\gamma)^2\,\sigma_2^2}\right).\end{aligned} \tag{3.22}$$

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

Second, the dividend-price ratio of the λ-asset dp_t decreases in the contemporaneous log endowment growth if $\lambda > \gamma$. As discussed in Section 2.3, the price of the asset increases in the agent's posterior mean if the leverage factor λ exceeds the curvature of the utility function as determined by the risk-aversion parameter γ.[29] A decreasing dividend-price ratio is consistent with substantial evidence showing that the price-dividend ratio (the reciprocal of dp_t) is positively correlated with subjective expectations of future growth (De La O and

[29] 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.

Myers 2021, Bordalo et al. 2023b).

Third, the subjectively expected return, which depends on the dividend-price ratio dp_t and on the subjectively expected payout growth $\mathcal{K}_t(\lambda) = \log\left(\pi_1\, e^{\lambda\,\hat{\mu}_{1,t}+\frac{1}{2}\,\lambda^2\,\sigma_1^2} + \pi_2\, e^{\lambda\,\hat{\mu}_{2,t}+\frac{1}{2}\,\lambda^2\,\sigma_2^2}\right)$, increases in the contemporaneous log endowment growth. The subjectively expected payout growth is increasing in the agent's posterior mean and thus also in the contemporaneous log endowment growth. For $\gamma > \lambda$, the dividend-price ratio also increases in contemporaneous log endowment growth, such that both components of the subjectively expected return increase in contemporaneous log endowment growth. For $\gamma < \lambda$, instead, the dividend-price ratio decreases in contemporaneous log endowment growth. However, $\lambda > \gamma > 0$ implies that the expected payout growth is more sensitive to the agent's state-wise posterior means $\hat{\mu}_{1,t}$ and $\hat{\mu}_{2,t}$ than the dividend-price ratio, such that the increase of the expected payout dominates the decreasing dp_t. Intuitively, the agent is more optimistic after observing a high contemporaneous log endowment growth, and the risky asset must deliver a high expected return to induce investment.[30] Consistent with survey evidence, the expected return of an agent with similarity-weighted memory is procyclical in that the agent subjectively expects a high return if contemporaneous log endowment growth is high.

Fourth, the subjective risk premium rp_t depends on the convexity of the cumulant-generating function over the interval $[-\gamma, \lambda]$, as highlighted by Martin (2013). Under similarity-weighted memory, however, the convexity of the cumulant-generating function changes over time with g_t. Consider the second-order approximation of the cumulant-generating function under the agent's subjective beliefs discussed in Section 2.3, $rp_t \approx \lambda\,\gamma\,\mathrm{Var}_t\left(g_{t+1}\right) = \lambda\,\gamma\,\left(\pi_1\,\sigma_1^2 + \pi_2\,\sigma_2^2 + \pi_1\,\pi_2\,\left(\hat{\mu}_{1,t} - \hat{\mu}_{2,t}\right)^2\right)$.[31] The agent knows the state-wise variances σ_1^2 and σ_2^2, but the wedge between the posterior mean of state 1 and 2 is time-

[30] Figure D.2 in the appendix shows that an increase in contemporaneous log endowment growth rotates the subjective cumulant-generating function. As the cumulant-generating function is convex (Martin 2013), the effect of an increase in the contemporaneous log endowment growth on the subjective expected return is not necessarily monotone, and Figure 3.1 below shows that the expected return is a convex function of g_t. The statements in the main text hold for a second-order approximation of the subjective cumulant-generating function.

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

varying with the contemporaneous log endowment growth, and the agent's posterior variance is a convex function of the contemporaneous log endowment growth. The subjective risk premium of the agent is, up to a second-order approximation of the subjective cumulant-generating function, proportional to the unconditional posterior variance, and thus a convex function of the contemporaneous log endowment growth. The agent requires a high risk premium in exceptionally good and bad times, while the risk premium is moderate in intermediate regions of log endowment growth.

Figure 3.1: Qualitative asset pricing implications of similarity-weighted memory

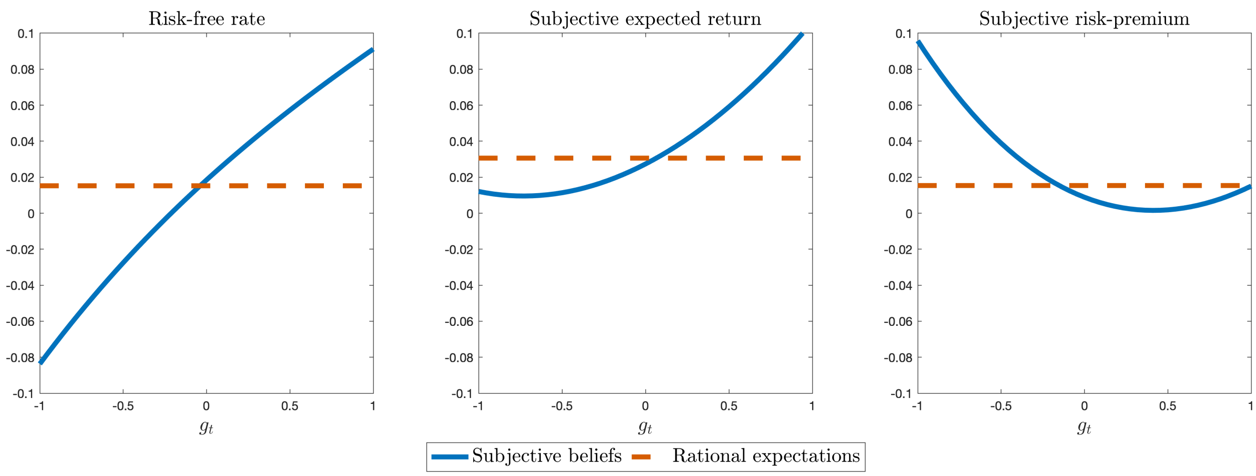

Figure 3.1 shows the equilibrium asset prices for for different values of the contemporaneous log endowment growth g_t. The left panel shows the risk-free rate, the middle panel the expected return, and the right panel the subjective risk premium under rational expectations (orange dashed line) and under similarity-weighted memory (blue solid line). The parameters of the log endowment growth process are $\mu_1 = 0.1, \mu_2 = 0.2, \sigma_1 = 0.1, \sigma_2 = 0.2, \pi_1 = 0.9, \lambda = 3$, and the preference parameters are $\beta = 0.98, \gamma = 4, \psi = 1.5, \kappa = 0.001$.

Figure 3.1 shows the risk-free rate, expected return, and subjective risk premium for varying levels of the risk premium. In contrast to the discussion so far, I consider a specification of the Epstein and Zin (1989)-preferences with $\psi \neq 1/\gamma$. The qualitative properties of asset prices highlighted for power utility continue to hold. The risk-free rate, subjective expected return, and subjective risk premium are constant in the i.i.d. economy under rational expectations, as shown by the dashed orange lines Figure 3.1. Under similarity-weighted memory, instead, the risk-free rate is increasing in the contemporaneous log endowment growth g_t, while the subjectively expected return is a convex function of g_t due to the effect of higher-order moments (see Footnote 30). Similarly, as discussed above, the subjective risk premium of the agent is a convex function of the contemporaneous log endowment growth, due to the convexity of the posterior variance.

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

$$\mathbb{E}\left(r_{t+1}\right) - r_t^f = \lambda\, \mathbb{E}(g_{t+1}) + \mathcal{K}_t(-\gamma) - \mathcal{K}_t(\lambda-\gamma) - \frac{\bar{p}}{1-\bar{p}} \left(\mathbb{E}\left(dp_{t+1}\right) - dp_t\right), \tag{3.23}$$

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

To gain intuition, consider the case of log-normal endowment growth ($\mu = \mu_1 = \mu_2$ and

$\sigma = \sigma_1 = \sigma_2$), since closed-form solutions exist in this case.[32] Under log-normality, it is

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

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

Figure 3.2: Objective and subjective risk premium

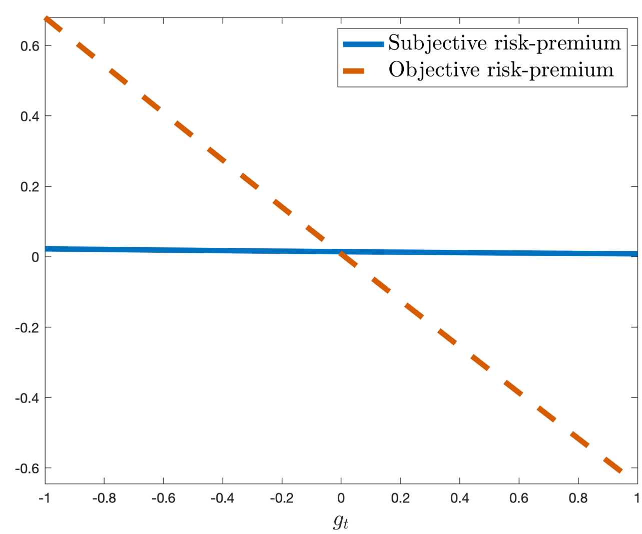

Figure 3.2 shows the objective (orange dashed line) and subjective (blue solid line) risk premium for different realizations of the contemporaneous log endowment growth g_t. The parameters of the log endowment growth process are $\mu_1 = 0.1, \mu_2 = 0.2, \sigma_1 = 0.1, \sigma_2 = 0.2, \pi_1 = 0.9, \lambda = 3$, and the preference parameters are $\beta = 0.98, \gamma = 4, \psi = 1.5, \kappa = 0.01$.

