Selective Memory Equilibrium*

Drew Fudenberg[†] Giacomo Lanzani[‡]
Philipp Strack[§]

September 15, 2022

Abstract

We study agents who are more likely to remember some experiences than others, but update beliefs as if the experiences they remember are the only ones that occurred. If the agent's behavior converges, their limit strategy is a *selective memory equilibrium*. We illustrate how selective memory equilibrium can be used to understand the long-run effects of several well-documented memory biases, such as ego-boosting bias, associativeness, interference, and extreme experience bias. We then extend our analysis to cases where the expected number of recalled experiences is bounded and experiences that are recalled once are more likely to be recalled again. Here the frequency of recalled experiences does not converge, but we characterize the long-run action frequencies that can arise.

JEL codes: D83, D90

*We thank Chiara Aina, Ian Ball, Abhijit Banerjee, Dirk Bergemann, Jonathan Brownrigg, Simone Cerreia-Vioglio, Gary Charness, Roberto Corrao, Glenn Ellison, Ignacio Esponda, Erik Eyster, Thomas Graeber, Jerry Green, Philippe Jehiel, Mats Koster, Roger Lagunoff, Shengwu Li, Stephen Morris, Ryan Oprea, Parag Pathak, Frank Schilbach, Giorgio Saponaro, Josh Schwartzstein, Andrei Shleifer, Rani Spiegler, Tomasz Strzalecki, Maximilian Voigt, Alex Wolitzky, Muhamet Yildiz, and Sevgi Yuksel for helpful comments and National Science Foundation grant SES 1951056 for financial support.

[†]Department of Economics, MIT
Drew.Fudenberg@gmail.com
77 Massachusetts Avenue, Cambridge MA 02139
Fax: 617-715-4582
[‡]Department of Economics, MIT
[§]Department of Economics, Yale University

1 Introduction

We provide a new conceptual framework for the study of agents who have selective memory in the sense that they are more likely to recall some events than others. We assume that selective memory is stochastic and exogenous, and allow the agent's actions to influence what they observe. In most of the paper, we also assume that agents are unaware of their selective memory, so they update their beliefs as if the experiences they remember are the only ones that occurred. These assumptions fit evidence from both experimental and real-world settings.[1] Although our work is inspired by the neuroscience and psychology literature on memory, we do not try to develop a model that fully matches the process of memory formation and retrieval. Instead, we develop a tractable model of how selective memory shapes behavior and beliefs in decision problems that allows us to analyze situations where the agent's actions determine the distribution of their observations.

Our focus is on selective memory's long-run implications. We show that if an agent's behavior converges, their beliefs converge to a *memory-weighted KL minimizer*, i.e., a distribution that minimizes the Kullback-Leibler ("KL") divergence from a version of the true outcome distribution that gives more weight to realizations that are more likely to be remembered.[2] Moreover, the agent's strategy then converges to a *selective memory equilibrium*, which is a strategy that myopically maximizes their expected payoff against a probability distribution over these minimizers. If all experiences are recalled with the same probability, then memory limitations have no long-run effect. However, if memory is selective and agents are more likely to remember some experiences than others, selective memory can have a persistent effect. For example, an agent who is more likely to recall when they performed well in a task than when they performed poorly will underestimate the task's difficulty and do it too often.

Our framework allows us to analyze the long-run consequences of important and widely documented forms of selective memory such as pleasant memory bias (Mischel,

[1]Memory has been informally described as stochastic since the early stages of the psychology literature, as in James [1890], and recent neuroscience (e.g. Shadlen and Shohamy [2016]) supports this interpretation. Schacter [2008] and Kahana [2012] discuss evidence that some experiences are recalled more often, and e.g. Reder [2014], Zimmermann [2020], Gödker, Jiao, and Smeets [2022] provide evidence of partial or complete unawareness of memory biases.

[2]Berk [1966] showed that when a Bayesian agent observes a sequence of exchangeable outcomes, their beliefs concentrate on the elements of their prior that minimize the Kullback-Leibler divergence with the true distribution. Our results extend this insight to an agent with selective memory whose observations can depend on their actions.

Ebbesen, and Zeiss [1976], Adler and Pansky [2020], and the related ego-boosting bias Zimmermann [2020]),[3] cognitive dissonance (Elkin and Leippe [1986], Chammat et al. [2017], Gödker, Jiao, and Smeets [2022]), associativeness (Thomson and Tulving [1970], Tulving and Schacter [1990], Enke, Schwerter, and Zimmermann [2020], Goetzmann, Watanabe, and Watanabe [2022]), confirmatory bias (Hastie and Park [1986]), interference (Bordalo, Gennaioli, and Shleifer [2019] and Bordalo, Conlon, Gennaioli, Kwon, and Shleifer [2021]), and the overweighting of extreme outcomes (Cruciani, Berardi, Cabib, and Conversi [2011]). In contrast, earlier papers on selective memory each studied a specific form of memory bias, and most only considered short-run effects.

Under positive memory bias, the agent is more likely to recall experiences that induced a larger utility. For example, Zimmermann [2020] finds that subjects who received poor scores on an IQ test are more likely to state that they "cannot recall" their test results, even though that answer is payoff dominated in the experiment, and there were only three things for subjects to recall. Gödker, Jiao, and Smeets [2022] find that investors over-remember positive returns of stocks they invested in, and that their selective memory distorts both their beliefs and their future investment decisions in the direction our model predicts.

We show that positive memory bias can endogenously generate the same long-run behavior as overconfidence in a fixed learning environment. However, we argue that the overconfidence that arises from selective memory is more susceptible to external manipulation through changes in the feedback provided to the agent. For example, coupling negative feedback on one dimension with positive feedback on another will lead to the negative feedback being recalled more often and less bias in long-run beliefs.[4]

Under associativeness, it is easier to recall situations that are similar to the current decision problem, for example, past experiences in which the agent had a similar mood (Matt, Vázquez, and Campbell [1992]). This is a "bias" if it leads the agent to underweight data relative to its true informativeness. The simplest version of associativeness, similarity weighting (Bordalo, Gennaioli, and Shleifer [2020]) does not alter the possible long-run outcomes for a correctly specified agent: We prove that all the selective memory equilibria are self-confirming.

We also study extreme experience bias, which makes experiences with more extreme

[3]Chew, Huang, and Zhao [2020] find that people selectively forget some negative events and create memories of fictitious positive ones.

[4]This practice is suggested in the management literature. See e.g. Procházka, Ovcari, and Durinik [2020].

payoffs more memorable. We then show that moderate risk aversion paired with this bias may explain the extreme risk aversion revealed by the prices of safe and risky assets in financial markets. Moreover, we show that if rarer experiences are more easily recalled the agent overweights small probabilities as assumed in prospect theory.

Selective memory equilibrium resembles Berk-Nash equilibrium (Esponda and Pouzo [2016]), which applies to agents with perfect memory but a misspecified prior. Indeed, we show that every uniformly strict Berk-Nash equilibrium (Fudenberg, Lanzani, and Strack [2021a]) is equivalent to a uniformly strict selective memory equilibrium for some memory function and a full-support prior, and that every uniformly strict selective memory equilibrium is equivalent to a uniformly strict Berk-Nash equilibrium with the appropriate prior support. However, this equivalence fails for Berk-Nash equilibria that are not uniformly strict.[5] In addition, unlike Berk-Nash equilibria, in general, selective memory equilibria do not reduce to self-confirming equilibria (Fudenberg and Levine [1993a]) when the agent is correctly specified. Importantly, the form of misspecification that would lead to the same behavior as a given form of selective memory depends on the environment. That is, particular forms of misspecification and selective memory that coincide under one information structure could lead to very different comparative statics with respect to changes in what the agent observes. We illustrate this point by showing that combining positive and negative feedback has qualitatively different effects on agents that have ego-boosting memory than on agents who are dogmatically overconfident.

In our baseline model, the agent typically recalls more and more periods as time goes on, and with probability 1 the number of recalled experiences converges to infinity, just as they would if the agent had perfect memory. We extend our model to allow for the possibility that the expected number of recalled periods is bounded. Here the agent's beliefs need not converge to a deterministic limit even if the strategy is held fixed, so behavior need not converge either. Instead, we show that whenever the frequency with which strategies are used converges, the limit is a *stochastic memory equilibrium*, meaning that the strategy distribution is generated by a best response to the distribution of memories it generates. We use this to model the effect of "rehearsal," where an experience recalled in one period is more likely to be recalled again. In the Appendix we extend the model to allow for agents who are only partially naïve about their selective memory. We find that if memory is either not selective (in that all

[5] A selective memory equilibrium is uniformly strict if it is the unique best reply to all the beliefs supported on the memory-weighted minimizers.

events are remembered with equal probability) or the agent is not naive (and aware of their selective memory) long-run beliefs will not be distorted and selective memory equilibrium coincides with self-confirming equilibrium.

Related Theoretical Work Mullainathan [2002] studies selective memory where the probability of recalling an observation is the linear sum of a base rate, an "associativeness" term that measures the experience's similarity to the current observation, and a "rehearsal" term that is an indicator for whether the experience was recalled in the previous period. Like us, the paper assumes that agents are naïve about their selective memory. It also assumes that signals are normal and are not influenced by the agent's actions. Afrouzi, Kwon, Landier, Ma, and Thesmar [2020] also studies an agent who is forecasting the next realization of an AR(1) process. It assumes the agent knows the data generating process and chooses which experiences to recall at a cost. Bordalo, Conlon, Gennaioli, Kwon, and Shleifer [2021] considers an agent with a fixed sample size who assesses the relative likelihood of a set of hypotheses by sampling their memory with replacement, and also models the "interference" of some memories with others. None of these papers addresses our question of determining the agent's long-run beliefs and actions.

Schwartzstein [2014] studies the long-run beliefs of an agent whose attention is based on perceived informational value. The agent recalls all of their observations, but naively does not realize that they haven't observed some realizations. As with selective memory and misspecified beliefs, this can lead the agent to make systematically biased forecasts, but it does not encompass biases unrelated to informational content, such as ego-boosting memory. Wilson [2014] and Jehiel and Steiner [2020] study the optimal use of a finite memory is a setting where the agent receives a stream of exogenous signals until they stop and take a single action. Battigalli and Generoso [2021] proposes a formalism to separate assumptions on the players' information and memory. Bénabou and Tirole [2002] studies deliberate memory distortion by a time-inconsistent agent in a two-period model with two signals, and Jehiel [2021] proposes a multi-self solution concept to model "forgetful liars." Further afield, Malmendier and Nagel [2016], Malmendier and Shen [2018], and Malmendier, Pouzo, and Vanasco [2020] consider models where agents apply a weight to events that depends on their age at the time the event happened, Nagel and Xu [2022] analyzes an asset pricing model where the representative agent has fading memory, Bordalo, Gennaioli, and Shleifer [2019] and Bordalo, Burro, Coffman, Gennaioli, and Shleifer [2022] study the effect of memory limitations on rental choice and the COVID pandemic respectively.

Technical contribution Our analysis of finite expected memory extends the Esponda, Pouzo, and Yamamoto [2021a] application of Benaim, Hofbauer, and Sorin [2005]'s stochastic approximation techniques for differential inclusion to settings where beliefs remain stochastic in the limit. To do this, we use the Poisson limit theorem to characterize the distribution of recalled beliefs at each recalled history. To deal with rehearsal we also need to combine previous stochastic approximation techniques with an extension of Doeblin's Theorem to non-homogeneous Markov chains.

2 Setup

We study a sequence of choices made by a single agent. In every period $t = 1, 2, \ldots$ the agent observes a signal s from the finite set S and then chooses an action a from the finite set A. The realized signal s and the chosen action a induce an objective probability distribution $p^*_{a,s} \in \Delta(Y)$ over the finite set of possible outcomes Y.[6] A (pure) strategy is a map $\sigma : S \to A$. The agent's flow payoff is given by $u : S \times A \times Y \to \mathbb{R}$.

We assume that the agent knows the fixed and i.i.d. full-support distribution $\zeta \in \Delta(S)$ over signals.[7] They also know that the map from actions and signals to probability distributions over outcomes is fixed and depends only on their current action and the realized signal, but they are uncertain about the outcome distributions each signal-action pair induces. To model this uncertainty, we suppose that the agent has a prior μ_0 over data generating processes $p \in \Delta(Y)^{A\times S}$, where $p_{a,s}(y)$ denotes the probability of outcome $y \in Y$ when signal s is observed and action a is played.[8] The support of μ_0 is Θ; its elements are the p that the agent initially thinks are possible. The prior is *correctly specified* if its support contains the true data generating process $p^* \in \Theta$; if not, the prior is *misspecified.*

[6]We denote objective distributions with a superscript *.

[7]This assumption lets us focus on our key points. When beliefs about the signal distribution are independent of beliefs about the contingent outcome distributions, the analysis of the paper is unchanged.

[8]For every $X \subseteq \mathbb{R}^k$, we let $\Delta(X)$ denote the set of Borel probability distributions on X endowed with the topology of weak convergence and the associated Prokhorov metric d_p. We also identify each $p \in \Delta(Y)^{A\times S}$ with an element of $\mathbb{R}^{|S|\times|A|\times|Y|}$ and for $x \in \mathbb{R}^k$ we let $||x||_\infty = \max\{|x_1|, \ldots, |x_n|\}$.

Assumption 1. For all $p \in \Theta$, $y \in Y$, $a \in A$, and $s \in S$, $p^*_{a,s}(y) > 0$ if and only if $p_{a,s}(y) > 0$.

This assumption means that the experiences that the agent initially thinks are impossible are objectively impossible, and vice versa.

Recalled Histories We assume that the agent always recalls the signal they just observed. The agent's memory is distorted by a collection of signal-dependent *memory functions* $m_{s'} : S \times A \times Y \to [0,1]$, where $m_{s'}(s,a,y)$ specifies the probability with which the agent remembers a past realization of the signal, action, outcome triplet (s,a,y) when they observe signal s'. We call these triplets *experiences*, and we assume that for each s, s', and a there is at least one y with $m_{s'}(s,a,y) > 0$.[9]

Let $H_t = (S \times A \times Y)^t$ denote the set of all histories of length t, and $H = \cup_t H_t$ the set of all histories. After history $h_t = (s_\tau, a_\tau, y_\tau)_{\tau=1}^t$ and signal s_{t+1}, the *recalled periods* r_t are a random subset of $\{1, ..., t\}$. Period τ with experience (s_τ, a_τ, y_τ) is remembered with probability $m_{s_{t+1}}(s_\tau, a_\tau, y_\tau)$, independently of which other periods are remembered.[10] The *recalled history* is the subsequence of recalled experiences $h_t(r_t) = (s_\tau, a_\tau, y_\tau)_{\tau \in r_t}$.

Beliefs We assume the agent is unaware of their selective memory and naïvely updates their beliefs as if the experiences they remember are the only ones that occurred,[11] so that their belief at $t+1$ equals

$$\mu(C|h_t(r_t)) = \frac{\int_{p \in C} \prod_{\tau \in r_t} p_{a_\tau, s_\tau}(y_\tau) d\mu(p)}{\int_\Theta \prod_{\tau \in r_t} p_{a_\tau, s_\tau}(y_\tau) d\mu(p)} \qquad \forall C \subseteq \Theta. \tag{1}$$

Remark. In our model, the set of recalled histories is not only stochastic but non-monotonic: the agent might remember a past event one day and not another, which

[9] Here we assume that the agent either remembers an experience perfectly or not at all, as opposed to remembering only certain aspects of it, such as one or two components. We relax this in Appendix B.9. We also suppress the possible dependence of the agent's memory function on an exogenous state, though that would be easy to add, and allow us to capture, e.g., the effect of an exogenous endowment on beliefs in Hartzmark, Hirshman, and Imas [2019].

[10] Thus, the distribution over subsets $B \subseteq \{1, .., t\}$ is given by $\mathbb{P}\left[r_t = B | (s_\tau, a_\tau, y_\tau)_{\tau=1}^t, s_{t+1}\right] = \prod_{\tau=1}^t \left(\mathbb{1}_{\tau \in B} m_{s_{t+1}}(s_\tau, a_\tau, y_\tau) + \mathbb{1}_{\tau \notin B}(1 - m_{s_{t+1}}(s_\tau, a_\tau, y_\tau))\right)$.

[11] See, e.g., Reder [2014] for evidence supporting this assumption. In particular, agents often do not make inferences about their forgotten observations from the actions they remember taking. We discuss partial naïveté briefly in the conclusion, and analyze it in Appendix Section A.9.

fits the evidence on memory retrieval Kahana [2012]. Online Appendix B.10 analyses the limit implications of an alternative model where the memory function determines the probability that an experience is recalled in the period just after it occurs. If it is recalled, it is never forgotten, and if it is not recalled, it is never remembered. Because experiences recalled at later dates include all those that were recalled earlier, in this alternative model the agent's past actions don't convey additional information. As with the model we present here, any limit action must be a selective memory equilibrium.

Best Responses and Optimal Policies The agent's belief μ determines the subjective expected utility of each action. Denote by $BR(s, \mu)$ the actions that maximize expected utility when signal s is observed:

$$BR(s, \mu) = \underset{a \in A}{\operatorname{argmax}} \int_{\Theta} \sum_{y \in Y} u(s, a, y) p_{a,s}(y) d\mu(p).$$

A *policy* $\pi : H \to A^S$ specifies a pure strategy for every recalled history. We assume that the agent is myopic and uses an *optimal policy*, i.e., a map $\pi : H \to A^S$ such that for all $s \in S$ and recalled histories $h_t \in H$, $\pi(h_t)(s) \in BR(s, \mu(\cdot|h_t))$.[12]

2.1 Examples

We illustrate our model with five commonly studied examples of memory bias. In this subsection, assumptions about the memory function m hold for all $s, s' \in S, y \in Y$ and $a, a' \in A$.