Figure 3.2 shows the subjective and objective risk premium for the general i.i.d. Markov-

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

switching process considered here. The intuition from the log-normal case discussed above extends to the more general i.i.d. Markov-switching process: When this period's endowment growth is high, such that the agent becomes very optimistic, the realized risk premium in the next period will, on average, be low. Thus, the objective risk premium in our setting is predicatbly countercyclical. Moreover, note two additional properties of the objective risk premium: Equation 3.23 shows that the dividend-price ratio positively predicts the realized risk premium; and Figure 3.2 indicates that the subjective risk premium varies significantly less than the objective risk premium.

Summary. Similarity-weighted memory explains salient empirical differences between asset prices and subjective expectations thereof, especially patterns of cyclicality, predictability, and the sensitivity of risk premia to risk measures. First, subjective expected returns under similarity-weighted memory are procyclical (Greenwood and Shleifer 2014) as a consequence of procyclical fundamental expectations (De La O and Myers 2021, Bordalo et al. 2022, Nagel and Xu 2022), while objective returns are countercyclical (Campbell and Shiller 1988, Fama and French 1988) due to the predictable mean-reversion of the agent's beliefs about fundamentals. Second, the subjective risk premium is not predictable by aggregate valuation ratios (Nagel and Xu 2023), since both are determined under the agent's contemporaneous subjective beliefs and the agent's beliefs are uninformative for future realizations of fundamentals; but the objective risk premium is predictable by the dividend-price ratio (Campbell and Shiller 1988) due to the predictability of the agent's forecast errors. Predictability of the objective risk premium implies highly volatile returns (Shiller 1981). Third, the subjective volatility of fundamentals is convex in the contemporeaneous log endowment growth, leading to a convex subjective risk premium (Lochstoer and Muir 2022, Nagel and Xu 2023). In contrast, the variation of the objective risk premium is unrelated to objective measures of risk and to measures of risk-aversion, since both are constant by assumption (Lettau and Ludvigson 2010).

3.4 Calibration

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

Parameters. Table 3.1 summarizes the parameters that I use for the simulation of the economy. There are two types of parameters: preference/belief parameters and endowment growth parameters. For the preference parameters, I mostly follow the existing literature and set γ, the relative risk-aversion coefficient, to ten (Jin and Sui 2022).[33] Next, I set ψ, the elasticity of intertemporal substitution (EIS), to 1.5. The asset pricing literature considers a range of different values for the EIS, but the majority of papers uses values above one (Beeler and Campbell 2012).[34] Finally, I set β, the discount factor, to 0.9967 as in Nagel and Xu (2022), and κ, the scrutiny parameter to 0.01, so that scrutiny is of the same magnitude as the volatility of endowment growth.

Moreover, I estimate the parameters of the i.i.d. Markov-switching process for log endowment growth in Equation 3.1 using the methods in Johannes et al. (2016), and obtain estimates that are close to the results reported therein. I measure endowment growth using standard data—the quarterly growth of services and non-durable consumption from the Bureau of Economic Analysis from Q1 1947 until Q1 2023—and estimate the parameters with a Markov-Chain-Monte-Carlo (MCMC) method (for details, see Appendix G.1). The parameters reported in Table 3.1 are the Bayesian maximum a-posterior parameters among

[33]The long-run risk literature often assigns value of up to ten to the relative risk-aversion coefficient, and Mehra and Prescott (1985) argue that values up to and around ten are reasonable. Lowering the relative risk-aversion to $\gamma = 4$ (Nagel and Xu 2022) yields a subjective risk premium of approximately 0.5 in the case without parameter uncertainty shown in Table 3.2, but otherwise does not affect the results qualitatively.

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

Table 3.1: Calibration parameters

Parameter	Symbol	Value	Source
Endowment growth process			
Mean growth in			
State 1	μ_1	0.49%	Estimated
State 2	μ_2	0.16%	Estimated
Growth volatility in			
State 1	σ_1	0.45%	Estimated
State 2	σ_2	2.36%	Estimated
Probability of state 1	π_1	0.86	Estimated
Leverage	λ	3.00	Nagel and Xu (2022)
Preferences and memory			
Risk aversion	γ	10	Jin and Sui (2022)
EIS	ψ	1.5	Nagel and Xu (2022)
Time discount factor	β	0.9967	Nagel and Xu (2022)
Memory scrutiny	κ	0.01	-

Table 3.1 reports the parameters used in the simulation. The parameters of the endowment growth process are estimated using the methods in Johannes et al. (2016). For preferences, most values used are the same as in Jin and Sui (2022) and Nagel and Xu (2022). Memory scrutiny κ is set to be of the same magnitude as the volatility of endowment growth.

$10,000$ parameter combinations. Furthermore, I estimate the leverage parameter λ by regressing quarterly aggregate dividends obtained from CRSP on endowment growth and find $\lambda \approx 3.29$, which is close to $\lambda = 3$ used in Collin-Dufresne et al. (2016) and Nagel and Xu (2022). I set$\lambda = 3$ for comparability.

Simulation. I simulate the model at a quarterly frequency. Table 3.2 reports the annualized average moments from $10,000$ simulations of the model, each for 304 quarters. The left panel of Table 3.2 shows that the model can qualitatively match salient facts about asset prices.

Figure 3.3 shows the posterior beliefs from one realization of the endowment growth process. In combination with Table 3.2, Figure 3.3 shows that the agent's posterior mean of the good state $\hat{\mu}_{1,t}$—which is the state that occurs most often—is unbiased, while the posterior mean of the recession state is biased upward and fluctuates more strongly than $\hat{\mu}_{1,t}$.

Table 3.2: Simulation results with baseline parameters

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

Table 3.2 reports the model moments obtained from 10,000 simulations of the model for 304 quarters. The left panel (No parameter uncertainty) reports results under the assumption that the agent has an infinite sample of observations, while the right panel (Parameter uncertainty) reports results when the agent learns from a finite sample, such that his Bayesian posterior has a strictily positive variance around the parameter values. I use a burn-in period of 120 quarterly observations in the parameter uncertainty simulations. For each of the 10,000 economies, I draw 10 realizations of the agent's memory and average over these realizations. Returns and expectations are annualized as follows: the means are multiplied by four and the standard deviations are multiplied by two. For the risk-free rate, I multiply the quarterly mean and the standard deviation by four.

Furthermore, the posterior variance of the agent in the right panel of Figure 3.3 tends to spike during recessions (the vertical lines) and is on average slightly higher during recessions

Figure 3.3: Posterior long-term beliefs

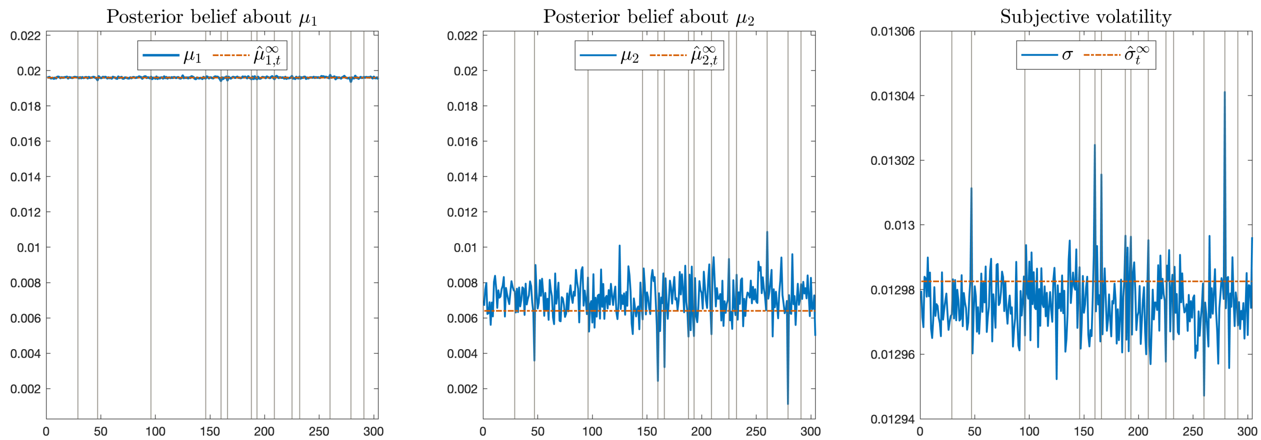

Figure 3.3 shows the annualized posterior beliefs of the agent for one realization of endowment growth. The left (middle) panel shows the posterior mean endowment growth in the good state (bad state) and the right panel shows the subjective volatility. The orange dash-dotted line in each panel plots the respective quantity under full-information rational expectations, and the vertical lines mark realized bad states. The simulation parameters are as in Table 3.1, except that I use $p_1 = 0.96$ to obtain fewer vertical lines.