Example 1 (Utility-Dependent Memory)**.** In some cases the probability of remembering an experience depends on its associated utility, so that $m_{s'}(s, a, y) = \Phi(u(s, a, y))$. Agents who are more likely to remember pleasant experiences correspond to monotone increasing Φ; agents who are more likely to remember extremely high or low utility realizations have Φ that is single-dipped. ▲

Example 2 (Positive Memory Bias)**.** Positive memory bias is the tendency to over-remember experiences that reflect positively on oneself, such as a high test score (see Mischel, Ebbesen, and Zeiss [1976] for early experimental evidence of positive memory bias and Adler and Pansky [2020] for a survey). To model this, we let one dimension

[12] Note that this requires the agent to choose a pure strategy.

y_1 of the outcome y reflect the self-image consequences of the experience, and specify that $m_{s'}(s,a,y) = \Phi(y_1)$ for some increasing Φ. ▲

Example 3 (Cognitive Dissonance and Ex-post Regret)**.** *Cognitive dissonance* is a memory bias where the pleasantness of a memory is relative instead of absolute, because the probability of recalling an experience depends on how well the chosen alternative performed compared to the counterfactual payoff the agent would have received under the ex-post optimal choice (Elkin and Leippe [1986]). This corresponds to $m_{s'}(s,a,y) = \Phi(\max_{a' \in A} u(s,a',y) - u(s,a,y))$ where $\Phi : \mathbb{R}_+ \to [0,1]$ is decreasing. ▲

Example 4 (Associative Memory and Similarity Weighting)**.** To model *associative memory* (Thomson and Tulving [1970]), assume that

$$m_s(s,a,y) > 0 \quad \text{and} \quad \frac{m_s(s,a,y)}{m_s(s',a,y)} > \frac{m_{s'}(s,a,y)}{m_{s'}(s',a,y)},$$

so that a signal is more likely to trigger memories of experiences where the signal was the same. In general, signals represent the conditions under which the choice is made. For example, when in a particular mood, agents tend to recall situations when they were in that mood before (Matt, Vázquez, and Campbell [1992], Mayer, McCormick, and Strong [1995]).[13,14]

A leading special case is *similarity-weighted memory*, where the probability of recalling a past experience only depends on the context in which the choice is taken: Here there is a metric d on the space of signals, and $m_{s'}(s,a,y) = \Phi(d(s,s'))$ for some decreasing function $\Phi : \mathbb{R}_+ \to [0,1]$. For example, professional economic forecasters overweight periods with a macroeconomic context similar to the current one, but only if they lived through them, see Goetzmann, Watanabe, and Watanabe [2022]. ▲

Example 5 (Confirmatory Memory Bias)**.** The agent has *confirmatory memory bias* (see Hastie and Park [1986] for evidence of the relevance of memory for confirmation bias) if they are more likely to remember experiences that the prior deems more likely. Suppose the agent only has two hypothesis, as in Lord, Ross, and Lepper [1979] and Rabin and Schrag [1999], so that $\Theta = \{p^0, p^1\}$, with $\mu(p^0) > \mu(p^1)$. Then confirmatory

[13] See Koszegi, Loewenstein, and Murooka [2021] for an approach that makes mood endogenous.

[14] Jehiel [2018] studies investors who only remember the outcomes of projects that were realized in a certain state and do not remember periods where the project was not realized.

memory bias corresponds to

$$\frac{p^0_{a,s}(y)}{p^1_{a,s}(y)} \geqslant (>) \frac{p^0_{a,s}(y')}{p^1_{a,s}(y')} \implies m_{s'}(s,a,y) \geqslant (>) m_{s'}(s,a,y') \,. \qquad \blacktriangle$$

Example 6 (Interference)**.** *Interference* between memories may lead the agent to be less likely to remember outcomes with frequently observed characteristics than outcomes with some rare characteristics. In Bordalo, Coffman, Gennaioli, Schwerter, and Shleifer [2021] an agent observing colored words and numbers is more likely to remember an orange number if they see more blue words, as then orange stands out more. To model this bias, we can suppress the signals, and set

$$m(a,y) = f(\mathbb{P}[y^1], \mathbb{P}[y^2], \ldots)$$

where outcomes are $y_t = (y^1_t, y^2_t, \ldots)$, $\mathbb{P}[y^k] = \sum_{z\in Y, z^k = y^k} p^*(z)$, and f is decreasing in each component.

3 Long-Run Outcomes

Let $\mathbb{P}_\pi$ denote the probability distribution on H induced by the objective action-contingent probability distribution p^* and the agent's policy π.

Definition 1. A strategy σ is a *limit strategy* if there is an optimal policy π such that $\mathbb{P}_\pi [\sup\{t \colon a_t \neq \sigma(s_t)\} < \infty] > 0$.

In words, a strategy is a limit strategy if there is positive probability that it will be played in every period after some random but finite time. This section gives some general results about limit strategies. These results make no assumptions about the memory function and thus can be used to study the effects of different memory biases. Section 4 then discusses the consequences of specific memory biases.

3.1 Selective Memory Equilibrium

To characterize the strategies that can arise as limit behavior, we define for each strategy σ the set of *memory-weighted KL minimizers* after signal s':

$$KL^{\Theta,m}_{s'}(\sigma) = \underset{p\in\Theta}{\operatorname{argmin}} \left(- \sum_{s\in S} \zeta(s) \sum_{y\in Y} m_{s'}(s,\sigma(s),y) p^*_{\sigma(s),s}(y) \log p_{\sigma(s),s}(y) \right) . \qquad (2)$$

These are the elements of Θ that maximize the log-likelihood of the memory-weighted outcome distribution induced by σ. Note that only the relative sizes of the weights m matter for determining $KL_{s'}^{\Theta,m}(\sigma)$: if $\hat{m}(\cdot) = \lambda m(\cdot)$ for some $\lambda > 0$ then $\hat{m}$ and m have the same memory-weighted KL minimizers.

Definition 2. A strategy σ is a

1. *Selective memory equilibrium* if for all $s \in S$ there is $\nu_s \in \Delta(KL_s^{\Theta,m}(\sigma))$ such that $\sigma(s) \in BR(s, \nu_s)$.

2. *Uniformly strict selective memory equilibrium* if for all $s \in S$ and all $\nu \in \Delta(KL_s^{\Theta,m}(\sigma))$, $\{\sigma(s)\} = BR(s, \nu)$.

In a selective memory equilibrium σ, the action played after each signal s is a best reply to some belief over memory-weighted KL minimizers from the distribution of experiences generated by σ. The uniformly strict version adds the restriction that there is the same unique best reply for each of these minimizers. Both concepts allow the actions played in response to different signals to be justified by different beliefs because which memories are triggered depends on the current realization of the signal.

Theorem 1. *Every limit strategy is a selective memory equilibrium.*

The idea of this theorem is that when the agent plays a fixed strategy, the empirical distribution converges, and the distribution of recalled experiences after each signal converges to a deterministic limit where the best-fitting models after each signal s are the memory-weighted KL minimizers $KL_s^{\Theta,m}(\sigma)$.

The theorem is proved by contradiction. We fix a strategy σ that is not a selective memory equilibrium and observe that there is at least one signal after which σ does not prescribe a best reply to the memory weighted KL-minimizers. A compactness-continuity lemma guarantees that this also holds for beliefs concentrated on a neighborhood of the minimizers. So it is enough to show that beliefs concentrate on these neighborhoods. A preliminary lemma shows it is sufficient to obtain this concentration under the counterfactual (and possibly suboptimal) policy where the agent always plays σ. Under such policies, the empirical frequency after each signal s converges to the distribution given by $p^*_{s,\sigma(s)}$, and the Borel-Cantelli lemma implies that almost surely the recalled history is large and representative. With this, we can extend Berk [1966]'s concentration result to the product of actual and recalled experiences to show that under representative recalled histories, the distributions that don't minimize the memory weighted KL divergence have vanishing posterior probability.

Proposition 10 in the Online Appendix establishes a partial converse: every uniformly strict selective memory equilibrium is stable, meaning that play converges to it with arbitrarily high probability for an open set of beliefs.

Theorem 1 provides a learning foundation for some equilibrium concepts that have been used in recent work. For example, Koszegi, Loewenstein, and Murooka [2021] proposes an equilibrium concept where the agent is more likely to remember successes than failures if they are in a good mood, and the agent's mood is determined by their self-esteem, which is a function of the number of past successes they remember. This is a case of our model where the agent's mood is an action chosen to match their perceived probability of succeeding at a task (i.e., their perceived ability). Our equilibrium concept then coincides with Koszegi, Loewenstein, and Murooka [2021]'s "self-esteem personal equilibrium," and Theorem 1 shows that any long-run learning outcome must be such an equilibrium.

Berk-Nash equilibrium is another example of an equilibrium concept for which we can provide a learning foundation based on selective memory. For example, Section 4.2 shows that positive memory bias can lead to overconfidence. Overconfidence has been modeled as the result of exogenous misspecification; the fact that it can be endogenously derived from a well-documented memory bias provides a micro-foundation for Berk-Nash equilibrium in this context. More generally, Proposition 6 shows that any Berk-Nash equilibrium can be micro-founded through selective memory, though how economically convincing such a micro-foundation is depends on the specific application. Finally, Section 5.1 shows that our model of selective memory leads to the same long-run predictions as underinference (Phillips and Edwards [1966]) and can thus micro-found it.

4 Specific Forms of Selective Memory

4.1 Similarity-Weighted Memory and Self-Confirming Equilibrium

Definition 3. Strategy σ is a *self-confirming equilibrium* if there is $\nu \in \Delta(\Theta)$ such that for all $s \in S$ and $p \in \operatorname{supp} \nu$, $p_{\sigma(s),s} = p^*_{\sigma(s),s}$ and $\sigma(s) \in BR(s, \nu)$.

Fudenberg and Levine [1993b] shows that self-confirming equilibria correspond to the steady states of a learning model with long-lived but myopic Bayesian agents; it requires that agents have correct beliefs about the consequences of their equilibrium action but allows them to have incorrect beliefs about actions they do not use.

Proposition 1. *For a correctly specified agent with similarity-weighted memory (Example 4), a strategy is a selective memory equilibrium if and only if it is a self-confirming equilibrium.*

This result follows from the fact that since $m_{s'}(s, a, y)$ does not depend on a or y, when the agent is correctly specified and has an accurate sample, the true distribution is the best fit for *every* signal, so the weight assigned to each signal does not matter. However, similarity weighting can change the set of selective memory equilibria when the agent is misspecified. Whether this occurs depends on how the agent thinks the outcome distribution varies with signals. For example, suppose the agent thinks that the outcome distributions associated with different signals are independent. In that case, the similarity is irrelevant whether or not the agent is correctly specified: The set of selective memory equilibria coincides with the set of Berk-Nash equilibria with or without this bias. At the other extreme, the agent might think that the distribution of outcomes is the same for all signals. In this case, similarity weighting can change the set of selective memory equilibria, as shown by Example 11 in the Online Appendix.[15]

4.2 Ego-Boosting Memory Bias and Overconfidence

It is well established that many people are more likely to recall situations that reflect positively on themselves.[16] This can lead to a particular kind of pleasant memory bias: they are more likely to remember experiences that boost their self-assessment than those that reduce it.

Consider a situation where the agent observes i.i.d. outcomes $y_t \in Y \subset \mathbb{R}$ that reveal information about an ego-relevant characteristic such as IQ or the ability to be a successful investor. There are no signals, A is endowed with a linear order, and the agent (correctly) believes that their action does not affect the realized outcome. The

[15] Also, even when there is a unique selective memory equilibrium, and it is objectively optimal, the speed of convergence to the equilibrium can be influenced by similarity weighting. This is similar to kernel density estimation, where the optimal bandwidth trades off having enough observations with relying too much on distant values.

[16] See, e.g. Mischel, Ebbesen, and Zeiss [1976].

next proposition shows that in this case, a larger bias leads to a more positive limit belief and higher limit action. This provides a memory foundation for the positive correlation between an agent happiness and the inaccuracy of their beliefs documented in Alloy and Abramson [1979].

Proposition 2. *Suppose that m, m' and p^* are constant in a, $m'(a,y) = f(y)m(a,y)$ for some increasing function f, $u(a,y)$ is supermodular, and that $\Theta = \Delta(\Delta(Y))$. The agent's long-run belief with memory m' concentrates on a distribution of outcomes higher in first-order stochastic dominance than the distribution under the long-run belief with memory m, and the limit action with memory m' will be higher than the limit action with memory m.*

Intuitively, because the prior has full support, the KL minimizer will be the distribution that exactly matches what the agent remembers. The agent's selective memory makes this recalled history more favorable than the true one, and because the agent's utility function is supermodular their limit action is higher than the objective optimum.

Example 7. Suppose that each period the agent takes an action $a \in \{0,1\}$, with $u(a,y) = a(y-z)$, $z \in (0,1)$. Here y is the outcome of an IQ test, which is either pass, $y = 1$, or fail, $y = 0$, so $a = 1$ is optimal if and only if the probability of passing the test exceeds z. The agent passes the test with probability p^*. They always recall passed IQ tests, and they recall failed tests with probability ϕ:

$$m(a,y) = \begin{cases} 1 & \text{if } y = 1 \\ \phi & \text{if } y = 0 \end{cases}.$$

In the long run the agent believes that the probability of passing an IQ test is

$$p = \frac{p^*}{\underbrace{p^*}_{\text{Successes}} + \underbrace{(1-p^*) \times \phi}_{\text{Failures}}} = p^* + \frac{p^*(1-p^*)}{\phi/(1-\phi) + p^*}.$$

For example, if the true probability p^* is .5, and the agent remembers failing an IQ test with probability .8, in the long run they believe that they pass the test with probability .556. Consequently, they will behave like an exogenously misspecified agent who dogmatically believes their ability to pass is at least .556. Moreover, the difference between p and p^* is monotonic in the agent's selectivity bias ϕ. ▲

This example relates to an experiment by Zimmermann [2020] in which subjects took an IQ test and received three noisy observations of how well they performed relative to other subjects. Zimmermann [2020] finds that all subjects can recall the signals immediately after observing them, but subjects who received negative feedback were less likely to recall the feedback a month later than subjects who received positive feedback: subjects are roughly 20% more likely to state that they "cannot recall" the result of the IQ test if the feedback was negative, even though that answer is payoff dominated in the experiment and there were only three things for subjects to try to recall.[17] Thus, here, selective memory is a better explanation for long-run overconfidence than selective attention.

Example 7 and Proposition 2 also relate to the literature on overconfidence and financial decision-making. Walters and Fernbach [2021] finds investors are 10% less likely to recall an investment that led to a loss compared to an investment that led to a gain. Moreover, selective forgetting predicts overconfidence, and overconfidence is reduced when investors rely less on memory. In an incentivized experiment, Gödker, Jiao, and Smeets [2022] finds that subjects over-remember good investment outcomes and under-remember bad investment outcomes. In line with the prediction of Proposition 2, this leads subjects to have overly optimistic beliefs about their investments and reinvest in bad investments more often.

Ego Boosting Bias and Misattribution We next show how an agent with ego-boosting bias can misinterpret data about other aspects of the world.

Example 8. Suppose that besides taking an IQ test, the agent works on a group project with a coworker. The outcome distributions $(p_1, q) \in [0,1]^2$ and outcome $(y_1, y_2) \in \{0,1\}^2$ are two dimensional, where the first component denotes whether or not the agent passed an IQ test and the second component denotes whether a group project succeeded. The agent passes the IQ test with probability p_1, and the group project succeeds with probability $p_2 = \alpha p_1 + (1-\alpha)q$, so the success of the group project depends on the ability of the agent p_1 and their coworker's ability q. The agent always remembers experiences with positive IQ test results, and remembers experiences with negative test results with probability $\phi \in (0,1)$. If (p_1^*, q^*) denote the true ability of

[17] Zimmermann [2020] finds that "negative feedback is indeed recalled with significantly lower accuracy, compared to positive feedback." Here lower accuracy means both that the agents are more likely to report that they do not recall the experience and that they misreport the experience.

the agent and the coworker then the agent's long-run beliefs will satisfy

$$p_1 = p_1^* + \frac{p_1^*(1-p_1^*)}{\phi/(1-\phi)+p_1^*} \qquad q = q^* - \frac{\alpha}{1-\alpha}\frac{p_1^*(1-p_1^*)}{\phi/(1-\phi)+p_1^*}\,.$$

The agent thus underestimates the coworker's ability, and the underestimation grows as memory becomes more selective. ▲

To generalize this example, we consider a two dimensional outcome space $Y = Z \times Z \subset \mathbb{R}^2$, where y_1 corresponds to an ego-relevant characteristic, and is distributed according to p_1^*. The second component y_2 is independently drawn, with $p_2^*(y_2) = \alpha p_1^*(y_2) + (1-\alpha)q^*(y_2)$ for some $\alpha \in (0,1)$. The agent knows that the outcomes are independently drawn each period according to a p^* satisfying these conditions, but does not know p_1^* or q^*, and their prior belief assigns positive probability to each of these distributions.

Proposition 3. *Suppose that m and p^* are constant in a, and y_2 and m is increasing in y_1. Then the agent's long-run belief about p_1 concentrates on a distribution that is higher in first-order stochastic dominance than p_1^* and the agent's long-run belief about p_2 concentrates on a distribution that is lower than p_2^*.*

This result provides an explanation for the Gilovich [2008] finding that "a survey of college professors found that 94% thought they were better than their average colleague".

Reinforcement through Actions We note that actions can play an important role in amplifying the misconceptions caused by selective memory. For example, suppose an agent starts out with an unbiased belief about their coworker's ability, and decides what fraction α of work to delegate to them. As the agent over-remembers their own successes, they become overconfident about their own ability, and to explain the disappointingly low frequency of successes in the group project, they became overly pessimistic about the coworker's. The agent thus delegates delegate less work to their coworker, whose ability thus has a smaller effect on output. Thus, to explain the disappointingly low output the agent becomes even more pessimistic about the coworkers ability, leading to even less delegation.

The induced long-run beliefs and effect of actions are similar to the attribution bias derived in Heidhues, Kőszegi, and Strack [2018] for exogenously overconfident agents, where low outcomes are attributed to an exogenous state. More generally, Section 5.2 shows in any fixed environment, that the long-run belief induced by selective memory can always be replicated by exogenous misspecification, and vice versa.

Despite this equivalence in a fixed environment, selective memory and exogenous misspecification can lead to very different predictions about the effect of changes in information. statics. Suppose for example that negative feedback is delivered along with positive feedback about an unrelated trait of the agent. Combining positive and negative information in this way is known as a "feedback sandwich;" it is suggested to make feedback more impactful in the management and psychology literature.[18] If the positive feedback makes the experiences with failed IQ tests less unpleasant, an agent with positive memory bias would be more likely to remember them, and their long-run belief would move closer to their actual ability. So they would become less biased about their coworker's ability. This is in contrast to exogenous misspecification, where positive feedback about an unrelated state would not affect the agent's beliefs about their own or their coworker's ability.

4.3 Extreme Experience Bias and Risk Attitudes

This section shows that for choices over lotteries, memory distortions can generate the same behavior as a distorted risk preference. We again simplify by supposing there are no signals, and let the outcome $y \in \mathbb{R}$ be the amount of money received by the agent, with $u(s, a, y) = v(y)$ for some concave v.

Extreme Experience Bias Suppose the agent chooses between a safe action $a = 0$ that induces outcome y_0, and a symmetric risky lottery $a = 1$ with expected value $\bar{y}$. We say that the agent has an *extreme experience bias* if the probability of remembering an experience m is an increasing function of the distance of the outcome y from its expected value, i.e.:

$$m(s, a, y) = h(|y - \bar{y}|) \tag{3}$$

for some increasing $h : \mathbb{R}_+ \to \mathbb{R}_+$.