than during normal times. Subjective volatility is negatively correlated with endowment growth, such that the agent perceives a more risky economy in bad times than in good times. The bottom row in Table 3.3 shows the coefficient of aCoibion and Gorodnichenko (2015)-regression as in Proposition 5 and indicates a marginal overreaction of the agent's beliefs in that $b_{CG} < 0$.[35]

The left panel of Table 3.2 shows that expected returns and the risk-free rate are procyclical (positive correlation with endowment growth), while subjective and objective risk premia are countercyclical (negative correlation with endowment growth), consistent with stylized empirical facts. Furthermore, the subjective risk premium of the agent is almost constant, with an average standard deviation of 0.005% and does, on average, also not vary across normal times (average of 1.199%) or recessions (average of 1.204%). In contrast to the subjective risk premium, the objective risk premium varies more strongly (standard de-

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

Table 3.3: Predictability and Coibion and Gorodnichenko (2015)-regressions

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

Table 3.3 reports the mean estimates from regressions for $10,000$ simulations of the model for 304 quarters plus a 110 quarters burn-in period in the parameter-uncertainty case. The first row shows the mean coefficients when regressing subjectively expected and objectively obtained risk premia on the log dividend-price ratio, as in Nagel and Xu (2023). The price-dividend ratio is rescaled to unit standard deviation. The second row shows the mean estimate from Coibion and Gorodnichenko (2015)-regressions of the forecast error on the forecast revision. The left panel shows the results obtained without parameter uncertainty, while the right panel shows the results with parameter uncertainty.

viation of 2.237%), but is slightly lower than the subjectively expected risk premium with an average of 0.998%. The countercyclicality of the objective risk premium is more pronounced, since the objective risk premium is lower in normal times (0.007%) than during recessions (2.360%). Table 3.3 shows that the objective risk premium is predictably countercyclical using the dividend-price ratio ($\hat{b}_{obj} = 1.119$), while the subjective risk premium does not vary with the dividend-price ratio ($\hat{b}_{subj} = 0.002$). The higher volatility and predictability of the objective risk premium than of the subjective risk premium is consistent with empirical findings in Nagel and Xu (2023).

The left panel of Table 3.2 highlights that asset prices in an economy with similarity-weighted memory distortions are qualitatively consistent with stylized facts, but they do not match the data well. Empirically, endowment growth—which is the sole factor that affect's the agent's long-term beliefs—is very smooth, such that the agent's beliefs are also very smooth. The low volatility of the agent's beliefs, in turn, restricts the extent to which the model can generate a high risk premium. **Parameter uncertainty.** A long literature attempts to match the level of the equity risk premium and of the risk-free rate, including models with habit formation (Campbell and Cochrane 1999), long-run risks (Bansal and Yaron 2004), or consumption disasters (Barro 2006). Collin-Dufresne et al. (2016) show that the parameter uncertainty emerging from Bayesian learning endogenously generates long-run risk under the agent's beliefs and a high risk premium. In this section, I consider a natural extension of my model to the case in which the agent learns with similarity-weighted memory from finitely many observations, thus generating Bayesian parameter uncertainty. In addition to matching the qualitative features of asset prices, the model with parameter uncertainty achieves a quantitative fit of empirical asset prices.

As before, log endowment growth g_t is normally distributed conditional on the state, but the agent does not know the state-wise mean μ_s. In any period, the agent recalls $|H^R_{t,s}| = k_{t,s} < \infty$ past endowment growth observations per state, but is naïve about the selectivity of her memory. The agent's prior about mean endowment growth in each state is $\mu_s \sim \mathcal{N}\left(\mu_{0,s}, \frac{\sigma_s^2}{\nu_s}\right)$, where ν_s scales the informativeness of the prior. The agent's posterior upon recalling the state-dependent history $H^R_{t,s}$ is

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

with

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

The agent is uncertain about the mean of log endowment growth when her prior has positive variance, $\nu_s < \infty$, and when she recalls a finite number of past observations, $k_{t,s} < \infty$. I follow the approach in Collin-Dufresne et al. (2016) and Nagel and Xu (2022) to analyze the effect of memory-induced Bayesian parameter uncertainty on asset prices. With parameter uncertainty, asset prices must be approximated numerically (see Appendix G.2).[36]

I simulate the model 10,000 times for 304 quarters. For each quarter and each simulation, I draw the agent's memory based on the memory function in Equation 3.2 for 10 times.

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

Throughout the simulations, I focus on the case in which the agent's prior is uninformative $\nu_s \to 0$. Additionally, I generate 120 quarterly endowment growth realizations as burn-in period, such that the agent has access to 30 years of data before "entering" the market.

Figure 3.4: Posterior beliefs with parameter uncertainty

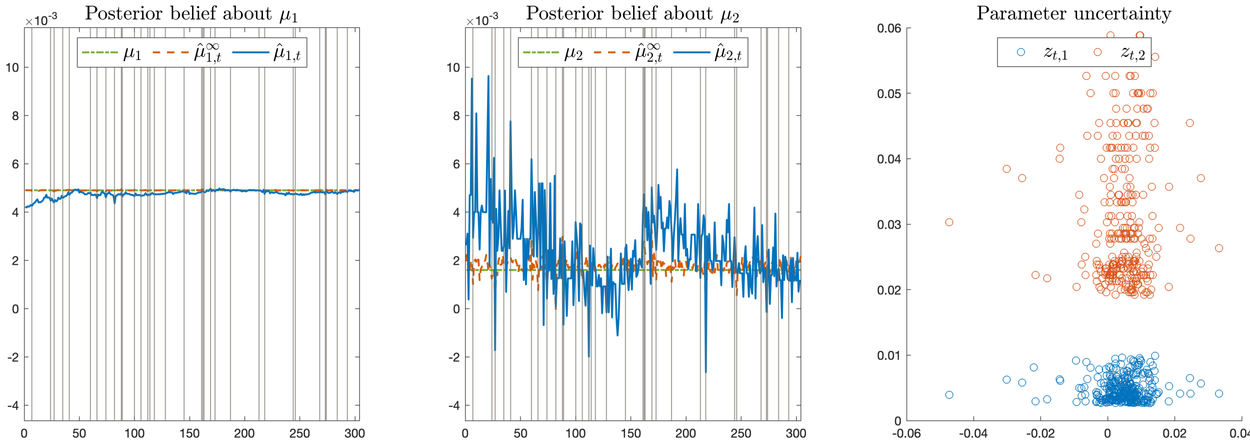

Figure 3.4 shows one time-series realization of the agent's posterior beliefs when allowing for parameter uncertainty. The left panel shows the beliefs about μ_1, and the middle panel shows the beliefs about μ_2. In both panels, the solid blue line shows the posterior with a limited number of observations, while the orange dotted line shows the beliefs in the $t \to \infty$-limit. The horizontal green dashed line plots the fundamentally true mean μ_s, and the vertical grey lines mark recession periods. The right panel plots the parameter uncertainty $z_{t,s}$ against the endowment growth g_t. I focus on an uninformative prior $\nu_s \to 0$. The simulation parameters are as in Table 3.1.

Figure 3.4 plots one realization of the agent's beliefs. The blue solid line in the left and middle panel shows the agent's posterior mean endowment growth in state 1 and state 2, respectively. The orange dotted line shows the posterior mean in the $t \to \infty$-limit (Proposition 2), while the green dashed line shows the true fundamental mean μ_s. The vertical lines mark periods in which recession state realizes. Finally, the right panel of Figure 3.4 plots the parameter uncertainty $z_{t,s}$ against the endowment growth g_t. Combined with the results in Table 3.2, Figure 3.4 highlights that the agent's beliefs when learning from a limited sample are close to the beliefs in the $t \to \infty$-case characterized in Proposition 2. Compared to the case without parameter uncertainty, the agent's posterior beliefs fluctuate more when learning from a limited sample, since the standard deviation of the posterior mean is higher in the right panel of Table 3.2 than in the left panel. Similarly, the properties of the subjective volatility of the economy discussed for the case without parameter uncertainty persist in the case with parameter uncertainty, but the correlation is lower (in absolute terms). The bottom row of Table 3.3 shows that the agent's beliefs overreact more strongly when learning from a limited sample (right panel), with a Coibion and Gorodnichenko (2015)-regression coefficient of $\hat{b}_{CG} = -0.419$.

The invariance of the qualitative features of beliefs when going from $t \to \infty$ to a limited-t implies that I find the qualitative features of asset prices discussed above also when considering parameter uncertainty: The expected return and risk-free rate are procyclical, whereas subjective and objective risk premia are countercyclical. Moreover, the objective risk premium fluctuates more than the subjective risk premium. Finally, the objective risk premium is predictable by the dividend-price ratio ($\hat{b} = 3.896$), while the subjective risk premium does not fluctuate much with the dividend-price ratio ($\hat{b} = 0.002$).

Furthermore, considering parameter uncertainty also achieves an excellent quantitative fit of empirical asset prices. Empirically, Jin and Sui (2022) report an average realized risk premium of 3.90%, while the average subjective risk premium, measured as log expected return minus the log risk-free rate as in the theoretical model, is around 1.90%.[37] In the simulations, I obtain an average objective risk premium of 3.21%, which is considerably higher during recessions (12.05%) than during normal times (1.77%). The subjective risk premium is lower than the objective risk premium at 1.11%, and does not vary much across normal times (1.08%) and recessions (1.34%). Similarly, the risk-free rate is within the range of empirical estimates with 2.17% (Jin and Sui 2022, Nagel and Xu 2022) and has a low volatility (Campbell and Cochrane (1999), for example, reverse-engineer a constant risk-free rate). Departing from rational expectations by considering subjective beliefs that arise from similarity-weighted memory explains the mismatch of subjective and objective asset prices regarding the cyclicality, predictability, and sensitivity to risk measures; and subjective beliefs under similarity-weighted memory generate a realistically high risk premium and low

[37] I measure the subjective risk premium using the data provided by Nagel and Xu (2022), which combines several individual surveys. The data is available on the website of the Review of Financial Studies.

risk-free rate.

4 Asset prices under a peak-end rule memory distortion

In this section, I demonstrate the flexibility of the general model introduced in Section 2 by analyzing another memory distortion. From now on, I assume that the agent's memory of past experiences is distorted by the peak-end rule. The peak-end rule describes the finding that humans evaluate the past based on the most valent moments (peak or trough) and "the end," which is justified by the higher likelihood of recalling extreme and recent experiences (Kahneman 2000, Alaybek et al. 2022). As before, I first outline the setup to then discuss the implications of the peak-end rule on the agent's beliefs and asset prices. The peak-end rule has not yet received as much attention and empirical support in the finance literature as similarity-weighted memory, such that this section is shorter than Section 3.

4.1 Framework

Consider log-normal endowment growth

$$g_t = \mu + \sigma\,\epsilon_t, \qquad \epsilon_t \overset{i.i.d.}{\sim} \mathcal{N}\left(0,1\right). \tag{4.1}$$

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

$$m^{\mathrm{PE}}\left(g_\tau, g_t\right) := \underbrace{\exp\left[-e^{-\frac{(g_\tau-\mu)^2}{2\,\sigma^2}}\right]}_{\text{Extreme experience bias}} \cdot \underbrace{\exp\left[\frac{(g_\tau - g_t)^2}{2\,\kappa}\right]}_{\text{Similarity}}. \tag{4.2}$$

The extreme experience bias models the higher likelihood of recalling "peaks." The inner exponential term is proportional to the probability density function (pdf) of a normal distribution, with higher values close to the true mean μ and lower values in the tails of the distribution of log endowment growth, and the outer exponential implies that the center of the distribution—values around μ—are underweighted, while values in the tails of the distribution are overweighted.[38] The similarity term models the higher likelihood of recalling experiences at the end.[39] Figure D.3 illustrates the memory function.

4.2 Subjective long-term beliefs and asset prices

In this section, I analyze the agent's subjective long-term beliefs under the peak-end memory function and highlight the implications for asset prices. Closed-form solutions for the agent's long-term belief do not exist under the memory function in Equation 4.2, such that I will mainly discuss the impact of the extreme experience bias. The effect of the similarity term on the agent's beliefs and asset prices is as in Section 3.

Subjective beliefs. The agent is more likely to recall extreme experiences and experiences that are similar to the time-varying end of the realized history H_t. The extreme experience bias does not systematically affect the posterior mean of the agent since both tails of the distribution of log endowment growth are overweighted symmetrically, but an extreme experience bias leads to an increased posterior variance.[40] Appendix C.3 shows that a pure extreme experience bias as specified by the first component in Equation 4.2 yields $\hat{\mu}_t = \mu$ and $\hat{\sigma}_t^2 \approx 1.41 \cdot \sigma^2 > \sigma^2$.

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

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

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

Figure 4.1 shows the posterior mean and variance of the agent under the peak-end memory distortion given in Equation 4.2, and I provide a numerical approximation of the agent's posterior mean in Appendix C.4. The left panel of Figure 4.1 highlights that the agent's posterior mean is an increasing function of this period's endwoment growth g_t, since the posterior mean under a peak-end memory distortion is closely related to the posterior mean under similarity-weighted memory (see Appendix C.4). The right panel of Figure 4.1 shows the posterior variance of the agent. Under peak-end memory, the posterior variance is higher than the true variance of the process due to the overweighting of extreme observations. Moreover, the posterior volatility is concave in contemporaneous endowment growth with a maximum at $g_t = \mu$ due to the combination of extreme experience bias and similiarty-weighting: If today's endowment growth $g_t = \mu$, the memory-function becomes identical to the pure overweighting of extreme experiences, and the posterior variance of the agent is $\hat{\sigma}_t^2 \approx 1.41 \cdot \sigma^2$. In contrast, if $g_t \neq \mu$, the agent recalls more endowment growth rates that are close to g_t, shifting the mass of recalled experiences closer to one tail of the distribution. The shift of the mass of recalled experiences to one tail of the distribution reduces the posterior variance, and the reduction is stronger the more mass is shifted towards the tails of the distribution.

Asset pricing implications. As a next step, I analyze the effect of the peak-end memory distortion on asset prices using the framework reviewed in Section 2.3. The cumulant-generating function of log endowment growth under the agent's time-t belief is given by

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

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

As in Section 3.4, I simulate the model $10,000$ times for 304 quarters and report average moments in Table 4.1. The parameters of the log endowment growth process are as in Nagel and Xu (2022) with a quarterly mean endowment growth of $\mu = 0.44\%$ and a quarterly

Figure 4.1: Posterior beliefs under peak-end memory

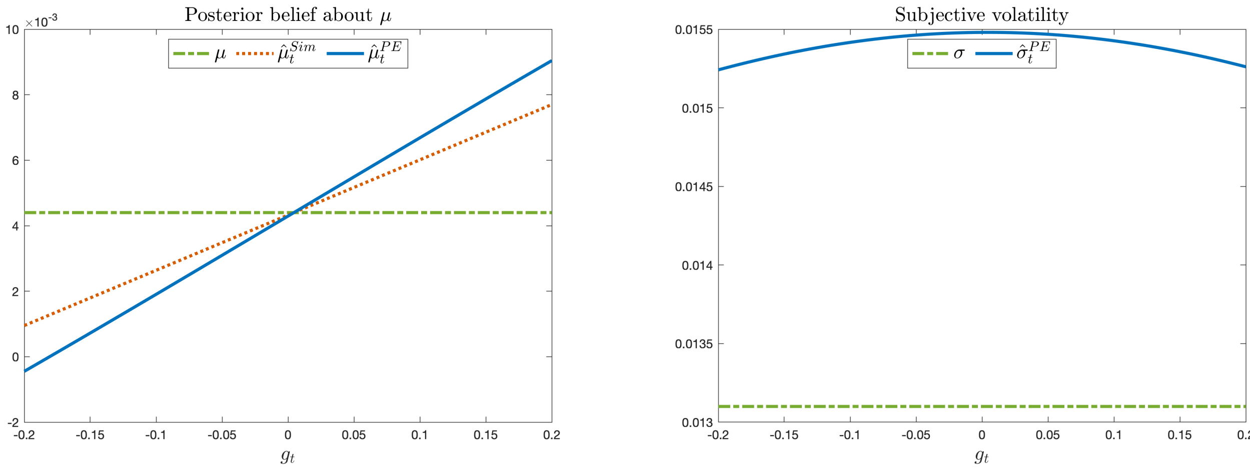

Figure 4.1 shows the posterior mean and variance of an agent with peak-end memory distortions as in Equation 4.2 for varying levels of contemporaneous endowment growth g_t. The parameters are $\mu = 0.44\%$, $\sigma = 1.31\%$, and $\kappa = 0.01$.

volatility of $\sigma = 1.31\%$. All other parameters are as in Table 3.1.

Table 4.1: Asset prices under the peak-end rule

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

Table 4.1 reports the model moments obtained from 10,000 simulations of the model for 304 quarters. I annualize the quantities as follows: Means are multiplied by four and the standard deviations are multiplied by two. For the risk-free rate, I multiply the quarterly mean and the standard deviation by four.

The simulation results in Table 4.1 highlight that the agent's posterior mean is an unbiased estimate of the true mean, with an average posterior mean of 1.760%, and is relatively stable with an average standard deviation of 0.063%. Time-variation in the agent's posterior mean is entirely driven by the similarity component of the memory function in Equation 4.2, such that the agent's posterior mean is perfectly correlated with this period's endowment growth g_t, leading to procyclical beliefs. Table E.2 in the Appendix shows the results of a Coibion and Gorodnichenko (2015)-regression and highlights that the agent's beliefs over-react to new information as in Section 3.4. The agent has a higher likelihood of recalling extreme experiences, which leads to a higher posterior volatility ($\hat{\sigma}_t = 3.096\%$) than fundamental volatility ($\sigma = 2.620\%$). The posterior volatility does barely vary over time (std. $< 0.001\%$) and is not correlated with contemporaneous endowment growth g_t.