Our next result, together with Theorem 1 shows that choosing the lottery is the long-run outcome with extreme experience bias only if it is the long-run outcome with perfect memory. Moreover, Example 12 in the Online Appendix shows that extreme

[18] Procházka, Ovcari, and Durinik [2020] describes an experiment where bundling negative feedback with positive feedback about an unrelated domain helps agents perform better.

experience bias can shift the long-run outcome from the lottery to the safe action. To state a result that holds for all concave utility functions, we assume that the true distribution of outcomes is symmetric.[19]

Proposition 4. *Suppose p_1^* is symmetric*[20] *and that the agent thinks all outcome distributions are possible under the risky action.*[21] *If choosing the lottery is not a self-confirming equilibrium, it is not a selective memory equilibrium with extreme experience bias.*

Because the agent over-remembers extreme experiences, the environment seems riskier than it truly is, so in the long run, they do not take the risky action if it would not be optimal for an agent without extreme experience bias.[22] By making the tail realizations relatively more memorable, extreme bias can make a risk-averse agent act as if they were even more risk-averse. This may help explain why the risk aversion needed to match the real-world investment choices is unrealistically high: the agents can be attracted by safe alternatives because they are moderately risk-averse, and their memory exaggerates the riskiness of the uncertain alternatives. For example, a single day where the stock market crashed might be more easily remembered than many days of average returns and lead to a biased perception of its riskiness. Indeed, the plausibility of this channel is supported by several studies that show that higher working memory is associated, either directly or through a proxy measure of cognitive ability, with lower risk aversion at both the intra- and interpersonal levels (see, e.g., Cokely and Kelley [2009], Boyle, Yu, Buchman, and Bennett [2012], and Benjamin, Brown, and Shapiro [2013]).

Rare Experience Bias Similarly, some forms of selective memory are equivalent to preferences that arise from distorting outcome probabilities. Suppose that the agent is more likely to remember experiences that happen more rarely, i.e., there is a decreasing function $h : [0,1] \to [0,1]$ such that $m(s, a, y) = h(p_1^*(y))$. In this case, in the long run the agent believes that the outcome distribution for the risky action is

$$\frac{h(p_1^*(y))}{\sum_{z \in Y} h(p_1^*(z))}.$$

[19] Extreme-experience bias can have the opposite effect of encouraging risk-taking behavior when the true distribution is very asymmetric, with a very low probability of a large payoff, as with lotteries.

[20] $p_1^*(\bar{y} + \delta) = p_1^*(\bar{y} - \delta)$ for all $\delta \in \mathbb{R}$.

[21] Thus $\Theta = \{p \in \Delta(Y) \times \Delta(Y) : p_0(y_0) = 1\}$.

[22] This behavior can also be induced by a misspecified belief with overly fat tails.

They will thus act as if they distort probabilities, as in prospect theory (Kahneman and Tversky [1979]).[23]

5 Alternative Models

This section compares our selective memory model with *underinference* and *misspecification*, which are two other ways to model similar effects.

5.1 Underinference

The phenomenon of underinference (Phillips and Edwards [1966]) is distinct from selective memory but has similar long-run implications, as we establish in Proposition 5. Here agents remember (or are presented with) a record of past observations, so memory is not an issue, and the agent's beliefs are a deterministic function of the sequence of observations. However, they underweight a given observation (s, a, y) when applying Bayes rule. In particular, they use the deterministic updating rule

$$\mu^U(C|(s_i, a_i, y_i)_{i=1}^t) = \frac{\int_{p\in C} \prod_{i=1}^t (p_{a_i,s_i}(y_i))^{m(s_i,a_i,y_i)} d\mu(p)}{\int_\Theta \prod_{i=1}^t (p'_{a_i,s_i}(y_i))^{m(s_i,a_i,y_i)} d\mu(p')}, \tag{4}$$

where $m(s, a, y) \in [0, 1]$ is the *underinference distortion* applied to experience (s, a, y).

As with selective memory, this memory distortion leads beliefs to concentrate on a set of weighted KL minimizers, and as the next result shows the underinference distortion maps directly to a selective memory function.

Proposition 5. *If σ is a limit strategy with underinference distortion m, it is a selective memory equilibrium with memory function m.*

A leading special case is *uniform underinference* where $m(s, a, y) = c < 1$ and the agent discounts all observations by the same amount. In this case, Propositions 1 and 5 imply that the limit strategy for a correctly specified agent must be a self-confirming equilibrium. A natural question is whether the two theories can be distinguished from observable data. Most of the current evidence on overconfidence only regards agent

[23] Rare experience bias is related to interference, as under both less likely outcomes are remembered more easily; one is about the overall frequency of an outcome and the other about the frequencies of its components.

beliefs and not the recalled histories, which makes it difficult to tell the two distortions apart (see Benjamin [2019]).

If signals are absent and actions are real-valued, the way actions respond to outcomes can be used to distinguish underinference and selective memory. Under overconfidence, the realization of y_t is sufficient to predict whether a_{t+1} is more or less than a_t. Under selective memory, the set of past experiences retrieved at time $t+1$ may differ from those at time t, so in general the previous period's outcome and action are not sufficient to predict how actions change. Moreover, the action sequence features a sort of regression to the mean: after a particularly high action, the next action will likely be lower.

In general, with an exogenous data generating process, the agent's beliefs will converge to the same limit as with underinference, so their limit action will be the same. If the data generating process is endogenous, random memory realizations can induce switches in actions, so actions are less likely to converge. The following example illustrates this possibility.

Example 9. Suppose $A = \{a, b\}$, $Y = \{0, 1\}$, $S = \varnothing$, $u(a, y) = y$, and that the agent knows that the probability of $y = 1$ given action a is some $c \in (0, 1)$, i.e. $p_a(1) = p^*_a(1) = c$ for all $p \in \Theta$. The agent does not know the probability of outcome 1 under action b. Their initial belief is that it is larger than that of action a, so $BR(\mu_0) = b$, although there is $p' \in \Theta$ with $p'_b(1) < c$. The truth is that $1 > p^*_b(1) > c$, so action b is optimal, but if m is strictly positive, both a and b are selective memory equilibria. In the underinference model both a and b have positive probability of being limit actions, and if $a_t = a$ then $a_\tau = a$ almost surely for all $\tau > t$, while if $m(b, 1) \geqslant m(b, 0)$ with the selective memory model, then $\lim_{t\to\infty} a_t = b$ almost surely. ▲

More generally, selective memory does not generate as much long-run inefficiency as underinference: Whenever the agent believes that the consequences of different actions are independent, if the expected utility of a selective memory equilibrium a under the memory-weighted KL minimizer is lower than the ex-ante value of an alternative b, then a is not a limit strategy.

5.2 Selective Memory and Misspecification

We can relate the long-run implications of selective memory to those of *misspecification* in the sense of the statistics literature, where the true model is not in the support of the agent's prior. Let

$$KL^{\Theta,\mathbf{1}}(\sigma) = \operatorname*{argmin}_{p\in\Theta} \left(-\sum_{s\in S} \zeta(s) \sum_{y\in Y} p^*_{\sigma(s),s}(y) \log p_{\sigma(s),s}(y) \right) \tag{5}$$

be the set of KL minimizers for Θ when the function m is identically equal to 1, so the agent does not have selective memory.

Definition 4. A strategy σ is a

1. *Berk-Nash equilibrium* if for all $s \in S$, there exists $\nu \in \Delta(KL^{\Theta,\mathbf{1}}(\sigma))$ such that $\sigma(s) \in BR(s,\nu)$.

2. *Uniformly strict Berk-Nash equilibrium* if for all $\nu \in \Delta(KL^{\Theta,\mathbf{1}}(\sigma))$ and all $s \in S$, $\{\sigma(s)\} = BR(s,\nu)$.

Esponda and Pouzo [2016] shows that Berk-Nash equilibrium is a necessary condition for a strategy to be the long-run outcome of a possibly misspecified learning process. Fudenberg, Lanzani, and Strack [2021a] shows that uniformly strict Berk-Nash equilibrium is a necessary condition for a strategy to be the long-run outcome with probability near 1.

Berk-Nash equilibrium is the special case of selective memory equilibrium where memory is perfect as opposed to selective. Moreover, even beyond this case there is a close relationship between the uniformly strict versions of these equilibrium concepts: For given prior support Θ and objective distribution p^*, every uniformly strict Berk-Nash equilibrium is equivalent to a selective memory equilibrium with full-support prior for some memory function, and every uniformly strict selective memory equilibria is equivalent to a Berk-Nash equilibrium for some support. To formalize this idea, we will need to consider two different priors μ_0 and μ_0' with possibly different supports Θ and Θ'.

Definition 5. A Berk-Nash equilibrium σ with support Θ and a selective memory equilibrium σ' with support Θ' are *belief equivalent* if $\sigma = \sigma'$, and for all $s \in S$ there exists a belief $\nu \in \Delta(KL^{\Theta,\mathbf{1}}(\sigma) \cap KL_s^{\Theta',m}(\sigma))$ such that $\sigma(s) \in BR(s,\nu)$.

Two equilibria are belief equivalent if they prescribe the same strategies, and behavior after each signal can be justified by the same belief. Note that in a selective memory equilibrium, the signal can influence the agent's memory and the set of KL minimizers, while in a Berk-Nash equilibrium, the KL minimizers are the same regardless of the signal.

Proposition 6.

1. *Every uniformly strict Berk-Nash equilibrium with support Θ is belief equivalent to a selective memory equilibrium with support $\Theta' = \Delta(Y)^{A\times S}$ for some memory function.*

2. *Every uniformly strict selective memory equilibrium with support Θ is belief equivalent to a uniformly strict Berk-Nash equilibrium for some Θ'.*

The idea behind the first part of the proposition is that if we start from a KL minimizer p with perfect memory but incomplete support, we can choose a memory function that rescales the probability of an outcome by $p_{\sigma(s),s}(y)/p^*_{\sigma(s),s}(y)$, making the recalled frequency equal to p. This makes p a weighted-memory minimizer, so σ is a best reply.[24] The second part of the proposition is trivial: to construct a strict Berk-Nash equilibrium that leads to the same beliefs and behavior as in the selective memory equilibrium, we can endow the agent with a degenerate belief that equals the long-run belief in the selective memory equilibrium.[25]

Remark 1. As we prove in Online Appendix B.4, the uniform strictness conditions of Proposition 6 are needed:

1. There are Berk-Nash equilibria that are not belief equivalent to any selective memory equilibrium with support $\Theta' = \Delta(Y)^{A\times S}$.

2. There are selective memory equilibria that are not belief equivalent to any Berk-Nash equilibrium.

3. Moreover, unlike Berk-Nash equilibria, selective memory equilibria do not reduce to self-confirming equilibria when the agent is correctly specified. In particular, selective memory equilibria need not be objectively optimal when the distribution of outcomes is independent of the agent's action.

[24] Every p'' that is outcome-equivalent under σ is also a minimizer, and these p'' may not have been elements of Θ. Because σ need not be a best response to some of them, it need not be a uniformly strict selective memory equilibrium.

[25] Formally, we take Θ' to be a singleton p with $p_{a,s}(y) = p'_{a,s}(y)$ where $p' \in KL_s^{\Theta,m}(\sigma)$.

To illustrate the equivalence result, consider a buyer who submits an offer for a good in a double-blind two-sided auction where the price z is set at the buyer's bid, so the seller's dominant strategy is to bid their value. Suppose that the buyer has an exogenously fixed conviction that the price sellers ask is independent of the quality of the good they are selling. If the buyer's value of the good is $x + v + \varepsilon$ where x is the value for the seller, v measures the gains from trade, and ε is a noise term, and if after every interaction, the buyer observes their payoff and the ask price x, then in the Berk-Nash equilibrium they submit a bid that is too low, as in Esponda [2008]. Proposition 6 shows that memory distortions can, over time, lead the agent to believe that value and bid are independent and thus have the same long-run behavior and beliefs. This is obtained with a memory function that gives more weight to experiences with a larger gap between buyer's values and ask prices.[26]

While Proposition 6 implies that selective memory and misspecification will have similar long-run implications in a *fixed* environment, in Section 4.2 we have shown that the two models have different comparative statics with respect to changes in the environment.[27] Therefore, empirical work that distinguishes between the two models based on variation in information is needed.

6 Finite Expected Memory

6.1 Stationary Bounded Memory

In our base model, the number of recalled experiences goes to infinity as the agent's sample size increases. This section modifies the model so that even when the actual sample size goes to infinity, the expected number of instances recalled by the agent remains bounded. This provides an explanation of why time-changing behavior may persist in apparently stationary environments that complements explanations based on recency bias (see Cheung and Friedman [1997], Fudenberg and Levine [2014], Fudenberg and Peysakhovich [2016], and Erev and Haruvy [2016]).

We suppress the signals for simplicity, and assume that the agent's memory at time $t + 1$ is distorted through a memory function that now depends on calendar time: For

[26] Note also that since memory cannot affect the perceived relative likelihoods of experiences that are not observed, and only some distributions for the counterfactual justify the equilibrium strategy, the resulting equilibrium may not be uniformly strict.

[27] Selective memory can arguably be viewed as a form of misspecification, as the agent is not aware of their memory limitations. From that perspective, our results show that the classic misspecification studied in Bayesian statistics is closely related to a psychologically-founded form of misspecification.

some fixed integer k,

$$m_{t+1}(a,y) = \min\{1, k/t\}\, m(a,y). \tag{6}$$

With this specification, as the number of experiences t grows to infinity, the probability that any specific experience is recalled decreases at rate k/t. Thus asymptotically the expected number of recalled experiences is no more than k, so even when the action process converges, the recalled outcome frequency can have a non-degenerate distribution.

After history $h_t = (a_i, y_i)_{i=1}^t$, the *recalled periods* r_t are a random subset B of $\{1, .., t\}$ with probability distribution

$$\mathbb{P}\left[r_t = B | h_t\right] = \prod_{i=1}^{t} \left(\mathbb{1}_{i \in B} m_{t+1}(a_i, y_i) + \mathbb{1}_{i \notin B}(1 - m_{t+1}(a_i, y_i))\right). \tag{7}$$

The *recalled history* is the subsequence of action-outcome pairs $h_t(r_t) = (a_i, y_i)_{i \in r_t}$ of the recalled periods, and μ_{t+1} is the random period-$t+1$ belief $\mu(\cdot | h_t(r_t))$ induced by the recalled history $h_t(r_t)$.

To simplify notation let $\eta_\alpha \in \Delta(\mathbb{N}^{A \times Y})$ be the product distribution where the marginal for the action outcome pair (a, y) is Poisson distributed with parameter $\alpha(a) p_a^*(y) k m(a, y)$. We are going to show that η_α is the limit distribution of the recalled experiences if the action frequencies converge to α. Intuitively, the expected number of times a pair (a, y) is recalled is proportional to the frequency of the action $\alpha(a)$, the probability of the outcome given the action $p_a^*(y)$, and how memorable that experience is, i.e., $m(a, y)$. Let F_α be the distribution of beliefs induced by η_α, i.e., $F_\alpha(B) = \eta_\alpha\left(\{h : \mu(\cdot|h) \in B\}\right)$ for all $B \subseteq \Delta(\Theta)$.[28]

Definition 6. A *stochastic memory equilibrium* is a mixed strategy α for which there is a Markovian mixed policy $\rho : \Delta(\Theta) \to \Delta(A)$ such that

1. at every belief μ the action is optimal, i.e., $\rho(\mu) \in BR(\mu)$, and

2. α equals the action frequencies induced by ρ, i.e., $\alpha = \mathbb{E}[\rho(\mu) | \mu \sim F_\alpha]$.

In words, the action distribution α in a stochastic memory equilibrium is characterized by a fixed point condition: there is an optimal Markovian policy that maps the

[28]Note a small abuse of notation: η_α defines a probability measure over the number of times each experience is recalled, and not over the set of histories, as histories with the same experiences in a different order are distinct. But all such histories induce the same posterior, so the definition of F_α is unambiguous.

distribution over beliefs F_α induced by η_α to α.

For every $t \in \mathbb{N}$, define the *action frequency* at time t by

$$\alpha_t(h_t)(a) = \frac{1}{t}\sum_{\tau=1}^{t} \mathbb{1}_a(a_\tau)$$

for all $h_t \in H_t$ and for all $a \in A$.

Theorem 2. *If $\alpha_t \to \alpha$ with positive probability, then α is a stochastic memory equilibrium.*

The proof shows that if play converges to a limit that is not a stochastic memory equilibrium, the distribution of actions would not correspond to the distribution of best responses generated by the agent's beliefs. We first show that when the α_t converge to α, the distribution of recalled experiences is the η_α defined above, which follows from the Poisson limit theorem on the sum of binomials. We then show that beliefs converge to F_α, and use the Benaim, Hofbauer, and Sorin [2005] extension of stochastic approximation to differential inclusions to show that if α is not a stochastic memory equilibrium, the agent's best response to their beliefs would lead the distribution of actions to move away from α.

In a stochastic memory equilibrium, the agent will sometimes rely on a small number of past instances to make decisions. This can induce long-run underreaction of beliefs and insensitivity to sample size, a form of representativeness bias first documented by Kahneman and Tversky [1972]. We can calibrate the memory capacity parameter k to match various aspects of the evidence. For example, the probability that the agent decides only on the basis of their prior decreases exponentially in k, and is equal to:

$$\prod_{(a,y)\in A\times Y} \exp(-\alpha(a)p_a^*(y)\, k\, m(a,y)).$$

There is a similar formula for the probability of making a decision based on at most seven experiences, which is claimed to be the "magical" number for working memory (Miller [1956]), so we can calibrate the model by choosing k that makes the probability of choosing based on 7 or fewer memories close to 1. Moreover, a falsifiable implication of the model is that the decrease in the probability of recalling an additional t-th experience decreases factorially regardless of the value of k.

The Role of the Prior under Finite and Infinite Expected Memory One important difference between finite and infinite expected memory is the role of the prior. To see this in the starkest form, we compare limit behavior with a selective memory function $\hat{m}$ and that in a stochastic memory equilibrium with memory $\hat{m}_t = \min\{1, k/t\}\hat{m}$. We first consider the case of an exogenous data generating process where the action does not influence the outcome distribution.

Corollary 1. *Fix $\Theta \subseteq \Delta(Y)^A$ and $m : A \times Y \to [0,1]$. Suppose that for all $p \in \Theta$, m, p^*, p do not depend on actions and there exists a unique uniformly strict selective memory equilibrium $\hat{a}$:*

1. *When the agent has infinite expected memory, $\mathbb{P}_\pi\left[\sup\{t\colon a_t \neq \hat{a}\} < \infty\right] = 1$ for every optimal policy π.*

2. *For every action a that is not weakly dominated in Θ and every expected memory capacity $k \in \mathbb{N}$, there exists a prior belief with support Θ such that if $\alpha_t \to \alpha$ with positive probability then $\alpha(a) > 0$.*

When the data generating process is exogenous and memory is unbounded, the empirical distribution of recalled outcomes converges almost surely, and the agent ends up playing the best reply to this distribution. With finite memory, there is a positive fraction of periods in which agent recalls so little that they play a best reply to their prior. Since weakly undominated actions are best replies to some prior, the result follows. More generally, when the action does influence the distribution of outcomes, the prior may influence the probability of converging to a specific selective memory equilibrium, but the set of selective memory equilibria is the same for every prior. This is not the case with bounded memory, because the prior influences the chosen action when the number of recalled experiences is small.