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

5 Conclusion

This paper explores the implications of selective memory for asset prices and shows that similarity-weighted memory simultaneously accounts for important facts about belief formation, survey data, and asset prices. With i.i.d. fundamentals and constant risk-aversion, my model explains the empirically observed discrepancies between subjectively expected and objectively realized returns using a simple mechanism: The agent selectively recalls past observations of the fundamental that are similar to its contemporaneous realization. A good realization of the fundamental causes the agent to become too optimistic in that she expects a high future growth with a low volatility, leading to a high expected return and a low risk premium. The agent's optimism implies that 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 general framework proposed in this paper can be used to analyze further selective memory distortions, and I demonstrate this flexibility by considering the peak-end rule. Under the peak-end rule, the subjective risk premium is high because the agent tends to recall extreme experiences. In addition, the model could be extended to active learning models in a portfolio choice context (Gödker et al. 2022, Fudenberg et al. 2023), to agents who access false memories, or to heterogeneous agent settings.

The paper complements empirical evidence from finance (Bordalo et al. 2023b, Jiang et al. 2023, Nagel and Xu 2023), experiments (Enke et al. 2022, Burro et al. 2023), and cognitive psychology (Kahana 2012) by showing that selective—and especially similarity-weighted—memory theoretically explains conceptually puzzling patterns of subjective and objective asset prices without resorting to time-varying risk-aversion, long-term risk, or disaster risk. The emerging pattern suggests that selective memory systematically affects aggregate economic outcomes, such as asset prices. Understanding the role of memory in the formation of subjective beliefs could thus help us to make sense of aggregate asset prices.

References

Abel, A. B. (1999). "Risk premia and term premia in general equilibrium". In: *Journal of Monetary Economics* 43.1, pp. 3–33.

Adam, K., A. Marcet, and J. Beutel (2017). "Stock price booms and expected capital gains". In: *American Economic Review* 107.8, pp. 2352–2408.

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

Alaybek, B., R. S. Dalal, S. Fyffe, J. A. Aitken, Y. Zhou, X. Qu, A. Roman, and J. I. Baines (2022). "All's well that ends (and peaks) well? A meta-analysis of the peak-end rule and duration neglect". In: *Organizational Behavior and Human Decision Processes* 170, p. 104149.

Amromin, G. and S. A. Sharpe (2014). "From the horse's mouth: Economic conditions and investor expectations of risk and return". In: *Management Science* 60.4, pp. 845–866.

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

Bacchetta, P., E. Mertens, and E. Van Wincoop (2009). "Predictability in financial markets: What do survey expectations tell us?" In: *Journal of International Money and Finance* 28.3, pp. 406–426.

Bansal, R. and A. Yaron (2004). "Risks for the long run: A potential resolution of asset pricing puzzles". In: *Journal of Finance* 59.4, pp. 1481–1509.

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

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer (2015). "X-CAPM: An extrapolative capital asset pricing model". In: *Journal of Financial Economics* 115.1, pp. 1–24.

Barberis, N., R. Greenwood, L. Jin, and A. Shleifer (2018). "Extrapolation and bubbles". In: *Journal of Financial Economics* 129.2, pp. 203–227.

Barro, R. J. (2006). "Rare disasters and asset markets in the twentieth century". In: *Quarterly Journal of Economics* 121.3, pp. 823–866.

Bastianello, F. and P. Fontanier (2022). "Expectations and learning from prices". In: *Working paper available at SRRN.*

Beeler, J. and J. Y. Campbell (2012). "The Long-Run Risks Model and Aggregate Asset Prices: An Empirical Assessment". In: *Critical Finance Review* 1.1, pp. 141–182.

Berk, R. H. (1966). "Limiting behavior of posterior distributions when the model is incorrect". In: *Annals of Mathematical Statistics* 37.1, pp. 51–58.

Bodoh-Creed, A. L. (2020). "Mood, memory, and the evaluation of asset prices". In: *Review of Finance* 24.1, pp. 227–262.

Bordalo, P., J. J. Conlon, N. Gennaioli, S. Y. Kwon, and A. Shleifer (2023a). "Memory and probability". In: *Quarterly Journal of Economics* 138.1, pp. 265–311.

Bordalo, P., N. Gennaioli, R. La Porta, and A. Shleifer (2019). "Diagnostic expectations and stock returns". In: *Journal of Finance* 74.6, pp. 2839–2874.

Bordalo, P., N. Gennaioli, R. La Porta, and A. Shleifer (2022). "Expectations of fundamentals and stock market puzzles". In: *NBER Working Paper*.

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

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020a). "Overreaction in macroeconomic expectations". In: *American Economic Review* 110.9, pp. 2748–2782.

Bordalo, P., N. Gennaioli, and A. Shleifer (2018). "Diagnostic expectations and credit cycles". In: *Journal of Finance* 73.1, pp. 199–227.

Bordalo, P., N. Gennaioli, and A. Shleifer (2020b). "Memory, attention, and choice". In: *Quarterly Journal of Economics* 135.3, pp. 1399–1442.

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

Campbell, J. Y. (1986). "Bond and stock returns in a simple exchange model". In: *Quarterly Journal of Economics* 101.4, pp. 785–803.

Campbell, J. Y. (1991). "A Variance Decomposition for Stock Returns". In: *Economic Journal* 101.405, pp. 157–179.

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

Campbell, J. Y. and J. H. Cochrane (1999). "By force of habit: A consumption-based explanation of aggregate stock market behavior". In: *Journal of Political Economy* 107.2, pp. 205–251.

Campbell, J. Y. and R. Shiller (1988). "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors". In: *Review of Financial Studies* 1.3, pp. 195–228.

Carr, H. A. (1931). "The Laws of Association". In: *Psychological Review* 38.3, p. 212.

Cassella, S. and H. Gulen (2018). "Extrapolation bias and the predictability of stock returns by price-scaled variables". In: *Review of Financial Studies* 31.11, pp. 4345–4397.

Charles, C. (2021). "Memory and trading". In: *Working paper available at SRRN.*

Charles, C. (2022). "Memory Moves Markets". In: *Working paper available at SRRN.*

Coibion, O. and Y. Gorodnichenko (2015). "Information rigidity and the expectations formation process: A simple framework and new facts". In: *American Economic Review* 105.8, pp. 2644–2678.

Collin-Dufresne, P., M. Johannes, and L. A. Lochstoer (2016). "Parameter learning in general equilibrium: The asset pricing implications". In: *American Economic Review* 106.3, pp. 664–98.

Cruciani, F., A. Berardi, S. Cabib, and D. Conversi (2011). "Positive and negative emotional arousal increases duration of memory traces: common and independent mechanisms". In: *Frontiers in behavioral neuroscience* 5, p. 86.

d'Acremont, M., W. Schultz, and P. Bossaerts (2013). "The human brain encodes event frequencies while forming subjective beliefs". In: *Journal of Neuroscience* 33.26, pp. 10887–10897.

Da, Z., X. Huang, and L. J. Jin (2021). "Extrapolative beliefs in the cross-section: What can we learn from the crowds?" In: *Journal of Financial Economics* 140.1, pp. 175–196.

De La O, R. and S. Myers (2021). "Subjective cash flow and discount rate expectations". In: *Journal of Finance* 76.3, pp. 1339–1387.

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

Ehling, P., A. Graniero, and C. Heyerdahl-Larsen (2018). "Asset prices and portfolio choice with learning from experience". In: *Review of Economic Studies* 85.3, pp. 1752–1780.

Enke, B., F. Schwerter, and F. Zimmermann (2022). "Associative Memory, Beliefs and Market Interactions". In: *Working paper.*

Epstein, L. G. and S. E. Zin (1989). "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework". In: *Econometrica* 57.4. Publisher: [Wiley, Econometric Society], pp. 937–969.

Esponda, I. and D. Pouzo (2016). "Berk–Nash equilibrium: A framework for modeling agents with misspecified models". In: *Econometrica* 84.3, pp. 1093–1130.

Fama, E. F. and K. R. French (1988). "Dividend yields and expected stock returns". In: *Journal of Financial Economics* 22.1, pp. 3–25.

Fudenberg, D., G. Lanzani, and P. Strack (2023). "Selective Memory Equilibrium". In: *Working paper available at SRRN.*

Gilboa, I. and D. Schmeidler (1995). "Case-based decision theory". In: *Quarterly Journal of Economics* 110.3, pp. 605–639.

Gödker, K., P. Jiao, and P. Smeets (2022). "Investor memory". In: *Working paper available at SRRN.*

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

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

Greenwood, R. and S. G. Hanson (2013). "Issuer quality and corporate bond returns". In: *Review of Financial Studies* 26.6, pp. 1483–1525.

Greenwood, R. and A. Shleifer (2014). "Expectations of returns and expected returns". In: *Review of Financial Studies* 27.3, pp. 714–746.

Hansen, L. P., J. C. Heaton, and N. Li (2008). "Consumption strikes back? Measuring long-run risk". In: *Journal of Political Economy* 116.2, pp. 260–302.

Happel, J., Y. Karabulut, L. Schäfer, and S. Tuzel (2022). "Shattered Housing". In: *Working paper available at SRRN.*

Heidhues, P., B. Kőszegi, and P. Strack (2021). "Convergence in models of misspecified learning". In: *Theoretical Economics* 16.1, pp. 73–99.