6.2 Rehearsal

We expand the model to incorporate the effect of *rehearsal*: If an experience is recalled in one period, it is more likely to be recalled in subsequent periods, see Kandel et al. [2000] and the references therein. To model this phenomenon we assume that the agent's memory at time $t+1$ is distorted through a *rehearsal memory function* that can depend on the history recalled in the previous period

$$m_{t+1}(a, y|(r_t, (a_t, y_t))) = \min\{1, k/t\}\left(m(a, y) + R\mathbb{1}_{(a,y)\in r_t \cup \{(a_t,y_t)\}}\right), \qquad (8)$$

where the experiences $r_t \cup \{(a_t, y_t)\}$ that were recalled or experienced last period have an additional probability of $0 \leqslant R \leqslant 1 - \max_{a \in A, y \in Y} m(a, y)$ of being recalled, and $R = 0$ reduces to equation (6).[29]

We now define a Markov chain over histories, where the expected number of times a pair (a, y) is recalled is proportional to the frequency of action, the probability of the outcome given the action, how memorable that realization is, and whether it occurred or was recalled in the last period.

Definition 7. For every $t \in \mathbb{N}$ and $h' = (a_i, y_i)_{i=1}^t$, let $\eta_{\alpha,h'}$ be the distribution where the number of occurrences of each action-outcome pair (a, y) has a Poisson distribution with parameter

$$\begin{array}{cl} \alpha(a)p_a^*(y)\, k\, [m(a, y) + R] & \text{if } \exists i \in \{1, ..., t\}\colon\ (a, y) = (a_i, y_i) \\ \alpha(a)p_a^*(y)\, k\, [m(a, y) + Rp_a^*(y))] & \text{if } \nexists i \in \{1, ..., t\}\colon\ (a, y) = (a_i, y_i) \text{ and } a = \pi(h') \\ \alpha(a)p_a^*(y)\, k\, m(a, y) & \text{otherwise} \end{array}$$

independent of the number of the instances for the other pairs. The *induced Markov chain* η_α has state space H and Markov kernel $\eta_{\alpha,h'}(h)$.

We will show that this Markov chain admits a unique stationary distribution, and that this distribution is the limit time-average distribution over recalled histories. Indeed, at any time every subhistory of what is currently recalled has positive probability of being the next recalled history, so the chain is irreducible on the subsets of histories that can be reached with positive probability starting from the empty history. A calculation in the appendix shows it is positive recurrent, which yields the following lemma.

Lemma 1. *η_α admits a unique stationary distribution $\mathcal{H}_\alpha \in \Delta(H)$.*

Let $F_{\alpha,h'}$ be the distribution of beliefs induced by $\eta_{\alpha,h'}$, and let $\chi_{\alpha,h'}$ denote the distributions over actions induced by an optimal Markovian mixed policy ρ and random beliefs μ:

$$\chi_{\alpha,h'} = \{\alpha' \in \Delta(A) : \exists \rho \in \Pi_o \text{ such that } \alpha' = \mathbb{E}[\rho(\mu) | \mu \sim F_{\alpha,h'}]\},$$

[29] Thus the recalled periods r_t are distributed as in (7), with $m_{t+1}(\cdot|h_t)$ replaced by $m_{t+1}(\cdot|r_t, h_t)$. Rehearsal is a key feature of Mullainathan [2002]'s model of memory, which analyzed the period-ahead impact of rehearsal but not its long-run implications. To make it easier to compare our results, we too assume that only what was recalled yesterday affects what is recalled today.

where Π_o is the set of all measurable selections from the correspondence $\mu \mapsto \Delta(BR(\mu))$.

Definition 8. An *ergodic memory equilibrium* is a mixed strategy α such that there exists $(\alpha^h)_{h \in H} \in \Delta(A)^H$ with $\alpha^h \in \chi_{\alpha,h}$ for all $h \in H$ and $\alpha = \mathbb{E}_{\mathcal{H}_\alpha}[\alpha^h]$.

Ergodic memory equilibria are fixed points: for every recalled history $h \in H$, a mixed action α determines a probability distribution over what is recalled next period, and thus through the Markovian policy determines a mixed action α^h. Ergodic memory equilibrium requires that the expectation of α^h with respect to the induced stationary distribution over recalled histories is α.

Theorem 3. *If $\alpha_t \to \alpha$ with positive probability, then α is an ergodic memory equilibrium.*

6.3 Income Forecasts and Asset Pricing

Our model of finite expected memory and rehearsal lets us generalize the findings of Mullainathan [2002] about income forecasts beyond the specific parametric structure it assumed. It also lets us provide a novel memory-based explanation of the equity-premium and equity-volatility puzzles. We suppose that the outcome y_t is i.i.d. $y_t = \theta + \epsilon_t$, independent of the action of the agent, where the ϵ_t are mean-0 shocks.[30]

Memory Rehearsal and correlated prediction errors The rehearsal memory function of equation (8) generates the same predictions about one-period correlations as Mullainathan [2002], without assuming associativeness. First, a high outcome last period triggers memories of equally high past realizations, so that the forecasting error will be negatively correlated with the most recent information.[31] Second, when the baseline probability of remembering an event is low, and the rehearsal effect is strong, the forecast errors in successive periods will be positively correlated for the same reason as in as Mullainathan [2002]: memories that are forgotten are more likely to be forgotten again.

[30] Both Mullainathan [2002] and Weitzman [2007] assumed that the outcomes follow an AR1 process, so that a standard Bayesian would always place non-vanishing weight on the most recent outcome; our assumption of finite expected memory has the same implication.

[31] Mullainathan [2002] supposes that y has a positive density on the real line, so that some form of associativeness is needed for rehearsal to have any effect.

Asset pricing Suppose that the agent can choose between purchasing a safe asset a_0 and purchasing a (representative) equity portfolio a_1 at prices $p_{0,t}$ and $p_{1,t}$. If held one period, the safe asset has net return i, while the risky asset provides net return $i + \theta$, where θ is a random variable whose distribution is unknown. In this setting, the equity premium puzzle is that if the distribution of θ were known and equal to that observed in the data, a very large amount of risk aversion would be needed to justify the observed difference in asset prices.

Weitzman [2007] explains this with the combination of an overly pessimistic prior and the assumption that the agent believes θ changes over time, so they discard old observations. Ergodic memory equilibrium predicts the same effect even with a perceived constant risk premium. Theorem 3 guarantees that even in the long run the agent will rely on a limited number of observations, so that the pessimistic prior is able to sustain the premium.[32]

Moreover, if we add signals back to the model and assume extreme experience bias and similarity-weighted memory, our model predicts excess volatility with respect to fundamentals as in the equity-volatility puzzle. As an example, suppose the signal s is a mean-zero shift to the outcome distribution, and the similarity-weighted memory takes the form $d(s, s') = 1$ if s' and s have the same sign and 0 otherwise, and that the extreme experience bias has the form assumed in (3), with reference point the return of the low asset, i.e., $1 + i$. Similarity-weighted memory implies that after a negative signal the agent is more likely to sample negative realizations, and extreme experience bias implies that the most negative realizations will be the most remembered, inducing an excessively low willingness to pay for the asset. The same qualitative feature holds if the similarity between negative signals is high compared to the correlation of different-sign signals. Here the prediction of excess volatility is not obtained from recency alone: If memory is limited but not selective, the prediction is underreaction to signals.

7 Discussion

This is the first paper to explore the long-run implications of selective memory. Our equilibrium concept and results make it easy to predict the long-run implications of arbitrary memory biases, which should be of broad use in applied work. We illustrated

[32] Of course, as the size of the average number of recalled events grows, the premium shrinks, just as the premium in Weitzman [2007] shrinks as the fundamental's rate of change goes to 0.

our framework by showing that it explains how overconfidence can arise from an ego-boosting memory bias, and why agents may underestimate their co-workers' abilities even when they are correctly specified. It also allows to explain the excessive levels of risk aversion implied by asset choice as the result of only moderate risk aversion paired with an extreme experience bias that leads agents to overestimate the riskiness of the assets.

Distinguishing Between Models To distinguish between selective memory and underinference, one can elicit agent's beliefs.[33] Underinference predicts that the likelihood ratio between two data generating processes θ and θ' always increases between period $t-1$ and t if the period $t-1$ outcome was more likely under θ. Selective memory allows for violations of this monotonicity, especially if at the beginning of period t a signal triggering experiences favoring θ' is observed, while this signal is irrelevant with underinference. A more direct way to distinguish selective memory from other sources of mistaken inference, including misspecification, is to elicit both what the agent remembers and what they believe, as in Gödker, Jiao, and Smeets [2022]. This allows one to estimate the memory function and qualitatively distinguish between selective memory, misspecification, and underinference.

Partial naïveté We have assumed that agents treat the experiences they remember as if these were the only ones that happened. The Online Appendix relaxes this assumption, and considers agents who are partially aware of their memory limitations.[34] To do this we assume that agents know calendar time, and therefore how many observations they have not remembered. We show how the degree of naïveté can amplify existing memory biases in the case of ego-boosting memory bias. For an agent who is aware of their own forgetfulness, but not aware that their memory is selective, the selective memory equilibria under partial and full naïveté coincide. At the other extreme, if agents are fully aware of their memory function and correctly specified, any action that is optimal for the true data generating process is always a selective memory equilibrium. It is only the *combination* of naïveté and the selective nature of memory that distort the agent's long-run beliefs.

[33]However, see Danz, Vesterlund, and Wilson [2022] for practical challenges in elicitating beliefs accurately.

[34]To focus on the learning implications of selective memory, we restrict to myopic agents. Without myopia, one has to take a stance on the commitment power of the agent and on how they try to affect their future memory.

Possible Extensions There are several natural extensions that could be pursued in future work. One is to generalize i.i.d. signals to a Markov process. This would let us capture some other relevant biases; for example, the gambler's fallacy (see Rabin and Vayanos [2010] and He [2022]) would arise if a signal is more memorable when it is different than the signal in the previous period.[35] Or it might be much easier for agents to recall whether an experience happened at all than whether it happened five or six times; we could capture this by using a memory function that is concave in the number of times an experience occurred. Another generalization would be to memory functions with recency bias, such as $m_{s',t}(s_\tau, a_\tau, y_\tau) = m_{s'}(s_\tau, a_\tau, y_\tau) f(t-\tau)$ where f is a decreasing function. As with associative memory, when the outcomes are exogenous this bias only leads to slower learning, but when actions are endogenous it can prevent the agent from locking on to the optimal action.

A Appendix

A.1 Preliminaries

We begin by stating a few lemmas whose proofs are in the Online Appendix. For every $h_t \in H$ let $f(h_t) \in \Delta(S \times A \times Y)$ denote the empirical distribution over signals, actions, and outcomes in history h_t, and for every nonempty recalled history r_t let

$$\hat{f}(h_t, r_t)(s, a, y) = \frac{1}{|r_t|} \sum_{i=1}^{t} \mathbb{1}_{(s_i, a_i, y_i)=(s,a,y)} \mathbb{1}_{i \in r_t}$$

denote the recalled empirical distribution in history h_t when the recalled periods are r_t. For all $C \subseteq \Theta$ equation (1) can be rewritten as

$$\begin{aligned}\mu(C|h_t(r_t)) &= \frac{\int_{p\in C} \prod_{(s,a,y)} (p_{a,s}(y))^{\sum_{i=1}^{t} \mathbb{1}_{(s_i,a_i,y_i)=(s,a,y)} \mathbb{1}_{i\in r_t}} \mathrm{d}\mu}{\int_{p\in\Theta} \prod_{(s,a,y)} (p_{a,s}(y))^{\sum_{i=1}^{t} \mathbb{1}_{(s_i,a_i,y_i)=(s,a,y)} \mathbb{1}_{i\in r_t}} \mathrm{d}\mu} = \frac{\int_{p\in C} \prod_{(s,a,y)} (p_{a,s}(y))^{|r_t| \hat{f}(h_t,r_t)(s,a,y)} \mathrm{d}\mu}{\int_{p\in\Theta} \prod_{(s,a,y)} (p_{a,s}(y))^{|r_t| \hat{f}(h_t,r_t)(s,a,y)} \mathrm{d}\mu} \\ &= \frac{\int_{p\in C} \exp\left(|r_t| \sum_{(s,a,y)} \log(p_{a,s}(y)) \hat{f}(h_t, r_t)(s, a, y)\right) \mathrm{d}\mu}{\int_{p\in\Theta} \exp\left(|r_t| \sum_{(s,a,y)} \log(p_{a,s}(y)) \hat{f}(h_t, r_t)(s, a, y)\right) \mathrm{d}\mu}. \qquad (9)\end{aligned}$$

Also, for every $\gamma \in \Delta(S \times A \times Y)$ and $p \in \Delta(Y)^{A\times S}$ let $D(\gamma||p) = -\sum_{(s,a,y)} \gamma(s, a, y) \log(p_{a,s}(y))$.

[35] This extension could make use of the analysis of belief concentration for misspecified agents with Markov models developed in Fudenberg, Lanzani, and Strack [2021b].

Lemma A.1. *For all Borel measurable $C, C' \subseteq \Delta(Y)^{A\times S}$, $t \in \mathbb{N}$, $h_t \in H_t$ and $r_t \subseteq \{1, ..., t\}$,*

$$\frac{\mu(C|h_t(r_t))}{1-\mu(C'|h_t(r_t))} \geqslant \frac{\mu(C)}{1-\mu(C')} \exp\left(|r_t| \left[\inf_{p\in\Theta\setminus C'} D(\hat{f}(h_t, r_t)||p) - \sup_{p\in C} D(\hat{f}(h_t, r_t)||p)\right]\right).$$

Let $KL_s^{\Theta,m}(\sigma, \varepsilon) = \{p \in \Theta : \exists q \in KL_s^{\Theta,m}(\sigma), ||p-q||_\infty \leqslant \varepsilon\}$. The next lemma shows that posteriors that assign sufficiently high probability to these balls around the minimizers induce the expected behavior: actions that are not selective equilibria are not best replies, and makes a uniformly strict selective equilibrium the unique best reply.

Lemma A.2. *If σ is not a selective memory equilibrium, there are $s' \in S$ and $\varepsilon, C \in \mathbb{R}_{++}$ such that*

$$\frac{\nu(KL_{s'}^{\Theta,m}(\sigma,\varepsilon))}{1-\nu(KL_{s'}^{\Theta,m}(\sigma,\varepsilon))} > C \implies \sigma(s') \notin BR(s', \nu).$$

If σ is a uniformly strict selective memory equilibrium, there are $\varepsilon, C \in \mathbb{R}_{++}$ such that for all $s \in S$

$$\frac{\nu(KL_{s}^{\Theta,m}(\sigma,\varepsilon))}{(1-\nu(KL_{s}^{\Theta,m}(\sigma,\varepsilon))} > C \implies \{\sigma(s)\} = BR(s, \nu).$$

For any $t \in \mathbb{N}$, strategy σ and sequence of actions a^t let π^{σ,a^t} be the decision rule that prescribes action a_τ at period $\tau \leqslant t$ and action $\sigma(s_\tau)$ at all periods $\tau > t$.

The next lemma shows that if σ is a limit strategy, then for each time t, there is an action sequence a^t such that if the agent plays a^t in the first t periods and then σ afterwards there is positive probability that the induced sequence of beliefs makes σ optimal at all periods $\tau \geqslant t+1$.

Lemma A.3. *If for every t, every sequence of actions a^t, and every optimal policy $\tilde{\pi}$, $\mathbb{P}_{\pi^{\sigma,a^t}}[\sigma(s_\tau) = \tilde{\pi}(h_\tau(r_\tau), s_\tau)$ for all $\tau \geqslant t] = 0$ then σ is not a limit strategy.*

For every $\sigma \in A^S$ and $s' \in S$ let

$$M_\sigma(s') = \min_{p\in\Theta}\left(-\sum_{s\in S}\zeta(s)\sum_{y\in Y} m_{s'}(s,\sigma(s),y)p^*_{\sigma(s),s}(y)\log p_{\sigma(s),s}(y)\right)$$

denote the divergence of the weighted memory minimizers under strategy σ.

A.2 Theorem 1

Proof of Theorem 1. Suppose towards a contradiction that σ is a limit strategy under the optimal policy π, but not a selective memory equilibrium. Then by Lemma A.2 there are $s' \in S$ and $\varepsilon, C \in \mathbb{R}_{++}$ such that if $\frac{\nu(KL_{s'}^{\Theta,m}(\sigma,\varepsilon))}{1-\nu(KL_{s'}^{\Theta,m}(\sigma,\varepsilon))} > C$ then $\sigma(s') \notin BR(s',\nu)$. Let h_t be a history with positive probability under the optimal policy π. We show that if the agent plays σ at every period after h_t almost surely the belief μ_t reaches a region where no optimal policy prescribes $\sigma(s')$ after signal s'. By Lemma $A.3$ this is enough to obtain the desired conclusion.

By the strong law of large numbers, the frequency $f(h_\tau)(s,a,y)$ converges a.s. on h_t to

$$\lim_{\tau\to\infty} f(h_\tau)(s,a,y) = \begin{cases} \zeta(s)p^*_{a,s}(y) & \text{if } a=\sigma(s) \\ 0 & \text{otherwise} \end{cases}.$$

Define $\tilde{p}(\sigma,s') \in \Delta(S\times A\times Y)$ to be the induced distribution over remembered experiences

$$\tilde{p}(\sigma,s')(s,a,y) = \begin{cases} \frac{\zeta(s)m_{s'}(s,\sigma(s),y)p^*_{\sigma(s),s}(y)}{\sum_{\hat{y},\hat{s}}\zeta(\hat{s})m_{s'}(\hat{s},\sigma(\hat{s}),\hat{y})p^*_{\sigma(\hat{s}),\hat{s}}(\hat{y})} & \text{if } a=\sigma(s) \\ 0 & \text{otherwise} \end{cases}.$$

Since for every two periods $\tau' > \tau$ and $B \subseteq \{1,...,\tau'\}$ the probability of recalling B at time τ' conditional on the actual history $h_{\tau'}$ is independent of the history r_τ recalled at period τ, i.e., $\mathbb{P}[r_{\tau'}=B|h_{\tau'}] = \mathbb{P}[r_{\tau'}=B|h_{\tau'},r_\tau]$, by the second Borel-Cantelli lemma (see, e.g., Theorem 8.3.4 in Dudley [2018]), for every $\varepsilon \in \mathbb{R}_{++}$, $s' \in S$ and $k \in \mathbb{N}_{++}$ almost surely there is a $\tau > t$ such that $s_\tau = s'$, $|r_\tau| > k$ and $||\hat{f}(h_\tau,r_\tau) - \tilde{p}(\sigma,s')||_\infty < \varepsilon$. We show that eventually $\frac{\nu(KL_{s'}^{\Theta,m}(\sigma,\varepsilon))}{1-\nu(KL_{s'}^{\Theta,m}(\sigma,\varepsilon))} > C$ on the histories where these two conditions are satisfied. Since they hold almost surely, the result follows.

Let $\varepsilon', \kappa \in \mathbb{R}_{++}$ be such that

$$\kappa < \inf_{\{p' \notin KL_{s'}^{\Theta,m}(\sigma,\varepsilon)\}} \left(-\sum_{s\in S}\zeta(s)\sum_{y\in Y} p^*_{\sigma(s),s}(y)m_{s'}(s,\sigma(s),y)\log p'_{\sigma(s),s}(y)\right) - M_\sigma(s')$$

and

$$\frac{\kappa}{2} > \sup_{\{p' \in KL_{s'}^{\Theta,m}(\sigma,\varepsilon')\}} \left(-\sum_{s\in S}\zeta(s)\sum_{y\in Y} p^*_{\sigma(s),s}(y)m_{s'}(s,\sigma(s),y)\log p'_{\sigma(s),s}(y)\right) - M_\sigma(s').$$

So, by Lemma A.1

$$\frac{\mu(KL_{s'}^{\Theta,m}(\sigma,\varepsilon)|h_\tau(r_\tau))}{1-\mu(KL_{s'}^{\Theta,m}(\sigma,\varepsilon)|h_\tau(r_\tau))}$$
$$\geqslant \frac{\mu\left(KL_{s'}^{\Theta,m}(\sigma,\varepsilon)\right)}{1-\mu\left(KL_{s'}^{\Theta,m}(\sigma,\varepsilon')\right)} \exp\left(|r_\tau|\left(\begin{array}{c}\inf_{\{p\notin KL_{s'}^{\Theta,m}(\sigma,\varepsilon)\}} D(\hat{f}(h_\tau,r_\tau)||p)\\ -\sup_{\{p\in KL_{s'}^{\Theta,m}(\sigma,\varepsilon')\}} D(\hat{f}(h_\tau,r_\tau)||p))\end{array}\right)\right).$$

The last expression goes to $+\infty$ as $|r_\tau|\to\infty$, as by the definitions of κ and ε'

$$\begin{aligned}
&-\lim_{t\to\infty} \sup_{\{p\in KL_{s'}^{\Theta,m}(\sigma,\varepsilon')\}} -\sum_{(s,a,y)} \hat{f}(h_\tau,r_\tau)\,(s,a,y)\log(p_{a,s}(y))\\
&\qquad+\lim_{t\to\infty} \inf_{\{p\notin KL_{s'}^{\Theta,m}(\sigma,\varepsilon)\}} -\sum_{(s,a,y)} \hat{f}(h_\tau,r_\tau)\,(s,a,y)\log(p_{a,s}(y))\\
&=-\sup_{\{p\in KL_{s'}^{\Theta,m}(\sigma,\varepsilon')\}} -\sum_{s\in S}\zeta(s)\sum_{y\in Y} m_{s'}(s,\sigma(s),y)p^*_{\sigma(s),s}(y)\log p'_{\sigma(s),s}(y)\\
&\qquad+\inf_{\{p\notin KL_{s'}^{\Theta,m}(\sigma,\varepsilon)\}} -\sum_{s\in S}\zeta(s)\sum_{y\in Y} m_{s'}(s,\sigma(s),y)p^*_{\sigma(s),s}(y)\log p_{\sigma(s),s}(y) > \frac{\kappa}{2} > 0\,.\square
\end{aligned}$$

A.3 Proposition 1

Proof of Proposition 1. We show that only data generating processes p for which $p_{\sigma(s),s} = p^*_{\sigma(s),s}$ are memory-weighted KL-minimizers after signal s.

Suppose that p is such that for some $\hat{s}\in S$, $p_{\sigma(\hat{s}),\hat{s}} \neq p^*_{\sigma(\hat{s}),\hat{s}}$. By the Gibbs inequality,

$$\sum_{y\in Y} p^*_{\sigma(s),s}(y)\log p^*_{\sigma(s),s}(y) \geqslant \sum_{y\in Y} p^*_{\sigma(s),s}(y)\log p_{\sigma(s),s}(y)$$

for all $s\in S$, with a strict inequality for $s=\hat{s}$. This, together with $d(\hat{s},\hat{s})=0$ and $f(0)>0$, implies that

$$\begin{aligned}
\sum_{s\in S}\zeta(s)\sum_{y\in Y} m_{\hat{s}}(s,\sigma(s),y)p^*_{\sigma(s),s}(y)\log p_{\sigma(s),s}(y) &= \sum_{s\in S}\zeta(s)f(d(s,\hat{s}))\sum_{y\in Y} p^*_{\sigma(s),s}(y)\log p_{\sigma(s),s}(y)\\
&< \sum_{s\in S}\zeta(s)f(d(s,\hat{s}))\sum_{y\in Y} p^*_{\sigma(s),s}(y)\log p^*_{\sigma(s),s}(y)\\
&= \sum_{s\in S}\zeta(s)\sum_{y\in Y} m_{\hat{s}}(s,\sigma(s),y)p^*_{\sigma(s),s}(y)\log p^*_{\sigma(s),s}(y)
\end{aligned}$$

proving that $p\notin KL_{\hat{s}}^{\Theta,m}(\sigma)$. $\square$

A.4 Proposition 2

Proof of Proposition 2. We first derive the long-run belief for $\tilde{m} \in \{m, m'\}$. Because the memory function $\tilde{m}$ and the probability distribution over outcomes p^* are independent of the agent's action, we suppress the dependence of p and $\tilde{m}$ on a.

Because $\Theta = \Delta(\Delta(Y))$, for every σ the KL minimizer is the distribution:

$$p^{\tilde{m}}(y) = \frac{\tilde{m}(y)p^*(y)}{\sum_{z \in Y} \tilde{m}(z)p^*(z)} .$$

Thus $p^{m'}(y) = w(y)p^m(y)$ where $w(y) = f(y)\frac{\sum_{z \in Y} m(z)p^*(z)}{\sum_{z \in Y} m'(z)p^*(z)}$ is non-decreasing, so the function

$$z \mapsto \sum_{x \leqslant z} (p^{m'}(x) - p^m(x)) = \sum_{x \leqslant z} p^m(x)(w(x) - 1)$$

is quasi-convex. As it equals 0 for $z < \min_{y \in Y} y$ and for $z \geqslant \max_{y \in Y} y$, it is non-positive for $z \in [\min_{y \in Y} y, \max_{y \in Y} y]$, so $p^{m'}$ dominates p^m in first-order stochastic dominance. Every limit action must be optimal given $p^{\tilde{m}}$ for $\tilde{m} \in \{m, m'\}$ by Theorem 1, so the agent's action must be higher under m' than under m. $\square$

A.5 Proposition 3

Proof of Proposition 3. Because (y_1, y_2) are subjectively independent conditional on the value of p_1, the learning problem decouples across the two dimensions. By Proposition 2 the long-run belief about p_1 is higher than the true distribution p_1^*. The probability with which an outcome is remembered is independent of the second component, so the agent learns p_2^*. They infer q to be $q(y_2) = \frac{p_2^*(y_2) - \alpha p_1(y_2)}{1-\alpha}$. Thus $q - q^* \equiv \frac{\alpha}{1-\alpha}(p_1^* - p_1)$, and as p is greater than p^* in first-order stochastic dominance, it follows that q is lower than q^* in first-order stochastic dominance. $\square$

A.6 Selective Memory Equilibrium and Misspecification

Proof of Proposition 6. To prove part (1), let $\sigma \in A^S$ be a uniformly strict Berk-Nash equilibrium, and let p' be an arbitrary element of $KL^{\Theta,\mathbf{1}}$. Since σ is a uniformly strict Berk-Nash equilibrium, for all $s \in S$, $\{\sigma(s)\} = BR(s, \delta_{p'})$. Moreover, by assumption (1), $p^*_{\sigma(s),s}(y) = 0$ implies $p'_{\sigma(s),s}(y) = 0$, so[36] $M := \max_{(s,y) \in S \times Y} \frac{p'_{\sigma(s),s}(y)}{p^*_{\sigma(s),s}(y)} < \infty$.

[36]We use the convention that $0/0 = 0$.

Define $\tilde{m}$ by $\tilde{m}_{s'}(s,a,y) = \frac{p'_{a,s}(y)}{Mp^*_{a,s}(y)}$. Then for an agent with a full-support prior and memory function $\tilde{m}$ the memory-weighted KL minimizers for strategy σ after signal s' are the elements of

$$\underset{p\in\Delta(Y)^{A\times S}}{\operatorname{argmin}} -\sum_{s\in S}\zeta(s)\sum_{y\in Y}\tilde{m}_{s'}(s,\sigma(s),y)p^*_{\sigma(s),s}(y)\log p_{\sigma(s),s}(y)$$

$$= \underset{p\in\Delta(Y)^{A\times S}}{\operatorname{argmin}} -\sum_{s\in S}\zeta(s)\sum_{y\in Y}\frac{p'_{\sigma(s),s}(y)}{M}\log p_{\sigma(s),s}(y) = \underset{p\in\Delta(Y)^{A\times S}}{\operatorname{argmin}} -\sum_{s\in S}\zeta(s)\sum_{y\in Y}p'_{\sigma(s),s}(y)\log p_{\sigma(s),s}(y).$$

Thus p' minimizes the KL divergence for all $s' \in S$, so σ is a selective memory equilibrium with a full-support prior.

Part (2), the converse direction, is trivial: simply take Θ' to be a singleton p such that for all $a \in A$ and $s \in S$, $p_{a,s'}(y) = p'_{a,s'}(y)$ for some $p' \in KL^{\Theta,m}_{s'}(\sigma)$. □

A.7 Theorem 2

The proof of Theorem 2 builds on the techniques of Esponda, Pouzo, and Yamamoto [2021a]. The key complication is that even when the action process converges beliefs remain stochastic, so we lack a counterpart to their Theorem 1. Instead, we prove a lemma on the convergence of the belief distribution that, paired with our assumption of expected utility maximization, lets us mimic the other steps of their proof. Let $(d_t)_{t\in\mathbb{N}}$ be a sequence of empirical joint distributions over actions and outcomes.

Lemma A.4. *If for some $\alpha \in \Delta(A)$, $d_t(a,y)$ converges to $\alpha(a)p^*_a(y)$ for all $(a,y) \in A \times Y$, then the distribution of μ_t given h_{t-1} weakly converges to $F_\alpha \in \Delta(\Delta(\Theta))$, and F is continuous.*

Proof. This is a special case of the more general A.5, which is proved below. □

Next, we write the empirical action frequency at time $t+1$ as a function of its expected error and a mean-0 error term:

$$\begin{aligned}\alpha_{t+1}(h_{t+1})(a) &= \alpha_t(h_t)(a) + \frac{1}{t+1}(\delta_{a_{t+1}}(a) - \alpha_t(h_t)(a)) \\ &= \alpha_t(h_t)(a) + \frac{1}{t+1}\left(\mathbb{E}[\delta_{a_{t+1}}(a)|\mu_{t+1}] - \alpha_t(h_t)(a)\right) \end{aligned} \tag{10}$$

where the second equality follows from the fact that the agent uses a stationary pure policy. Note that equation (10) is a particular case of the differential inclusion

$$\alpha_{t+1}(h_{t+1})(a) \in \left\{ \alpha_t(h_t)(a) + \frac{1}{t+1} \left(\delta_{a'}(a) - \alpha_t(h_t)(a) \right) : a' \in BR(\mu_{t+1}) \right\}.$$

The continuous-time interpolation of α_t is the function $w : \mathbb{R}_+ \to \Delta(A)$

$$w(\tau_t + c) = \alpha_t + c\frac{\alpha_{t+1} - \alpha_t}{\tau_{t+1} - \tau_t}, \quad \forall c \in \left[0, \frac{1}{t+1}\right], \tag{11}$$

where $\tau_0 = 0$ and $\tau_t = \sum_{i=1}^{t} \frac{1}{i}$ for all $t \in \mathbb{N}$.

Proof of Theorem 2. Let χ_α be the distribution of actions induced by the distribution of beliefs F_α and some best response selection $\rho \in \Pi_o$:

$$\chi_\alpha = \{\alpha' \in \Delta(A) : \exists \rho \in \Pi_o \text{ such that } \alpha' = \mathbb{E}_{F_\alpha}(\rho(\mu))\}.$$

We will use the results of Benaim, Hofbauer, and Sorin [2005] to show that (11) can be approximated by a solution to

$$\dot{\boldsymbol{\alpha}}_t \in \chi_{\boldsymbol{\alpha}_t} - \boldsymbol{\alpha}_t. \tag{12}$$

A solution to (12) with initial point $x^* \in \Delta(A)$ is a mapping $x : \mathbb{R}_+ \to \Delta(A)$ that is absolutely continuous over compact intervals, with $x(0) = x^*$, and (12) satisfied for almost every t. Since by Lemma A.4 F is upper hemicontinuous, a solution exists. Let $X_{x^*}^T$ be the set of solutions to (12) over $[0, T]$ with initial conditions x^*.

Now we show that the continuous-time interpolation of α defined in (11) can in the long run be approximated arbitrarily well by a solution to (12). Define the random variable $U_t = \alpha_{t+1} - \tilde{U}_t$, where $\tilde{U}_t$ is an arbitrary measurable selection from $\operatorname{argmin}_{\alpha' \in \chi_{\alpha_t}} ||\alpha_{t+1} - \alpha'||$. Since both χ_{α_t} and α_{t+1} are uniformly bounded, U_t is uniformly bounded. Moreover, by Lemma A.4 and the definition of χ_{α_t}, U_t converges almost surely to 0, so condition (i) of Proposition 1.3 in Benaim, Hofbauer, and Sorin [2005] is satisfied. Condition (ii) is also satisfied because $||\alpha_{t+1} - \alpha_t||_\infty < 1/(t+1)$, w is Lipschitz continuous of order 1, and α_t is uniformly bounded because it takes values in $\Delta(A)$, so w is a perturbed solution of (12). Thus by Theorem 4.2 in Benaim, Hofbauer, and Sorin [2005],[37]

$$\lim_{t\to\infty} \inf_{\tilde{\boldsymbol{\alpha}}\in X^T_{w(t)}} \sup_{0\leqslant s\leqslant T} ||w(t+s) - \tilde{\boldsymbol{\alpha}}(s)|| = 0 \quad \mathbb{P}_\pi\text{-almost surely for all } T \in \mathbb{N}. \tag{13}$$

If $\alpha \in \Delta(A)$ is not a stochastic memory equilibrium, since χ_α is convex by the separating hyperplane theorem there exists $f \in \mathbb{R}^A$ with $\alpha \cdot f > \max_{\bar{\alpha}\in\chi_\alpha} \bar{\alpha} \cdot f$. Let $K = \alpha \cdot f - \max_{\bar{\alpha}\in\chi_\alpha} \bar{\alpha} \cdot f$. By Lemma A.4, there exists $\varepsilon \in \mathbb{R}_{++}$ such that for all $\alpha' \in B_\varepsilon(\alpha)$, $\max_{\bar{\alpha}\in\chi_{\alpha'}} \bar{\alpha} \cdot f < \max_{\bar{\alpha}\in\chi_\alpha} \bar{\alpha} \cdot f + K/4$ and $\alpha \cdot f > \alpha \cdot f - K/4$. Therefore, for every initial condition $\alpha^* \in B_\varepsilon(\alpha)$ and every solution in $X^T_{\alpha^*}$, $\alpha \cdot f$ decreases at rate at least $K/2$ until the solution leaves $B_\varepsilon(\alpha)$. So there exists $T \in N$ such that for every initial condition $\alpha^* \in B_\varepsilon(\alpha)$ and every solution in $X^T_{\alpha^*}$, the differential inclusion leaves $B_\varepsilon(\alpha)$ by time T.

We conclude the proof by combining an argument similar to Proposition 1 of Esponda, Pouzo, and Yamamoto [2021a] with equation (13) to rule out convergence to a non-equilibrium point. We will prove that α_t does not converge to α on any path $(h_t)_{t\in\mathbb{N}}$ where (13) applies. Since the set of such sample paths has probability 1, α_t can only converge to stochastic memory equilibria.

Let $\hat{T} \in \mathbb{N}$ be such that on the chosen path $(h_t)_{t\in\mathbb{N}}$

$$\inf_{\tilde{\boldsymbol{\alpha}}\in X^T_{w(\hat{T})}} \sup_{0\leqslant s\leqslant T} ||w(\hat{T}+s) - \tilde{\boldsymbol{\alpha}}|| \leqslant \varepsilon/2. \tag{14}$$

If there is no $t > \hat{T}$ such that $w_t \in B_{\varepsilon/2}(\alpha)$, α_t does not converge to α. But if $w_t \in B_{\varepsilon/2}(\alpha)$ for some $t > \hat{T}$, then the differential inclusion leaves $B_\varepsilon(\alpha)$ by time $T + t$, and by (14), α_t does not stay in $B_{\varepsilon/2}(\alpha)$. This proves Theorem 2. □

A.8 Theorem 3

Proof of Lemma 1. Let H' denote the set of histories that can be reached with positive probability starting from the empty history. At any recalled history $h = (a_i, y_i)_{i=1}^n \in H'$, the probability of a transition to the empty history is bounded below by $Q := \prod_{(a,y)\in A\times Y} \exp(-\alpha(a)p^*_a(y)\, k\, [m(a,y) + R]) > 0$, so H' is a closed communicating class. Moreover, for any $h' \in H'$ there is $\tau \in \mathbb{N}$ such that the probability of a simple

[37]The proof of Theorem 4.2 in Benaim, Hofbauer, and Sorin [2005] invokes an implication of their Theorem 4.1 that is not correct. However, the statement is correct, as shown in Part 3 of the proof of Theorem 1 in Esponda, Pouzo, and Yamamoto [2021b].

path of length τ from the empty history to h' is some $M > 0$. Therefore the expected time of return to h' is bounded from above by $\tau + \sum_{i=\tau+1}^{\infty} i(1-QM)^i \leqslant \infty$, so h is positive recurrent. Since there is zero probability to leave H', the Markov chain is irreducible on H' and all the states in H' are positive recurrent, η_α has a unique invariant distribution. $\square$

To prove Theorem 3 we extend Doeblin's Theorem[38] to non-homogeneous Markov chains whose entries in one column are uniformly bounded away from 0.[39] For every $R \in \mathbb{R}^{\mathbb{N}\times\mathbb{N}}$ let $\|R\|_{TV} = \sup_{i\in\mathbb{N}} \left|\sum_{j\in\mathbb{N}} r_{ij}\right|$. Given $\varepsilon > 0$, let $\mathcal{T}_\varepsilon$ be the set of (row) stochastic matrices R for which there exists $l \in \mathbb{N}$ such that $r_{jl} \geqslant \varepsilon$ for all $j \in \mathbb{N}$. Define $V_0 = \left\{x \in \mathbb{R}^{\mathbb{N}} : \sum_i x_i = 0\right\}$.