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

Jin, L. J. and P. Sui (2022). "Asset pricing with return extrapolation". In: *Journal of Financial Economics* 145.2, pp. 273–295.

Johannes, M., L. A. Lochstoer, and Y. Mou (2016). "Learning about consumption dynamics". In: *Journal of Finance* 71.2, pp. 551–600.

Johnson, T. C. (2007). "Optimal learning and new technology bubbles". In: *Journal of Monetary Economics* 54.8, pp. 2486–2511.

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

Kahana, M. J. (2012). *Foundations of human memory.* Oxford University Press.

Kahneman, D. (2000). "Evaluation by moments: Past and future". In: *Choices, values, and frames*, pp. 693–708.

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

Kleijn, B. and A. Van Der Vaart (Jan. 2012). "The Bernstein-Von-Mises theorem under misspecification". In: *Electronic Journal of Statistics* 6.none.

Kuchler, T. and B. Zafar (2019). "Personal experiences and expectations about aggregate outcomes". In: *Journal of Finance* 74.5, pp. 2491–2542.

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

Lewellen, J. and J. Shanken (2002). "Learning, asset-pricing tests, and market efficiency". In: *Journal of Finance* 57.3, pp. 1113–1145.

Li, K. and J. Liu (2023). "Extrapolative asset pricing". In: *Journal of Economic Theory* 210, p. 105651.

Lochstoer, L. A. and T. Muir (2022). "Volatility expectations and returns". In: *Journal of Finance* 77.2, pp. 1055–1096.

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

Malmendier, U. and S. Nagel (2011). "Depression babies: do macroeconomic experiences affect risk taking?" In: *Quarterly Journal of Economics* 126.1, pp. 373–416.

Malmendier, U. and S. Nagel (2016). "Learning from inflation experiences". In: *Quarterly Journal of Economics* 131.1, pp. 53–87.

Malmendier, U., D. Pouzo, and V. Vanasco (2020). "Investor experiences and financial market dynamics". In: *Journal of Financial Economics* 136.3, pp. 597–622.

Malmendier, U. and L. S. Shen (2018). *Scarred consumption.* Tech. rep. National Bureau of Economic Research.

Malmendier, U., G. Tate, and J. Yan (2011). “Overconfidence and early-life experiences: the effect of managerial traits on corporate financial policies”. In: *Journal of Finance* 66.5, pp. 1687–1733.

Malmendier, U. and J. A. Wachter (2021). “Memory of past experiences and economic decisions”. In: *Working paper available at SRRN*.

Malmendier, U. and J. A. Wachter (2022). “Memory of past experiences and economic decisions”. In: *Working paper available at SRRN*.

Martin, I. (2013). “Consumption-based asset pricing with higher cumulants”. In: *Review of Economic Studies* 80.2, pp. 745–773.

Mehra, R. and E. Prescott (1985). “The equity premium: A puzzle”. In: *Journal of Monetary Economics* 15.2, pp. 145–161.

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

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

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

Mullainathan, S. (2002). “A memory-based model of bounded rationality”. In: *Quarterly Journal of Economics* 117.3, pp. 735–774.

Müller, G. E. and A. Pilzecker (1900). *Experimentelle Beiträge zur Lehre vom Gedächtniss*. Vol. 1. Barth.

Nagel, S. and Z. Xu (2022). “Asset Pricing with Fading Memory”. In: *Review of Financial Studies* Forthcoming.

Nagel, S. and Z. Xu (2023). “Dynamics of subjective risk premia”. In: *Journal of Financial Economics* 150.2.

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

Rietz, T. A. (1988). “The equity risk premium: a solution”. In: *Journal of Monetary Economics* 22.1, pp. 117–131.

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

Shadlen, M. N. and D. Shohamy (2016). “Decision making and sequential sampling from memory”. In: *Neuron* 90.5, pp. 927–939.

Shiller, R. (1981). “Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?” In: *American Economic Review* 71.3, pp. 421–36.

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

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

Timmermann, A. G. (1993). “How learning in financial markets generates excess volatility and predictability in stock prices”. In: *Quarterly Journal of Economics* 108.4, pp. 1135–1145.

Tulving, E. and D. L. Schacter (1990). “Priming and human memory systems”. In: *Science* 247.4940, pp. 301–306.

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

Wachter, J. A. and M. J. Kahana (2023). “A retrieved-context theory of financial decisions”. In: *Quarterly Journal of Economics* Forthcoming.

Weitzman, M. L. (2007). “Subjective expectations and asset-return puzzles”. In: *American Economic Review* 97.4, pp. 1102–1130.

Zimmermann, F. (2020). “The dynamics of motivated beliefs”. In: *American Economic Review* 110.2, pp. 337–61.

A Proofs for Section 2

A.1 Proposition 1

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

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

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

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

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

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

$$\mu_{LM} = \tilde{\mu}_m,$$
$$\sigma^2_{LM} = \tilde{\sigma^2}_m.$$

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

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

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

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

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

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

We now solve for

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

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

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

From the second to the third line, I used the assumptions that (1) g_τ is i.i.d. and (2) that the memory of $g_j, j \neq \tau$ is independent of g_τ. Next, note that for $\tau_0 \to -\infty$, we have that

$\hat{\mu}_t = \mathbb{E}(\hat{\mu}_t)$ almost surely, such that we find

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

Overall, we thus have

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

and inserting into Equation A.2 yields the claim.

B Proofs for Section 3

B.1 Proposition 2

I derive the posterior belief of an agent with similarity-weighted memory as given by Equation 3.2. Note that the state $s \in \{1, 2\}$ follows an observable Markov chain, such that the agent performs state-wise inference. The memory-weighted probability distribution is given by

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

with integration constant z. The exponential term is Gaussian with

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

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

B.2 Proposition 3

It is

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

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

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

Combining both expressions, it is

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

and inserting yields the claim in Proposition 4.

B.3 Proposition 4

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

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

In addition, it is

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

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

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

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

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

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

The upper bound on the expected subjective variance is found as

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

B.4 Proposition 5

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

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

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

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

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

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

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

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

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

Finally, using Equation 3.10 and noting that $\mathbb{E}\left(\tilde{\mathbb{E}}_{\mathrm{t}}(g_{t+h}) - \mathbb{E}_{t-1}(g_{t+h})\right) = 0$, it is $a_{GC} = \pi_1\,\pi_2\,(\mu_2 - \mu_1)\;(\alpha_1 - \alpha_2)$.

B.5 Proposition 6

The results follow from inserting the cumulant-generating function given in Equation 3.5 into Result 1 in Martin (2013), as shown in Section 2.3. Under Epstein and Zin (1989)-

preferences, it is

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

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

C Proofs for Section 4: Peak-end rule

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

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

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

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

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

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

The CDF of X^*_n is

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

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

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

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

We can tighten the bound using Gordon's Inequality as

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

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

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

We can then find

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

and

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

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

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

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

C.2 Memory-weighted probability distribution

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

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

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

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

Restrict attention to the extreme-experience bias $m^P(g_\tau)$ with $a = \mu$ and $b = \sigma^2$.[41] 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.

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

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

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

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

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

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

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

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

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

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

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

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

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

Putting terms together, it is

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

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

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

Let us consider the memory function defined in Equation 4.2. First, I show that the memory-weighted probability distribution exists by showing that the integration constant $\mathcal{V}$ exists:

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

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

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

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

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

$$\begin{aligned} m^{PE}(g_\tau, g_t) &= e^{-\frac{(g_\tau-\mu)^2}{2\,\sigma^2}}\, e^{-\frac{(g_\tau-g_t)^2}{2\,\kappa}}\, e^{-e^{-\frac{(g_\tau-\mu)^2}{2\,\sigma^2}}} \\ &= e^{\frac{(\mu-g_t)^2}{2}}\, e^{\frac{(g_\tau-\hat{\mu}_t^S)^2}{2\,(\hat{\sigma}_t^S)^2}}\, e^{-e^{-\frac{(g_\tau-\mu)^2}{2\,\sigma^2}}}. \end{aligned}$$