Proposition 7. *If $\{P_t\}_{t\in\mathbb{N}} \subseteq \mathcal{T}_\varepsilon$ and $P_t \to P^*$, then there are $\bar{\xi} \in \Delta(\mathbb{N})$ and $K \in \mathbb{N}$ such that for each $\xi \in \Delta(\mathbb{N})$, $\xi_n \to \bar{\xi}$, where $\xi_n = \prod_{t=1}^{n} P_t \xi$ and $\left\|\xi_n - \bar{\xi}\right\|_{TV} \leqslant 3\,(1-\varepsilon)^n$ for all $n > K$. Moreover, $\bar{\xi} P^* = \bar{\xi}$.*

Proof of Theorem 3. We use stochastic approximation for differential inclusions (Benaim, Hofbauer, and Sorin, 2005) to show that (11) can be approximated by

$$\dot{\boldsymbol{\alpha}}_t \in \mathbb{E}_{\mathcal{H}_{\boldsymbol{\alpha}_t}}[\boldsymbol{\alpha}_t^h] - \boldsymbol{\alpha}_t. \tag{15}$$

Lemma A.5. *Suppose there is $\alpha \in \Delta(A)$ such that $d_t(a,y)$ converges to $\alpha(a)p_a^*(y)$ for all $(a,y) \in A \times Y$. The distribution of time t belief μ_t conditional on a recalled history h' weakly converges to $F_{\alpha,h'} \in \Delta(\Delta(\Theta))$, and $F_{\cdot,h'}$ is continuous.*

Proof. Given a length l history $(a_i,y_i)_{i=1}^l$, let $n_{a,y}((a_i,y_i)_{i=1}^l) = \sum_{i=1}^l \mathbb{1}_{a,y}(a_i,y_i)$ be the number of times the action-outcome pair (a,y) appears in $(a_i,y_i)_{i=1}^l$. Given a history $h' = (\tilde{a}_i,\tilde{y}_i)_{i=1}^{\tilde{l}}$ recalled in period $t-1$, the recalled history at time $t \geqslant k$ is distributed

[38] Doeblin's theorem (see, e.g., Theorem 2.2.1 in Stroock [2013]) says that if all entries in a column of a (row) stochastic matrix R are at least ε, then the chain converges to its invariant distribution π at rate at least $1-\varepsilon$: $\|R^n\mu - \pi\|_{TV} \leqslant 2(1-\varepsilon)^n \|\mu\|_{TV}$.

[39] Cerreia-Vioglio, Corrao, and Lanzani [2021] proves a related extension for finite-state chains.

as a product of multinomial distributions:

$$\mathbb{P}_t\left[(a_i,y_i)_{i=1}^l|(\tilde{a}_i,\tilde{y}_i)_{i=1}^{\tilde{l}}\right]=\prod_{(a,y)\in A\times Y}\binom{d_t(a,y)t}{n_{a,y}((a_i,y_i)_{i=1}^l)}$$
$$\times\left(\frac{k}{t}(m(a,y)+R(\mathbb{1}_{(a,y)\in\{\tilde{a}_i,\tilde{y}_i\}_{i=1}^{\tilde{l}}}+\mathbb{1}_{(a,y)\notin\{\tilde{a}_i,\tilde{y}_i\}_{i=1}^{\tilde{l}}}\mathbb{1}_{(a=\pi(h'))}p_a^*(y)))\right)^{n_{a,y}((a_i,y_i)_{i=1}^l)}$$
$$\times\left(1-\frac{k}{t}(m(a,y)-R(\mathbb{1}_{(a,y)\in\{\tilde{a}_i,\tilde{y}_i\}_{i=1}^{\tilde{l}}}+\mathbb{1}_{(a,y)\notin\{\tilde{a}_i,\tilde{y}_i\}_{i=1}^{\tilde{l}}}\mathbb{1}_{(a=\pi(h'))}p_a^*(y)))\right)^{d_t(a,y)t-n_{a,y}((a_i,y_i)_{i=1}^l)}.$$

By the Poisson limit theorem, the probability that (a,y) is recalled $n_{a,y}$ times when the previous recalled history was $h'=(\tilde{a}_i,\tilde{y}_i)_{i=1}^{\tilde{l}}$ converges to $e^{-\lambda(h')_{a,y}}\frac{\lambda(h')_{a,y}^{n_{a,y}}}{n_{a,y}!}$, where

$$\begin{aligned}\lambda(h')_{a,y}&=\lim_{t\to\infty}d_t(a,y)\,t\,\left(\frac{k}{t}(m(a,y)+R(\mathbb{1}_{(a,y)\in(\tilde{a}_i,\tilde{y}_i)_{i=1}^{\tilde{l}}}+\mathbb{1}_{(a,y)\notin(\tilde{a}_i,\tilde{y}_i)_{i=1}^{\tilde{l}}}\mathbb{1}_{(a=\pi(h'))}p_a^*(y)))\right)\\&=\alpha(a)p_a^*(y)\,k\,(m(a,y)+R(\mathbb{1}_{(a,y)\in(\tilde{a}_i,\tilde{y}_i)_{i=1}^{\tilde{l}}}+\mathbb{1}_{(a,y)\notin(\tilde{a}_i,\tilde{y}_i)_{i=1}^{\tilde{l}}}\mathbb{1}_{(a=\pi(h'))}p_a^*(y)))\,.\end{aligned}\tag{16}$$

Thus the random number of times (a,y) is recalled when the previous history is h' converges to a random variable $N_{a,y}^{\alpha}(h')$ that is Poisson distributed with parameter $\lambda(h')_{a,y}$. Moreover, let $(\alpha_n)_{n\in\mathbb{N}}\in\Delta(A)$ be a sequence converging to α^*, and fix some $\varepsilon>0$. Since all the $N_{a,y}^{\alpha^*}(h')$ have Poisson distributions, there is a $K\in\mathbb{N}$ such that $\mathbb{P}[\max_{(a,y)\in A\times Y}N_{a,y}^{\alpha^*}(h')>K]<\varepsilon$. Let $M\in\mathbb{N}$ be such that $\mathbb{P}[\max_{(a,y)\in A\times Y}N_{a,y}^{\alpha_n}(h')>K]<\varepsilon$ and $|\mathbb{P}[N_{a,y}^{\alpha_n}(h')=c]-\mathbb{P}[N_{a,y}^{\alpha^*}(h')=c]|<\varepsilon$ for all $(a,y)\in A\times Y$, for all $c\leqslant K$ and $n>M$. Then for any continuous and bounded $f:\Delta(\Theta)\to\mathbb{R}$, for all $n>M$ we have

$$\left|\int_{\Delta(\Theta)}f(\nu)dF_{\alpha_n,h'}-\int_{\Delta(\Theta)}f(\nu)dF_{\alpha,h'}\right|<\max_{\nu\in\Delta(\Theta)}|f(\nu)|((K+1)|A\times Y|)\varepsilon,\tag{17}$$

so $F_{\alpha_n,h'}$ weakly converges to $F_{\alpha,h'}$. □

We now apply Proposition 7 to prove that the distribution of recalled histories converges to $\mathcal{H}_{\boldsymbol{\alpha}}$. Lemma A.5 shows that the transition matrices over histories converge. Moreover, from the definition of m_t, the transition from histories that contain the same experiences (but possibly a different number of each) is the same, and since the set of experiences is finite, this convergence is uniform, as required by Proposition 7. Transition to the null history always has positive probability, so the transition matrices are in $\mathcal{T}_\varepsilon$ and thus the proposition guarantees convergence to the stationary distribution

$\mathcal{H}_{\boldsymbol{\alpha}}$. With this the rest of the proof follows the same steps of Theorem 2. $\square$

A.9 Partial naïveté

So far we have assumed that agents treat the experiences they remember as if these were the only ones that happened. This section considers agents who are partially aware of their memory limitations. We suppose throughout this section that actions have no effect on the outcome distribution. We also assume that the agent either does not remember their own actions or believes their actions convey no information, so they do not draw inferences about the experiences they forgot from their actions.[40] Finally, we suppose that agents know the current period and so know how many observations they have forgotten.

Suppose the agent believes that they remember an occurrence of signal s and outcome y with probability $\hat{m}(s,y)$, instead of the true probability $m(s,y)$. The subjective log-likelihood of a remembered history $h_t(r_t)$ after signal s' under data generating process p is then

$$\binom{t}{|r_t|}\left[\sum_{s\in S}\zeta(s)\sum_{z\in Y}p_s(z)(1-\hat{m}_{s'}(s,z))\right]^{t-|r_t|}\prod_{i\in r_t}\zeta(s_i)p_{s_i}(y_i)\hat{m}_{s'}(s_i,y_i)\,.$$

Thus up to a constant the subjective log-likelihood of outcome frequency f equals

$$\begin{aligned}(t-|r_t|)\log&\left[\sum_{s\in S}\zeta(s)\sum_{z\in Y}p_s(z)(1-\hat{m}_{s'}(s,z))\right]\\ &+\sum_{y\in Y}\sum_{s\in S}\sum_{\tau=1}^{|r_t|}\mathbb{1}_{s,y}(s_\tau,y_\tau)\log(p_s(y)\hat{m}_{s'}(s,y))\,,\end{aligned}\tag{18}$$

where $|r_t|$ is the number of events the agent remembers. (Note that the first term does not appear when the agent believes they remember everything, which corresponds to $|r_t|=t$.)

Because the expected value of $|r_t|/t$ is $1-\sum_{y\in Y}\sum_{s\in S}p(y)m_{s'}(s,y)$, (18) suggests

[40] By assuming that the agent makes no inference from their own past actions we rule out that the agent is sophisticated enough to make inference about the signals they saw from the actions they took.

the following generalization of

$$KL_{s'}^{\Theta,m,\hat{m}}(\sigma) = \operatorname*{argmin}_{p\in\Theta} -\Big(1-\sum_{s\in S}\sum_{y\in Y} m_{s'}(s,y)\zeta(s)p_s^*(y)\Big)\log\Big(1-\sum_{s\in S}\sum_{y\in Y}\zeta(s)p_s(y)\hat{m}_{s'}(s,y)\Big)$$
$$-\sum_{y\in Y}\sum_{s\in S} m_{s'}(s,y)\zeta(s)p_s^*(s,y)\log\left(\hat{m}_{s'}(s,y)p_s(y)\right).$$

Notice that when $\hat{m}_{s'}(s,y)=1$ for all s and y, the second term of the equation is equal to the minimized function in (2), and the first term is independent of p and has no impact on the minimization.

Definition 9. A *selective memory equilibrium for a partially naïve agent* is a strategy σ such that for every $s\in S$ there exists a belief $\nu\in KL_s^{\Theta,m,\hat{m}}(\sigma)$ with $\sigma\in BR(s,\nu)$.

For an agent who is aware of their own forgetfulness, but not aware that their memory is selective, i.e., who believes that their memory function m is constant, $KL_s^{\Theta,m,\hat{m}} = KL_s^{\Theta,m}$ and the selective memory equilibria of a partially and fully naïve agent coincide.[41] This shows that for our results what is important is that the agent is unaware of the selective nature of their memory, not that they are unaware of their forgetfulness. At the other extreme, if agents are fully aware of their memory function and correctly specified then $\delta_{p*}\in KL_{s'}^{\Theta,m,m}$ and thus any action that is optimal for the true data generating process is always a selective memory equilibrium.[42]

The next result, whose proof is omitted, follows from an argument analogous to the proof of Theorem 1.

Proposition 8. *When the agent is partially naïve, every limit strategy is a selective memory equilibrium.*

Moreover, as with notions of partial naïveté in cursed equilibrium and quasi-hyperbolic discounting, one can define a parametric notion of partial naïveté by assuming that $\hat{m}_{s'}(s,y)=(1-\alpha)+\alpha m_{s'}(s,y)$. For $\alpha=0$ the agent is fully naïve and unaware of their

[41] To see why, note that for $\hat{m}_{s'}$ constant, the minimized function is $-\sum_{y\in Y}\sum_{s\in S} m_{s'}(s,y)\zeta(s)p_s^*(y)\log\left(p_s(y)\right) - \left[\sum_{s\in S}\sum_{y\in Y}\zeta(s)p_s^*(y)m_{s'}(s,y)\log(\hat{m}_{s'}(s,y))\right] - \left[\sum_{s\in S}\zeta(s)\sum_{y\in Y}p_s^*(y)(1-m_{s'}(s,y))\right]\log(1-\hat{m}_{s'}(s,y))$. The last two terms are independent of p so $KL_{s'}^{\Theta,m,\hat{m}} = KL_{s'}^{\Theta,m}$.

[42] More generally, if the agents are misspecified but fully aware then selective memory equilibrium reduces to Berk-Nash equilibrium.

memory limitations. For $\alpha = 1$ the agent is sophisticated and understands their memory limitations which in consequence implies that the agent will have correct long-run beliefs.

As the next proposition illustrates, the degree of naïveté can amplify existing memory biases. Consider again the setting of Section 4.2 which studied positive memory bias by assuming that y is a scalar and $m(y)$ is increasing in y.

Proposition 9. *Suppose m and p^* are constant in a and m is increasing in y, that $\hat{m}(y) = (1-\alpha)+\alpha m(y)$, $\Theta = \Delta(\Delta(Y))$, and the utility function is supermodular. Then the agent's long-run belief concentrates on a distribution of outcomes that is increasing in first-order stochastic dominance in $(1-\alpha)$, i.e., the naïveté of the agent.*

Proof of Proposition 9. We first derive the long-run belief for a given subjective memory function $\hat{m}$. For every σ,

$$KL^{\Theta,m,\hat{m}}(\sigma) = \underset{p\in\Delta(Y)}{\operatorname{argmin}} - \sum_{y\in Y} \log(p(y)\hat{m}(y))m(y)p^*(y)\,.$$

Taking first-order conditions of the associated Lagrangian shows there is a unique element p of $KL^{\Theta,m,\hat{m}}(\sigma)$, given by

$$p(y) = \frac{\frac{m(y)}{\hat{m}(y)}p^*(y)}{\sum_{z\in Y}\frac{m(z)}{\hat{m}(z)}p^*(z)}\,.$$

Thus the long-run beliefs under the objective memory function m and subjective memory function $\hat{m}$ are the same as those of a fully naïve agent with memory function $\tilde{m}(y) = \frac{m(y)}{\hat{m}(y)}$ who is not aware of their selective memory. Note that for $\hat{m}_\alpha(y) = (1-\alpha)+\alpha m(y)$ and $\alpha > \alpha'$,

$$\frac{\frac{m(y)}{\hat{m}_\alpha(y)}}{\frac{m(y)}{\hat{m}_{\alpha'}(y)}} = \frac{\hat{m}_{\alpha'}(y)}{\hat{m}_\alpha(y)} = \frac{(1-\alpha')+\alpha' m(y)}{(1-\alpha)+\alpha m(y)}$$

is decreasing in $m(y)$ and hence in y. This lets us apply Proposition 2 to conclude that the long-run belief under the subjective memory function $\hat{m}_{\alpha'}$ will be higher in FOSD than that under the subjective memory function $\hat{m}_\alpha$. $\square$

References

Adler, O. and A. Pansky (2020). "A "rosy view" of the past: Positive memory biases." *Cognitive Biases in Health and Psychiatric Disorders*. Elsevier, 139–171.

Afrouzi, H., S. Y. Kwon, A. Landier, Y. Ma, and D. Thesmar (2020). *Overreaction and working memory*. Tech. rep. National Bureau of Economic Research.

Aliprantis, C. and K. Border (2013). *Infinite dimensional analysis: A hitchhiker's guide*. Berlin. Springer-Verlag.

Alloy, L. B. and L. Y. Abramson (1979). "Judgment of contingency in depressed and nondepressed students: Sadder but wiser?" *Journal of experimental psychology: General*, 108, 441.

Battigalli, P. and N. Generoso (2021). *Information Flows and Memory in Games*. Tech. rep.

Bénabou, R. and J. Tirole (2002). "Self-confidence and personal motivation." *The Quarterly Journal of Economics*, 117, 871–915.

Benaim, M., J. Hofbauer, and S. Sorin (2005). "Stochastic approximations and differential inclusions." *SIAM Journal on Control and Optimization*, 44, 328–348.

Benjamin, D. J. (2019). "Errors in probabilistic reasoning and judgment biases." *Handbook of Behavioral Economics: Applications and Foundations*, 2, 69–186.

Benjamin, D. J., S. A. Brown, and J. M. Shapiro (2013). "Who is 'behavioral'? Cognitive ability and anomalous preferences." *Journal of the European Economic Association*, 11, 1231–1255.

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

Bordalo, P., G. Burro, K. B. Coffman, N. Gennaioli, and A. Shleifer (2022). *Imagining the future: memory, simulation and beliefs about COVID*. Tech. rep. National Bureau of Economic Research.

Bordalo, P., K. Coffman, N. Gennaioli, F. Schwerter, and A. Shleifer (2021). "Memory and representativeness." *Psychological Review*, 128, 71.

Bordalo, P., J. Conlon, N. Gennaioli, S. Kwon, and A. Shleifer (2021). "Memory and probability." *NBER WP 29273*.

Bordalo, P., N. Gennaioli, and A. Shleifer (2019). "Memory and reference prices: An application to rental choice." *AEA Papers and Proceedings*. Vol. 109, 572–76.

— (2020). "Memory, attention, and choice." *The Quarterly Journal of Economics*, 135, 1399–1442.

Boyle, P., L. Yu, A. Buchman, and D. Bennett (2012). "Risk aversion is associated with decision making among community-based older persons." *Frontiers in Psychology*, 3, 205.

Cerreia-Vioglio, S., R. Corrao, and G. Lanzani (2021). "Dynamic opinion aggregation: Long-run stability and disagreement." *Mimeo*.

Chammat, M. et al. (2017). "Cognitive dissonance resolution depends on episodic memory." *Scientific Reports*, 7, 1–10.

Cheung, Y.-W. and D. Friedman (1997). "Individual learning in normal form games: Some laboratory results." *Games and economic behavior*, 19, 46–76.

Chew, S. H., W. Huang, and X. Zhao (2020). "Motivated false memory." *Journal of Political Economy*, 128, 3913–3939.

Cokely, E. T. and C. M. Kelley (2009). "Cognitive abilities and superior decision making under risk: A protocol analysis and process model evaluation." *Judgment and Decision Making*, 4, 20–33.

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

Danz, D., L. Vesterlund, and A. J. Wilson (2022). "Belief elicitation and behavioral incentive compatibility." *American Economic Review*.

Dudley, R. M. (2018). *Real analysis and probability*. Chapman and Hall/CRC.

Elkin, R. A. and M. R. Leippe (1986). "Physiological arousal, dissonance, and attitude change: Evidence for a dissonance-arousal link and a "Don't remind me" effect." *Journal of Personality and Social Psychology*, 51, 55.

Enke, B., F. Schwerter, and F. Zimmermann (2020). "Associative memory and belief formation." *NBER WP 26664*.

Erev, I. and E. Haruvy (2016). "Learning and the economics of small decisions." *Handbook of Experimental Economics, Volume 2*. Princeton University Press, 638–716.

Esponda, I. (2008). "Information feedback in first price auctions." *The RAND Journal of Economics*, 39, 491–508.

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

Esponda, I., D. Pouzo, and Y. Yamamoto (2021a). "Asymptotic behavior of Bayesian learners with misspecified models." *Journal of Economic Theory*, 195, 105260.

— (2021b). "Corrigendum: Asymptotic behavior of Bayesian learners with misspecified models." *Journal of Economic Theory*, 195, 105260.

Fudenberg, D., G. Lanzani, and P. Strack (2021a). "Limit points of endogenous misspecified learning." *Econometrica*, 89, 1065–1098.

— (2021b). "Pathwise concentration bounds for misspecified Bayesian beliefs." *Mimeo*.

Fudenberg, D. and D. K. Levine (1993a). "Self-confirming equilibrium." *Econometrica*, 61, 523–545.

— (1993b). "Steady state learning and Nash equilibrium." *Econometrica*, 547–573.

Fudenberg, D. and D. K. Levine (2014). "Learning with recency bias." *Proceedings of the National Academy of Sciences*, 111, 10826–10829.

Fudenberg, D. and A. Peysakhovich (2016). "Recency, records, and recaps: Learning and nonequilibrium behavior in a simple decision problem." *ACM Transactions on Economics and Computation (TEAC)*, 4, 1–18.

Gilovich, T. (2008). *How we know what isn't so.* Simon and Schuster.

Gödker, K., P. Jiao, and P. Smeets (2022). "Investor memory." *SSRN 3348315*.

Goetzmann, W. N., A. Watanabe, and M. Watanabe (2022). *Evidence on Retrieved Context: How History Matters.* Tech. rep. National Bureau of Economic Research.

Hartzmark, S. M., S. Hirshman, and A. Imas (2019). "Ownership, learning, and beliefs." *Available at SSRN*, 3465246.

Hastie, R. and B. Park (1986). "The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line." *Psychological Review*, 93, 258.

He, K. (2022). "Mislearning from censored data: The gambler's fallacy and other correlational mistakes in optimal-stopping problems." *Theoretical Economics*, 17, 1269–1312.

Heidhues, P., B. Kőszegi, and P. Strack (2018). "Unrealistic expectations and misguided learning." *Econometrica*, 86, 1159–1214.

James, W. (1890). *The principles of psychology.* Vol. 1. Cosimo, Inc.

Jehiel, P. (2018). "Investment strategy and selection bias: An equilibrium perspective on overoptimism." *American Economic Review*, 108, 1582–97.

— (2021). "Communication with forgetful liars." *Theoretical Economics*, 16, 605–638.

Jehiel, P. and J. Steiner (2020). "Selective Sampling with Information-Storage Constraints." *The Economic Journal*, 130, 1753–1781.

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

Kahneman, D. and A. Tversky (1972). "Subjective probability: A judgment of representativeness." *Cognitive Psychology*, 3, 430–454.

— (1979). "Prospect theory: An analysis of decision under risk." *Econometrica*, 47, 263–292.

Kandel, E. R. et al. (2000). *Principles of neural science.* Vol. 4. McGraw-hill.

Koszegi, B., G. Loewenstein, and T. Murooka (2021). "Fragile self-esteem." *Review of Economics Studies.*

Lord, C. G., L. Ross, and M. R. Lepper (1979). "Biased assimilation and attitude polarization: The effects of prior theories on subsequently considered evidence." *Journal of personality and social psychology*, 37, 2098.

Malmendier, U. and S. Nagel (2016). “Learning from inflation experiences.” *The Quarterly Journal of Economics*, 131, 53–87.

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

Malmendier, U. and L. S. Shen (2018). “Scarred consumption.” *NBER WP 24696*.

Matt, G. E., C. Vázquez, and W. K. Campbell (1992). “Mood-congruent recall of affectively toned stimuli: A meta-analytic review.” *Clinical psychology review*, 12, 227–255.

Mayer, J. D., L. J. McCormick, and S. E. Strong (1995). “Mood-congruent memory and natural mood: New evidence.” *Personality and Social Psychology Bulletin*, 21, 736–746.

Miller, G. A. (1956). “The magical number seven, plus or minus two: Some limits on our capacity for processing information.” *Psychological Review*, 63, 81.

Mischel, W., E. B. Ebbesen, and A. M. Zeiss (1976). “Determinants of selective memory about the self.” *Journal of consulting and clinical Psychology*, 44, 92.

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

Nagel, S. and Z. Xu (2022). “Asset pricing with fading memory.” *The Review of Financial Studies*, 35, 2190–2245.

Phillips, L. D. and W. Edwards (1966). “Conservatism in a simple probability inference task.” *Journal of Experimental Psychology*, 72, 346.

Procházka, J., M. Ovcari, and M. Durinik (2020). “Sandwich feedback: The empirical evidence of its effectiveness.” *Learning and Motivation*, 71, 101649.

Rabin, M. and J. L. Schrag (1999). “First impressions matter: A model of confirmatory bias.” *The Quarterly Journal of Economics*, 114, 37–82.

Rabin, M. and D. Vayanos (2010). “The gambler’s and hot-hand fallacies: Theory and applications.” *The Review of Economic Studies*, 77, 730–778.

Reder, L. M. (2014). *Implicit memory and metacognition*. Psychology Press.

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

Schwartzstein, J. (2014). “Selective attention and learning.” *Journal of the European Economic Association*, 12, 1423–1452.

Seneta, E. (2006). *Non-negative matrices and Markov chains*. Springer Science & Business Media.

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

Stroock, D. W. (2013). *An introduction to Markov processes*. Springer Science & Business Media.

Thomson, D. M. and E. Tulving (1970). "Associative encoding and retrieval: Weak and strong cues." *Journal of Experimental Psychology*, 86, 255.

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

Walters, D. J. and P. M. Fernbach (2021). "Investor memory of past performance is positively biased and predicts overconfidence." *Proceedings of the National Academy of Sciences*, 118.

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

Wilson, A. (2014). "Bounded Memory and Biases in Information Processing." *Econometrica*, 82, 2257–2294.

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

B For Online Publication

B.1 Proof of Lemmas

Lemma A.1. *For all Borel measurable $C, C' \subseteq \Delta(Y)^{A\times S}$, $t \in \mathbb{N}$, $h_t \in H_t$ and $r_t \subseteq \{1, ..., t\}$,*

$$\frac{\mu(C|h_t(r_t))}{1-\mu(C'|h_t(r_t))} \geqslant \frac{\mu(C)}{1-\mu(C')} \exp\left(|r_t| \left[\inf_{p\in\Theta\setminus C'} D(\hat{f}(h_t, r_t)||p) - \sup_{p\in C} D(\hat{f}(h_t, r_t)||p)\right]\right).$$

Proof. Using (9) we have that

$$\begin{aligned}\frac{\mu(C|h_t(r_t))}{1-\mu(C'|h_t(r_t))} &= \frac{\int_{p\in C} \exp\left(|r_t| \sum_{(s,a,y)} \log(p_{a,s}(y))\hat{f}(h_t, r_t)(s,a,y)\right) \mathrm{d}\mu}{\int_{p\in\Theta\setminus C'} \exp\left(|r_t| \sum_{(s,a,y)} \log(p_{a,s}(y))\hat{f}(h_t, r_t)(s,a,y)\right) \mathrm{d}\mu} \\ &= \frac{\int_{p\in C} \exp\left(-|r_t| D(\hat{f}(h_t, r_t)||p)\right) \mathrm{d}\mu}{\int_{p\in\Theta\setminus C'} \exp\left(-|r_t| D(\hat{f}(h_t, r_t)||p)\right) \mathrm{d}\mu} \\ &\geqslant \frac{\mu(C)}{1-\mu(C')} \frac{\exp\left(-|r_t| \sup_{p\in C} D(\hat{f}(h_t, r_t)||p)\right)}{\exp\left(-|r_t| \inf_{p\in\Theta\setminus C'} D(\hat{f}(h_t, r_t)||p)\right)}.\end{aligned}$$

□

Lemma A.2. *If σ is not a selective memory equilibrium, there are $s' \in S$ and $\varepsilon, C \in \mathbb{R}_{++}$ such that*

$$\frac{\nu(KL_{s'}^{\Theta,m}(\sigma,\varepsilon))}{1-\nu(KL_{s'}^{\Theta,m}(\sigma,\varepsilon))} > C \implies \sigma(s') \notin BR(s', \nu).$$

If σ is a uniformly strict selective memory equilibrium, there are $\varepsilon, C \in \mathbb{R}_{++}$ such that for all $s \in S$

$$\frac{\nu(KL_{s}^{\Theta,m}(\sigma,\varepsilon))}{(1-\nu(KL_{s}^{\Theta,m}(\sigma,\varepsilon))} > C \implies \{\sigma(s)\} = BR(s, \nu).$$

Proof of Lemma A.2. First observe that every $\sigma \in A^S$, $s \in S$, and $\varepsilon > 0$, $KL_s^{\Theta,m}(\sigma)$ and $KL_s^{\Theta,m}(\sigma,\varepsilon)$ are nonempty and compact. Indeed, by Assumption 1, the function

$$p \mapsto \sum_{s'\in S} \zeta(s') \sum_{y\in Y} m_s(s', \sigma(s), y) p^*_{\sigma(s'),s'}(y) \log p_{\sigma(s'),s'}(y)$$

is finite-valued and continuous on the compact set Θ. Therefore, $KL_s^{\Theta,m}(\sigma)$ is nonempty and compact by Theorem 2.43 in Aliprantis and Border [2013]. The result for $KL_s^{\Theta,m}(\sigma,\varepsilon)$ is an immediate consequence given the continuity of the supnorm.

For the first part of the lemma, suppose σ is not a selective memory equilibrium. Then there is an $s' \in S$ such that $\sigma(s') \notin BR(s', \Delta(KL_{s'}^{\Theta,m}(\sigma)))$. The upperhemicontinuity of the best reply map $BR(s,\cdot)$ and the compactness of $KL_{s'}^{\Theta,m}(\sigma,\varepsilon)$ imply that there are $\varepsilon, C \in \mathbb{R}_{++}$ such that if $\frac{\nu(KL_{s'}^{\Theta,m}(\sigma,\varepsilon))}{1-\nu(KL_{s'}^{\Theta,m}(\sigma,\varepsilon))} > C$ then $\sigma(s') \notin BR(s',\nu)$.

For the second part of the lemma, suppose σ is a uniformly strict selective memory equilibrium. The upperhemicontinuity of the best reply map $BR(s,\cdot)$ and the compactness of $KL_{s'}^{\Theta,m}(\sigma,\varepsilon)$ imply that there are $C \in \mathbb{R}_{++}$ and ε such that for all $s \in S$ if $\nu(KL_s^{\Theta,m}(\sigma,\varepsilon)) > C(1-\nu(KL_s^{\Theta,m}(\sigma,\varepsilon))$ then $\{\sigma(s)\} = BR(s,\nu)$. □

Lemma A.3. *If, for every t, every sequence of actions a^t, and every optimal policy $\tilde{\pi}$, $\mathbb{P}_{\pi^{\sigma,a^t}}[\sigma(s_\tau) = \tilde{\pi}(h_\tau(r_\tau), s_\tau)$ for all $\tau \geqslant t] = 0$ then σ is not a limit strategy.*

Proof of Lemma A.3. Fix an arbitrary optimal policy $\tilde{\pi}$ and a history (s^t, a^t, y^t) with $\mathbb{P}_{\tilde{\pi}}(s^t, a^t, y^t) > 0$. Let $\tau = \min\{t' > t \colon \sigma(s_{t'}) \neq \tilde{\pi}((s^{t'}, a^{t'}, y^{t'})(r_{t'}))(s_{t'+1})\}$ be the first time after (s^t, a^t, y^t) when $\tilde{\pi}$ does not prescribe σ. Note that since $\tilde{\pi}((s^{\hat{t}}, a^{\hat{t}}, y^{\hat{t}})(s_{\hat{t}+1}) = \sigma(s_{\hat{t}+1}) = \pi^{\sigma,a^t}(s^{\hat{t}+1}, a^{\hat{t}}, y^{\hat{t}})$ for all $\hat{t} \in [t, \tau-1]$, the agent's belief until period τ is the same under π^{σ,a^t} and $\tilde{\pi}$. As $\mathbb{P}_{\tilde{\pi}}(s^{t'+1}, a^{t'}, y^{t'}, r_{t'}) > 0$ implies $\mathbb{P}_{\pi^{\sigma,a^t}}(s^{t'+1}, a^{t'}, y^{t'}, r_{t'}) > 0$, the probability that the agent uses strategy σ forever (i.e., $\tau = \infty$) after history (s^t, a^t, y^t) equals 0 by the assumption of the lemma. So under every arbitrary optimal policy, after every history where σ is played a strategy different from σ is played with probability 1, so σ is not a limit strategy. □

B.2 Partially Recalled Histories

Here we suppose that the outcome space has a product structure, i.e., $Y = \times_{i\in I} Y_i$, and that the agent may recall only some components of the outcome. To model this case, we use a collection of signal-dependent *memory functions* $m_{s'} : (S \times A \times Y \times 2^I) \to [0,1]$, where $m_{s'}(s,a,y,B)$ specifies the probability with which an agent remembers the B outcome components of a past realization of the signal, action, outcome triplet (s,a,y) and with

$$\sum_{B\in 2^I} m_{s'}(s,a,y,B) = 1.$$

The case studied in the main body of the paper is $m_{s'}(s,a,y,B) = 0$ for every $B \notin \{\varnothing, I\}$ with the interpretation that $m_{s'}(s,a,y,\varnothing)$ means that the agent does not recall the experience (s, a, y) at all.

More generally, the *recalled history* at time t is the of recalled experiences $r_t = (s_\tau, a_\tau, y_\tau, B_{\tau,t})_{\tau=1}^t$ where $B_{\tau,t}$ denotes the components of the period τ outcome recalled at time t. If we maintain that the agent is unaware of their selective memory and naïvely updates their beliefs as if the experiences they remember are the only ones that occurred, we obtain

$$\mu(C|r_t) = \frac{\int_{p \in C} \prod_{\tau=1}^t p_{a_\tau, s_\tau}(\prod_{i \in I} \tilde{Y}_{\tau,i}) d\mu(p)}{\int_\Theta \prod_{\tau=1}^t p_{a_\tau, s_\tau}(\prod_{i \in I} \tilde{Y}_{\tau,i}) d\mu(p)} \qquad \forall C \subseteq \Theta$$

where $\tilde{Y}_{\tau,i} = Y_i$ if $i \notin B_{\tau,t}$ and $\tilde{Y}_{\tau,i} = \{y_{\tau,i}\}$ if $i \in B_{\tau,t}$.

With this, the results of the paper carry through with the following adaptation of the concept of memory-weighted KL-minimizers:

$$KL_{s'}^{\Theta,m}(\sigma) = \operatorname*{argmin}_{p \in \Theta} \sum_{s \in S} \zeta(s) \sum_{B \in 2^I} \sum_{y \in Y} m_{s'}(s, \sigma(s), y, B) p^*_{\sigma(s),s}(\tilde{Y}(y)) \log p_{\sigma(s),s}(\tilde{Y}(y))$$

where $\tilde{Y}(y) = Y_i$ if $i \notin B$ and $\tilde{Y}(y) = \{y_i\}$ if $i \in B$.

B.3 Proposition 3

Proposition 3. *Suppose that m and p^* are constant in a and y_2 and m is increasing in y_1. Then the agent's long-run belief about p_1 concentrates on a distribution that is higher in first-order stochastic dominance than p_1^* and the agent's long-run belief about p_2 concentrates on a distribution that is lower than p_2^*.*

Proof of Proposition 3. Because (y_1, y_2) are subjectively independent conditional on the value of p_1, the learning problem decouples across the two dimensions. By Proposition 2 the long-run belief about p_1 is higher than the true distribution p_1^*. The probability with which an outcome is remembered is independent of the second component, so the agent learns p_2^*. They infer q to be

$$q(y_2) = \frac{p_2^*(y_2) - \alpha p_1(y_2)}{1 - \alpha}.$$

Thus $q - q^* \equiv \frac{\alpha}{1-\alpha}(p_1^* - p_1)$, and as p is greater than p^* in first-order stochastic dominance, it follows that q is lower than q^* in first-order stochastic dominance. $\square$

B.4 Proof of the Remark

Remark 1. *The uniform strictness conditions of Proposition 6 are needed:*

1. *There are Berk-Nash equilibria that are not belief equivalent to any selective memory equilibrium with support $\Theta' = \Delta(Y)^{A\times S}$.*

2. *There are selective memory equilibria that are not belief equivalent to any Berk-Nash equilibrium.*

3. *Moreover, unlike Berk-Nash equilibria, selective memory equilibria do not reduce to self-confirming equilibria when the agent is correctly specified. In particular, selective memory equilibria need not be objectively optimal when the distribution of outcomes is independent of the agent's action.*

Proof of the Remark 1. To prove the statements we give three examples with a null signal space S.

1. Suppose that $Y = \{-1, 1\} = A$, the probability of 1 is 0.5 regardless of a, and that the agent does not have selective memory, but is misspecified, with $[0, .2] \cup [.8, 1]$ as the support of the prior beliefs over the probability of 1 under both actions. Then both .2 and .8 are KL minimizers, which cannot arise from selective memory with a full-support prior. This is immediate if $m = 0$ for some experience, and follows from full support and the strict convexity of the memory-weighted KL divergence if $m \gg 0$.

2. Suppose that $Y = \{-1, 0, 1\} = A$, $u(a, y) = ay + 0.05\mathbb{1}_{a=-1} - \mathbb{1}_{a=1}/12$, the probability over outcomes is uniform regardless of a, with $\Theta = \{(1/3, 1/3, 1/3), (1/3, 1/6, 1/2)\}$ and $m(a, y) = \mathbb{1}_{y=-1}$. Then both $(1/3, 1/3, 1/3)$ and $(1/3, 1/6, 1/2)$ are memory-weighted KL minimizers, but they can never be both KL minimizers with perfect memory. Therefore, 0 is a selective memory equilibrium that is not belief equivalent to any Berk-Nash equilibrium.