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

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

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

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

Thus, putting terms together, it is

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

D Additional figures

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

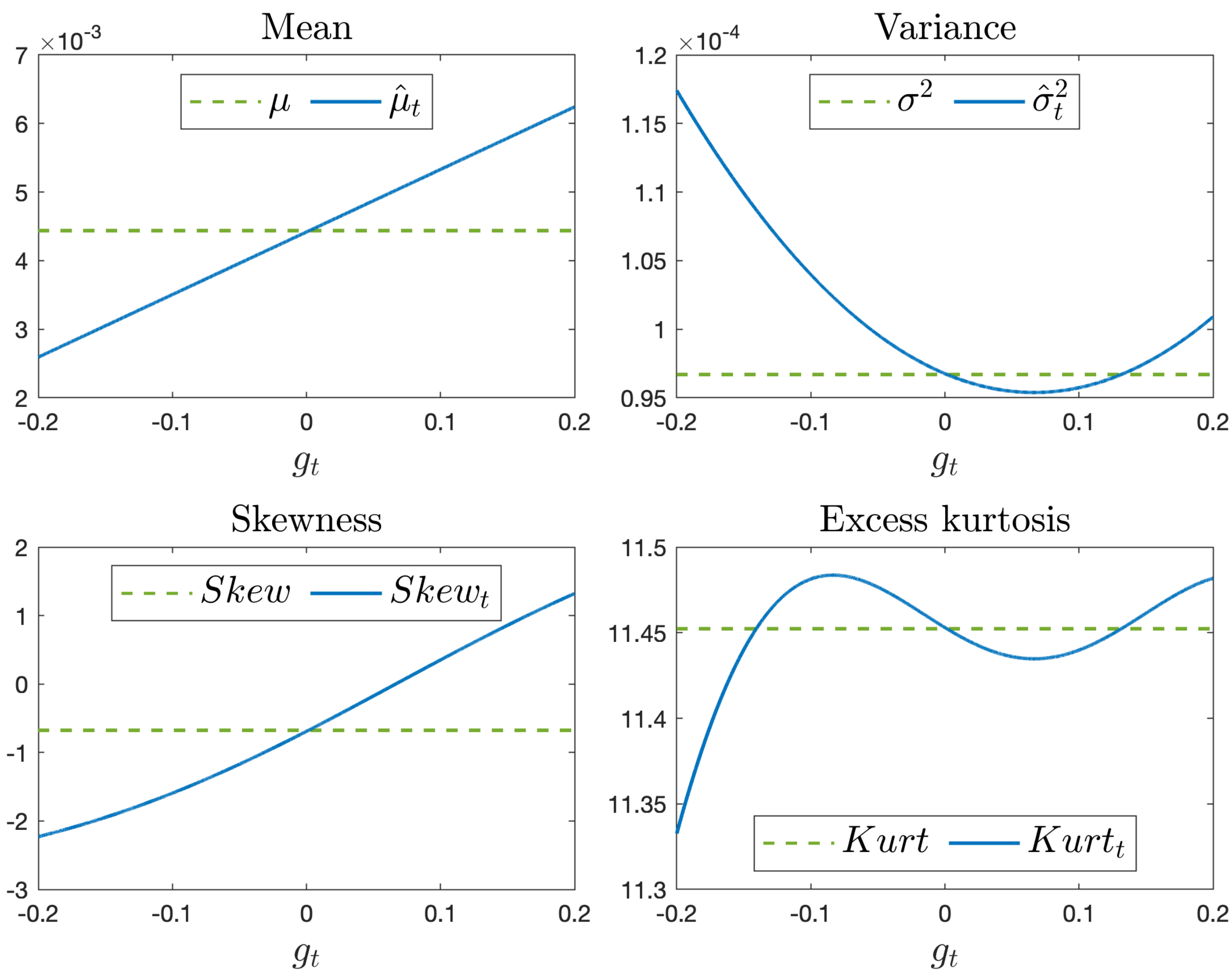

Figure D.1 shows the first four moments of endowment growth under the agent's subjective beliefs derived from similarity-weighted memory (solid blue line) and the true underlying values (green dashed line). Endowment growth is distributed as in Equation 3.1, and the similarity-weighted memory function is as in Equation 3.2. The parameters used to generate Figure D.1 are as in Table 3.1.

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

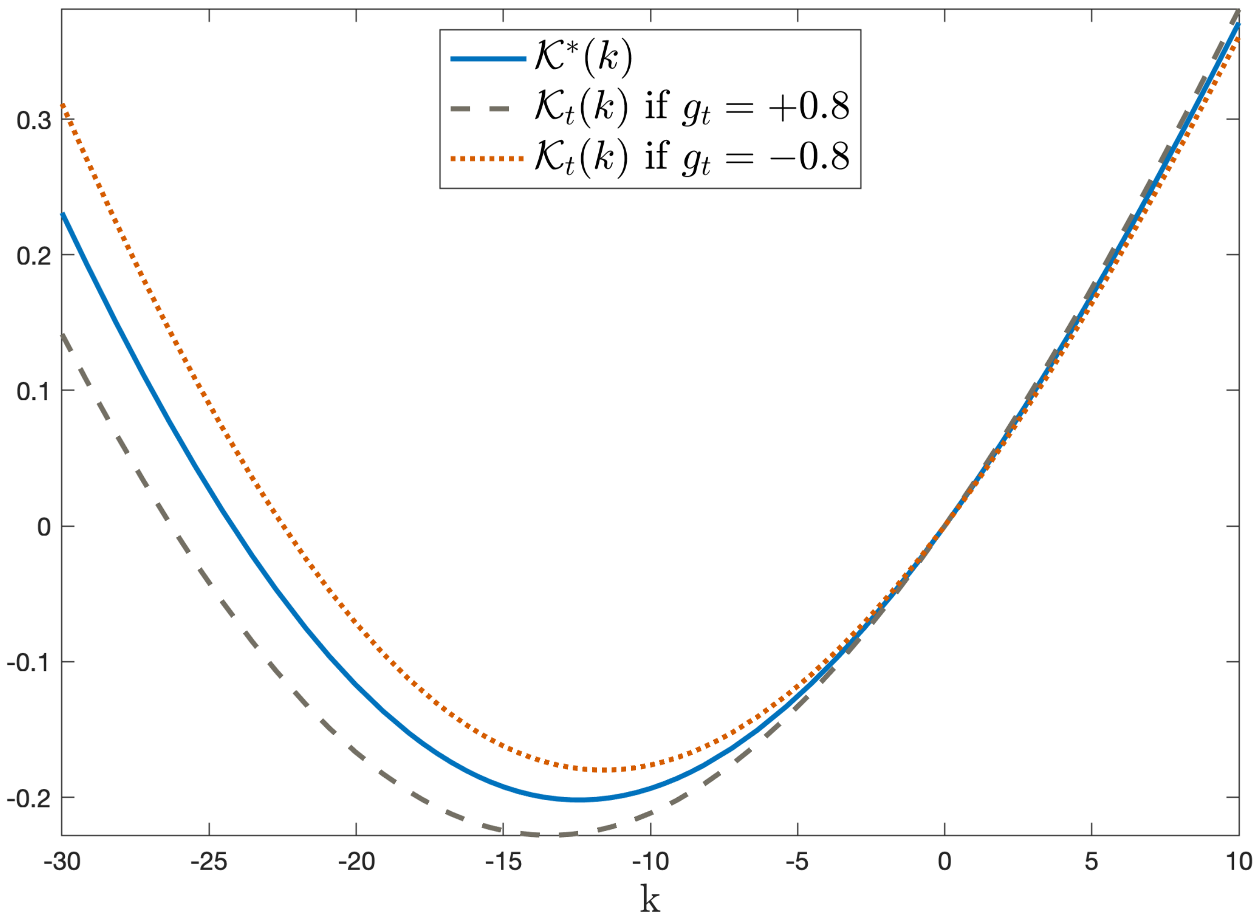

Figure D.2 shows the true cumulant-generating function of endowment growth $\mathcal{K}^*(k)$ (blue solid line) and the agent's subjective cumulant-generating function for a highly positive (grey dashed line) and negative (orange dotted line) current endowment growth g_t. The cumulant-generating functions given in Equation 3.5. The parameters are $\mu_1 = 0.05$, $\mu_2 = -0.05$, $\sigma_1 = 0.01$, $\sigma_2 = 0.02$, $\pi_1 = 0.8$, and $\kappa = 0.1$. Note that for any cumulant-generating function $\mathcal{K}(0)$, it is $\mathcal{K}(0) = 0$, such that changes of the current endowment growth lead to a rotation of the cumulant-generating function.

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

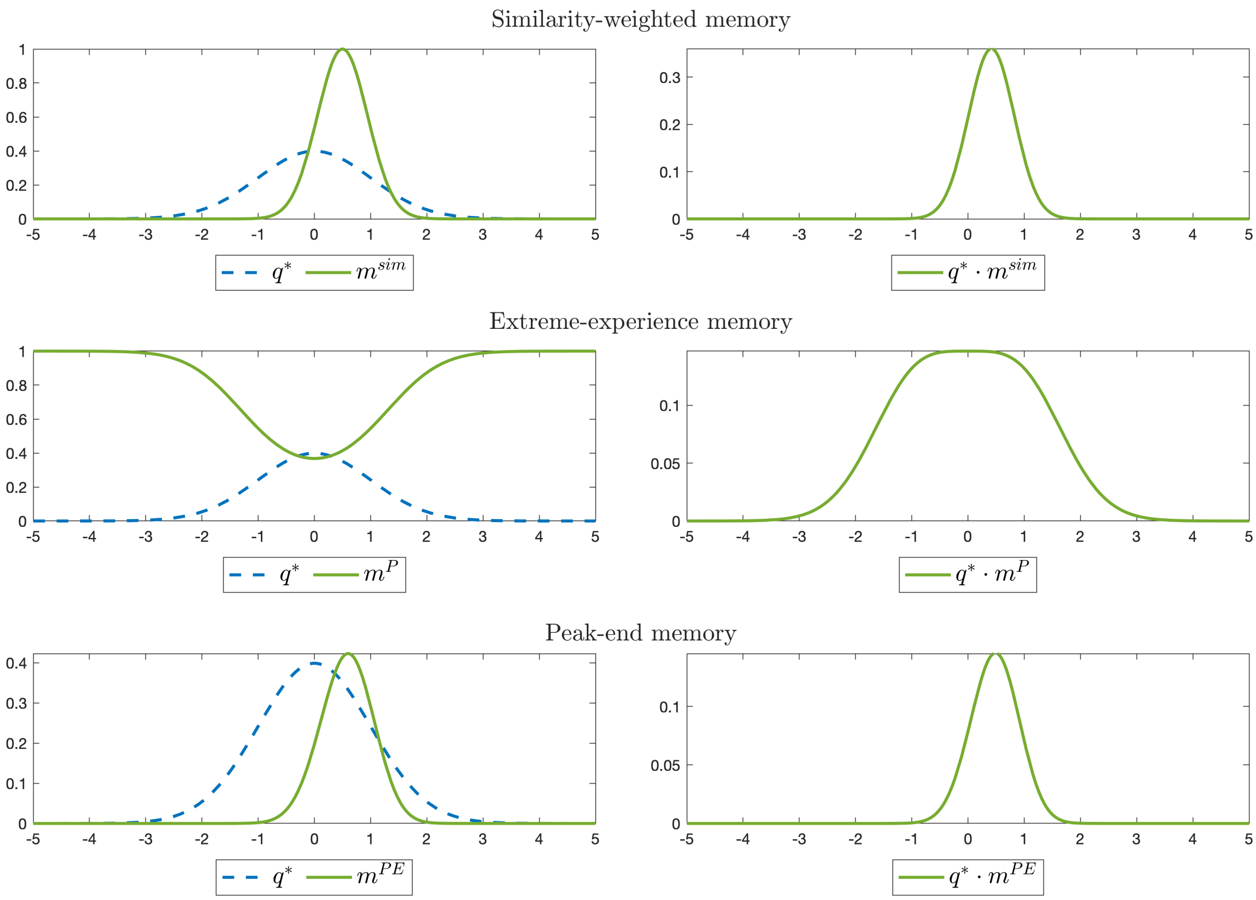

Figure D.3 shows the graphical construction of the memory-weighted true probability distribution. The first row shows the case of a similarity-weighted memory function, as considered in Section 3, the middle row shows the extreme-experience bias and the bottom row shows the peak-end memory function,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

E Additional tables

Table E.1: Average asset pricing quantities per quantile of endowment growth under parameter uncertainty

Quantile	$\overline{g}_t$	$\overline{\hat{\mu}}_t$	$\overline{rp}_t$	$\overline{rp}_t^O$
1	-2.954	1.719	1.289	18.639
2	0.813	1.769	1.081	4.674
3	1.920	1.784	1.069	1.962
4	3.011	1.798	1.070	-0.812
5	6.090	1.839	1.060	-8.414

Table E.1 reports average endwoment growth $\overline{g}_t$, posterior mean $\overline{\hat{\mu}}_t$, subjective risk premium $\overline{rp}_t$ and objective risk premium $\overline{rp}_t^O$ for each quantile of endowment growth. The moments are obtained from 10,000 simulations of the model for 304 quarters and 110 quarters burn-in period. For each of the 10,000 economies, I draw 10 realizations of the agent's memory and average over these realizations. All quantities are annualized by multiplying quarterly means by four.

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

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

Table E.2 reports the mean estimates from regressions for 10,000 simulations of the model for 304 quarters. The first row shows the mean coefficients when regressing subjectively expected and objectively obtained risk premia on the log dividend-price ratio, as in Nagel and Xu (2023). The price-dividend ratio is rescaled to unit standard deviation. The second row shows the mean estimate from Coibion and Gorodnichenko (2015)-regressions of the forecast error on the forecast revision. The agent's expectations are obtained under the peak-end memory distortion given in Equation 4.2.

F Asset pricing model and further asset-pricing results

In this appendix, I derive the asset-pricing results in Section 2.3 following Martin (2013). Consider the objective function in Equation 2.5 with $\psi \neq 1$[42] Under this formulation, the

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

stochastic discount factor becomes

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

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

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

where I conjecture that the consumption-wealth ratio CW is constant under the agent's beliefs.[43] I verify this conjecture below.

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

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

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

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

which is constant under the agent's time-t beliefs as conjectured because the agent expects

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

that $\mathcal{K}_{t+h}(k) = \mathcal{K}_t(k)$ for all $h \geq 1$. The dividend-price ratio is then

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

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

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

and the log of the expected return is

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

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

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

and the risk premium on any asset is

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

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

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

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

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

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

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

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

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

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

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

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

with

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

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

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

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

Case 1: Log-normally distributed endowment growth

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

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

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

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

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

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

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

Inserting into the previous equations, I find

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

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

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

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

Case 2: Two-state Markov process

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

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

where $s_t \in \{1, 2\}$ follows a two-state observable Markov chain with constant transition probabilities that ensure that endowment growth is i.i.d. Equation 3.5 in the main text gives the cumulant-generating function under the agent's beliefs. The momemt-generating function under the agent's beliefs follows from $\mathcal{M}_t(k) = \exp\left[\mathcal{K}_t(k)\right]$, and Equation 3.18 gives the expected moment-generating function under the econometrician's beliefs. Thus, the asset-pricing quantities under Epstein-Zin preferences are given as above.

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

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

and I find the objective expectation as

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

with

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

We can then insert the expressions into the objectively expected risk premium to derive numerical approximations of the expected risk premium. The following Figure **??** plots the expected risk premium against this period's endowment growth for the baseline parameters as in Table 3.1. As in the previous case with log-normal consumption growth, I find that the objective risk premium is strictly decreasing in this period's consumption growth. Moreover, compared to the objective risk premium, the subjective risk premium is virtually flat.

G Estimation and simulation procedures used in Section 3.4

G.1 Data and estimation

In this section, I describe the data and estimation procedure that I use to obtain the parameters for the simulations. The parameter values are given in Table 3.1.

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

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

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

and set the parameters of these distributions as

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

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

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

Thus, I find the parameters as follows:

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

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

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

such that I find

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

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

Table G.1: Prior parameters for estimation

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

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

Intuitively, the MCMC is solving the conjugate Bayesian posterior with the prior distributions given above. The algorithm iteratively varies the parameters of the model, and computes the log-likelihood of the posterior on the BEA endowment growth data. I then select the model that has the highest log-likelihood over 10,000 iterations, which corresponds to the Bayeisan maximum a posteriori (MAP) estimate. The parameters of the MAP are given in Table 3.1.

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

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

which implies

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

and I consequently run the regression

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

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

G.2 Parameter uncertainty

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

Case 1: Log-normally distributed endowment growth

Let consumption growth g_t be given by

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

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

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

where

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

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

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

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

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

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

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

Thus, I find that

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

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

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

where

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

Shocks to the log SDF are thus

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

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

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

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

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

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

The expected log-return on the consumption-claim is

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

The log risk-free rate is

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

and the risk premium on the consumption-claim is

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

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

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

and the log of the expected return is

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

The risk premium on the consumption-claim is then

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

Next, let us derive objective risk premium by taking the expectation of the return on the consumption-claim. The econometrician knows the true underlying process, such that $\mathbb{E}(R_{c,t+1}) = \frac{1}{\beta}\, e^{\mu + \frac{1}{2}\sigma^2}$ and the objective risk premium is

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

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

Case 2: Markov-process

Let us now consider the Markov process for endowment growth as in Equation 3.1. Closed-form solutions for asset prices with parameter uncertainty cease to exist, such that I detail the numerical procedure to obtain asset pricing equations in this Appendix. The procedure follows Collin-Dufresne et al. (2016).

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

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

with

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

The agent's perceived dynamics of endowment and her beliefs are[44]

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

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.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

as well as the following objective quantity

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

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

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

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

Online Appendix

OA.1 Selective memory for continuous distributions

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

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

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

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

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

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

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

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

experience as

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

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

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

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

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

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

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

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

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

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

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

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

Consequently,

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

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

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

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

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

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

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

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

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

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

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

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

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

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 2.2. If the prior support contains the memory-weighted density, the agent's beliefs will then concentrate on the memory-weighted density.

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

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

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

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

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

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

such that

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

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

OA.2 Extensions

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

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

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

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

The agent learns about both parameters of endowment growth, the mean and the volatility, from her recalled observations. The agent's memory is distorted by the similiarty-weighted memory function in Equation 3.2.

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

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

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

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

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

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

As a next step, I simulate the model $10,000$ times for 304 quarters and report average moments in Table 4.1. The parameters of the endowment growth process are as in Nagel and Xu (2022) with a quarterly mean endowment growth of $\mu = 0.44\%$ and a quarterly volatility of $\sigma = 1.31\%$. All other parameters are as in Table 3.1.

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

OA.2.2 State-dependent similarity-weighted memory

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

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

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

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

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

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

If $s_t = s_\tau$, the memory function is as in Equation 3.2. On the contrary, if $s_t \neq s_\tau$, the memory function is $\exp\left[-\frac{(g_\tau - g_t)^2}{\kappa}\right]$. Under the memory function in Equation .1, the agent is more likely to remember past growth rates that are closer to today's endowment growth rate, and to remember growth rates that occurred in the same state as today's state.

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

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

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

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

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

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

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