3. Suppose $Y = \{-1, 1\} = A$ and $u(a, y) = ya$. Then if $m(a, -1) = 0 < m(a, 1)$ for all $a \in A$, and the agent has a full-support prior over the action-independent outcome distributions, the only selective memory equilibrium is $a = 1$ even if the true probability of 1 under both actions is less than $1/2$ so that the objectively optimal action is -1. □

B.5 Proposition 4

Proposition 4 *Suppose $p_1^*(\bar{y}+y) = p_1^*(\bar{y}-y)$ and that the agent thinks all outcome distributions are possible under the risky action. If choosing the lottery is not a selective memory equilibrium with perfect memory, it is not a selective memory equilibrium with extreme experience bias.*

Proof of Proposition 4. If $a = 1$ is not a selective memory equilibrium with perfect memory, then the safe action $a = 0$ is preferred to the risky action $a = 1$, so $\sum_{y \in Y} v(y)p_1^*(y) < v(y_0)$. Because the prior assigns positive probability to all distributions induced by action a_1, the unique memory-weighted KL minimizer $\hat{p}$ under action a is such that

$$\hat{p}_1(y) := \frac{p_1^*(y)h(|y-\bar{y}|)}{\sum_{z \in Y} p_1^*(z)h(|z-\bar{y}|)}.$$

Therefore, if $a = 1$ is a selective memory equilibrium with selective memory $m(y) = h(|y - \bar{y}|)$, then

$$v(y_0) \leqslant \sum_{y \in Y} \hat{p}_1(y)v(y)\,.$$

We prove that this cannot be the case by showing that the distribution $\hat{p}_1$ is second-order stochastically dominated by p_1^*. To see this, observe that as p_1^* is symmetric around $\bar{y}$ and $h(|y - \bar{y}|)$ is symmetric around $\bar{y}$ it follows that $\hat{p}_1$ is symmetric around $\bar{y}$. As h is increasing it follows that $\hat{p}_1 - p_1^*$ changes its sign from positive to negative and back to positive so $\sum_{y \leqslant z} p_1^*(y)$ and $\sum_{y \leqslant z} \hat{p}_1(y)$ cross only once, at $z = \bar{y}$. This implies that $\hat{p}_1$ is a mean-preserving spread of p_1^*, so the risky action is perceived as more risky. As v is concave, this implies that

$$\sum_{y \in Y} v(y)p^*(y) \geqslant \frac{\sum_{y \in Y} p_1^*(y)h(|y-r|)v(y)}{\sum_{y \in Y} p_1^*(y)h(|y-r|)}$$

and the risky action cannot be a selective memory equilibrium. □

B.6 Proposition 5

Proposition 5. *If σ is a limit strategy with underinference distortion m, it is a selective memory equilibrium with memory function m.*

Proof of Proposition 5. Suppose towards a contradiction that σ is a limit strategy under the optimal policy π, but not a selective memory equilibrium. Then by Lemma

A.2 there are $s' \in S$ and $c, C \in \mathbb{R}_{++}$ such that if $\frac{\nu(KL^{\Theta,m}_{s'}(\sigma,c))}{1-\nu(KL^{\Theta,m}_{s'}(\sigma,c))} > C$ then $\sigma(s') \notin BR(s',\nu)$. Let h_t be a history with positive probability. We show that if the agent plays the strategy $\tilde{\pi}$ that coincides with π until h_t and prescribes σ at every period after h_t, then almost surely μ_t reaches a region where no optimal policy prescribes σ after signal s'. By Lemma A.3 this is enough to obtain the desired conclusion.

Under strategy $\tilde{\pi}$, by the Strong Law of Large Numbers,

$$\lim_{\tau\to\infty} f(h_\tau)(s,a,y) = \begin{cases} \zeta(s)p^*_{a,s}(y) & \text{if } a = \sigma(s) \\ 0 & \text{otherwise} \end{cases} \tag{19}$$

almost surely on the cylinder h_t. Now we express the posterior as a function of the observed frequencies, and show that it concentrates on the memory-weighted KL minimizers, so that the result follows from the upperhemicontinuity of BR.

Choose $\kappa \in \mathbb{R}_{++}$ such that

$$\kappa < \inf_{\{p' \notin KL^{\Theta,m}_{s'}(\sigma,c)\}} \left(-\sum_{s\in S} \zeta(s) \sum_{y\in Y} p^*_{\sigma(s),s}(y) m_{s'}(s,\sigma(s),y) \log p'_{\sigma(s),s}(y) \right) - M_\sigma$$

and $c' < c$ such that

$$\kappa/2 > \sup_{\{p' \in KL^{\Theta,m}_{s'}(\sigma,c')\}} \left(-\sum_{s\in S} \zeta(s) \sum_{y\in Y} p^*_{\sigma(s),s}(y) m_{s'}(s,\sigma(s),y) \log p'_{\sigma(s),s}(y) \right) - M_\sigma.$$

By equation (19) and the definition of κ and c', almost surely on the cylinder h_t we have

$$\begin{aligned} K: &= \lim_{t\to\infty} \inf_{\{p' \notin KL^{\Theta,m}_{s'}(\sigma,c)\}} -\sum_{(s,a,y)} f(h_t(s,a,y)) m(s,a,y) \log(p'_{a,s}(y)) \\ &- \lim_{t\to\infty} \sup_{\{p' \in KL^{\Theta,m}_{s'}(\sigma,c')\}} -\sum_{(s,a,y)} f(h_t(s,a,y)) m(s,a,y) \log(p'_{a,s}(y)) \\ &= -\inf_{\{p' \notin KL^{\Theta,m}_{s'}(\sigma,c)\}} \sum_{s\in S} \zeta(s) \sum_{y\in Y} m_{s'}(s,\sigma(s),y) p^*_{\sigma(s),s}(y) \log p'_{\sigma(s),s}(y) \\ &- \sup_{\{p' \in KL^{\Theta,m}_{s'}(\sigma,c')\}} -\sum_{s\in S} \zeta(s) \sum_{y\in Y} m_{s'}(s,\sigma(s),y) p^*_{\sigma(s),s}(y) \log p'_{\sigma(s),s}(y) > \kappa/2 > 0. \end{aligned}$$

By Lemma A.1 we have

$$
\begin{aligned}
&\frac{\mu(KL^{\Theta,m}_{s'}(\sigma,c)|(h_t))}{1-\mu(KL^{\Theta,m}_{s'}(\sigma,c)|(h_t))}\\
&\geqslant \frac{\mu\left(KL^{\Theta,m}_{s'}(\sigma,c')\right)\exp\left(\sup_{\{p'\in KL^{\Theta,m}_{s'}(\sigma,c')\}} -\sum_{(s,a,y)} tf(h_t(s,a,y))m(s,a,y)\log(p'_{a,s}(y))\right)}{\left(1-\mu\left(KL^{\Theta,m}_{s'}(\sigma,c')\right)\right)\exp\left(\inf_{\{p'\notin KL^{\Theta,m}_{s'}(\sigma,c)\}} -\sum_{(s,a,y)} tf(h_t(s,a,y))m(s,a,y)\log(p'_{a,s}(y))\right)}\\
&= \frac{\mu\left(KL^{\Theta,m}_{s'}(\sigma,c')\right)}{\left(1-\mu\left(KL^{\Theta,m}_{s'}(\sigma,c')\right)\right)}\exp(tK),
\end{aligned}
$$

which goes to ∞ as t grows since $K > 0$. □

B.7 Proposition 7

Since the elements of V_0 can be interpreted as difference between stochastic vectors, the next two lemmas show that the matrices in $\mathcal{T}_\varepsilon$ have a contraction property.

Lemma B.1. *If $R \in \mathcal{T}_\varepsilon$ and $x \in V_0$, then $xR \in V_0$ and $\|xR\|_{TV} \leqslant (1-\varepsilon)\|x\|_{TV}$.*

This follows immediately from the opening arguments in the proof of Theorem 2.2.1 in Stroock [2013] by setting $n = 1$.[43]

Lemma B.2. *If $\{R_t\}_{t=1}^m \subseteq \mathcal{T}_\varepsilon$, then for each $x \in V_0$*

$$
x\left(\prod_{t=1}^{m} R_t\right) \in V_0 \text{ and } \left\|x\left(\prod_{t=1}^{m} R_t\right)\right\|_{TV} \leqslant (1-\varepsilon)^m \|x\|_{TV}. \tag{20}
$$

Proof of Lemma B.2. We proceed by induction.

Initial Step. The statement follows from Lemma B.1.

Inductive Step. We assume the statement holds for $m \in \mathbb{N}$. We show it holds for $m+1$. Define $y = x\left(\prod_{t=1}^{m} R_t\right)$. By Lemma B.1 and since $R_{m+1} \in \mathcal{T}_\varepsilon$, this implies that $y \in V_0$, $yR_{m+1} \in V_0$ and $x\left(\prod_{t=1}^{m+1} R_t\right) = x\left(\prod_{t=1}^{m} R_t\right) R_{m+1} = yR_{m+1} \in V_0$. By inductive

[43] Given a sequence $\{P_t\}_{t\in\mathbb{N}}$ of stochastic matrices, $\prod_{t=1}^{n+1} P_t = \prod_{t=1}^{n} P_t P_{n+1}$ for all $n \in \mathbb{N}$

hypothesis and Lemma B.1, this implies that

$$\begin{aligned}\left\|x\prod_{t=1}^{m+1}R_t\right\|_{TV} &= \|yR_{m+1}\|_{TV} \leqslant (1-\varepsilon)\|y\|_{TV} = (1-\varepsilon)\left\|x\left(\prod_{t=1}^{m}R_t\right)\right\|_{TV}\\ &\leqslant (1-\varepsilon)(1-\varepsilon)^m\|x\|_{TV} = (1-\varepsilon)^{m+1}\|x\|_{TV},\end{aligned}$$

proving the inductive step. □

Given Lemma B.2 the proposition follows by a straightforward repetition of the arguments in the proof of Theorem 4.14 of Seneta [2006] with the only difference that the state space is infinite but countable.

B.8 Stability

Definition 10. A strategy σ is *stable* if for every $\varepsilon \in (0,1)$, and every prior ν with support Θ there is an n such that if σ is used in the first n periods the probability that the best reply at period $n+1$ is σ is larger than $1-\varepsilon$.

Proposition 10. *Every uniformly strict selective memory equilibrium is stable.*

Proof of Proposition 10. If σ is a uniformly strict selective memory equilibrium, Lemma A.2 implies that there are $\hat{C} \in \mathbb{R}_{++}$ and ε such that for all $s \in S$ if $\nu(KL_s^{\Theta,m}(\sigma,\varepsilon)) > \hat{C}(1-\nu(KL_s^{\Theta,m}(\sigma,\varepsilon))$ then $\{\sigma(s)\} = BR(s,\nu)$.

By the Law of Large Numbers and the finiteness of $S \times A \times Y$, for every C there exists $n_1 \in \mathbb{N}$ such that if σ is used for the first n periods, $n > n_1$

$$\mathbb{P}\left(||f(h_{n_1})(s,a,y) - f^*(s,a,y)|| < C\right) > 1-\frac{\varepsilon}{2} \tag{21}$$

where

$$f^*(s,a,y) = \begin{cases}\zeta(s)p^*_{a,s}(y) & \text{if } a=\sigma(s)\\ 0 & \text{otherwise}\end{cases}.$$

Define

$$\tilde{p}(\sigma,s')(s,a,y) = \frac{\zeta(s)m_{s'}(s,\sigma(s),y)p^*_{\sigma(s),s}(y)}{\sum_{\hat{y}\in Y,\hat{s}\in S}\zeta(\hat{s})m_{s'}(\hat{s},\sigma(\hat{s}),\hat{y})p^*_{\sigma(\hat{s}),\hat{s}}(\hat{y})}$$

if $a=\sigma(s)$ and $\tilde{p}(\sigma,s')(s,a,y) = 0$ otherwise. Since for every τ' and τ with $\tau' > \tau$ and every history $B \subseteq \{1,...,\tau'\}$, $\mathbb{P}[r_{\tau'} = B|h_{\tau'}] = \mathbb{P}[r_{\tau'} = B|h_{\tau'}, r_\tau]$, by the second

Borel-Cantelli lemma and equation (21), for every $\varepsilon \in \mathbb{R}_{++}$, $s' \in S$ and $k \in \mathbb{N}_{++}$ there is a $n_{k,C} > n_1$ such that for all $\tau \geqslant n_{k,C}$

$$\mathbb{P}\left[|r_\tau| > k \text{ and } ||\hat{f}(h_\tau, r_\tau) - \tilde{p}(\sigma, s')||_\infty < C\right] > 1 - \varepsilon. \tag{22}$$

As in the proof of Theorem 1 choose κ, $C' \in \mathbb{R}_{++}$ and $\varepsilon' < \varepsilon$ such that

$$\begin{aligned}\kappa < \inf_{\substack{f:\forall s' \in S \\ ||f - \tilde{p}(\sigma,s')|| < C'}} \Bigg(& \inf_{\{p' \notin KL^{\Theta,m}_{s'}(\sigma,\varepsilon)\}} \left(- \sum_{s \in S, a \in A, y \in Y} f(s,a,y) \log p'_{a,s}(y) \right) \\ & - \sup_{\{p' \in KL^{\Theta,m}_{s'}(\sigma,\varepsilon')\}} \left(- \sum_{s \in S} \zeta(s) \sum_{y \in Y} p^*_{\sigma(s),s}(y) \log p'_{\sigma(s),s}(y) \right) \Bigg).\end{aligned}$$

On the set of remembered histories where $\left\{|r_\tau| > k \text{ and } ||\hat{f}(h_\tau, r_\tau) - \tilde{p}(\sigma, s')||_\infty < C'\right\}$ identified by equation (22), Lemma A.1 implies that

$$\begin{aligned}&\frac{\nu(KL^{\Theta,m}_{s'}(\sigma,\varepsilon)|h_t)}{1 - \nu(KL^{\Theta,m}_{s'}(\sigma,\varepsilon)|h_t)} \\ &\geqslant \frac{\nu\left(KL^{\Theta,m}_{s'}(\sigma,\varepsilon)\right)}{\left(1 - \nu\left(KL^{\Theta,m}_{s'}(\sigma,\varepsilon')\right)\right)} \exp\left(|r_\tau| \left(\begin{array}{c} -\sup_{\{p' \in KL^{\Theta,m}_{s'}(\sigma,\varepsilon')\}} D(\hat{f}(h_\tau, r_\tau)|| \log(p')) \\ +\inf_{\{p' \notin KL^{\Theta,m}_{s'}(\sigma,\varepsilon)\}} D(\hat{f}(h_\tau, r_\tau)|| \log(p')) \end{array} \right)\right) \\ &\geqslant \frac{\nu\left(KL^{\Theta,m}_{s'}(\sigma,\varepsilon)\right)}{\left(1 - \nu\left(KL^{\Theta,m}_{s'}(\sigma,\varepsilon')\right)\right)} \exp(k\kappa).\end{aligned}$$

The result follows by setting $n = n_{k',C'}$ where $k' = \left\lfloor \log\left(\frac{\left(1-\nu\left(KL^{\Theta,m}_{s'}(\sigma,\varepsilon')\right)\right)}{\nu\left(KL^{\Theta,m}_{s'}(\sigma,\varepsilon)\right)} \hat{C} \right) / \kappa \right\rfloor + 1$. □

B.9 Partially Recalled Histories with Partial naïveté

Here we suppose that the outcome space has a product structure, i.e., $Y = \times_{i \in I} Y_i$ and that the agent may recall only some components of the outcome. Moreover, we continue to allow for partial naïveté as in Section A.9. To model this case, we use a collection of signal-dependent *memory functions* $m_{s'} : (S \times Y \times 2^I) \to [0,1]$, where $m_{s'}(s, y, B)$ specifies the probability with which an agent remembers the B outcome components of a past realization of the signal, outcome pair (s, y) and with

$$\sum_{B \in 2^I} m_{s'}(s, y, B) = 1.$$

Moreover, the agent believes that they remember an occurrence of signal s and outcome y with probability $\hat{m}_{s'}(s, y, B)$. The case studied in the main body of the paper is the one in which $m_{s'}(s, a, y, B) = 0$ for every $B \notin \{\varnothing, I\}$ with the interpretation that $m_{s'}(s, a, y, \varnothing)$ means that the agent does not recall the experience (s, a, y) at all. Thus the *recalled history* at time t is the of recalled experiences $r_t = (s_\tau, y_\tau, B_{\tau,t})_{\tau=1}^t$ where $B_{\tau,t}$ denotes the components of the period τ outcome recalled at time t. We obtain

$$\mu(C|r_t, s') = \frac{\int_{p \in C} \prod_{\tau=1}^t \hat{m}_{s'}(s_\tau, \prod_{i \in I} \tilde{Y}_{\tau,i}, B_{\tau,t}) p_{s_\tau}(\prod_{i \in I} \tilde{Y}_{\tau,i}) d\mu(p)}{\int_\Theta \prod_{\tau=1}^t \hat{m}_{s'}(s_\tau, \prod_{i \in I} \tilde{Y}_{\tau,i}, B_{\tau,t}) p_{s_\tau}(\prod_{\tau=1}^t \tilde{Y}_{\tau,i}) d\mu(p)} \qquad \forall C \subseteq \Theta$$

where $\tilde{Y}_{\tau,i} = Y_i$ if $i \notin B_{\tau,t}$ and $\tilde{Y}_{\tau,i} = \{y_{\tau,i}\}$ if $i \in B_{\tau,t}$. With this, the results of the paper carries through with the following adaptation of the concept of memory-weighted KL-minimizers:

$$KL_{s'}^{\Theta,m}(\sigma) = \operatorname*{argmin}_{p \in \Theta} - \sum_{s \in S} \zeta(s) \sum_{B \in 2^I} \sum_{y \in Y} m_{s'}(s, y, B) p_s^*(\tilde{Y}(y)) \log \hat{m}_{s'}(s, y, B) p_s(\tilde{Y}(y))$$

where $\tilde{Y}(y) = Y_i$ if $i \notin B$ and $\tilde{Y}(y) = \{y_i\}$ if $i \in B$.

Example 10 (Ego-Boosting Memory plus Cognitive Dissonance Reduction)**.** *Suppose that there two tasks, each of which the agent can either pass or fail, i.e.,* $Y_1 = Y_2 = \{0, 1\}$*, and there is no signal. The agent is more likely to recall successes in each component, but they are also more likely to recall the outcome of task 2 (a secondary task) if it confirms the outcome of the first task. For example, we could have*

$$\begin{aligned}
&m((1,1), \{1,2\}) = 1, \\
&m((1,0), \{1,2\}) = 0.1, \quad m((1,0), \{1\}) = 0.8, \quad m((1,0), \varnothing) = 0.1 \\
&m((0,1), \{2\}) = 0.7, \quad m((0,1), \varnothing) = 0.3 \\
&m((0,0), \varnothing) = 0.9, \quad m((0,0), \{1,2\}) = 0.1.
\end{aligned}$$

As in the case with a unique component, a partially naïve agent can be described by a perceived memory function that combines perfect memory with the true memory func-

tion:

$$\hat{m}(y, B) = \alpha m(y, B) + (1-\alpha)1 \quad \forall y \in Y, B \subseteq \{1, .., I\}.$$

Suppose that the initial belief of the agent is that the probability of success is independent and equal across tasks and is either $p = 0.9$ or $p' = 0.1$ with equal prior probability. Then after one period, if there was success only in task one, if the agent only recalls component 1 their posterior belief is

$$\begin{aligned}
&\mu(p|((1,0),\{1\}) \\
= &\frac{[\alpha m((1,1),\{1\}) + (1-\alpha)]\, p\,(1,1) + [\alpha m((1,0),\{1\}) + (1-\alpha)]\, p\,(1,0)}{\begin{array}{c}[\alpha m((1,1),\{1\}) + (1-\alpha)]\, p\,(1,1) + [\alpha m((1,0),\{1\}) + (1-\alpha)]\, p\,(1,0) + \\ [\alpha m((1,1),\{1\}) + (1-\alpha)]\, p'\,(1,1) + [\alpha m((1,0),\{1\}) + (1-\alpha)]\, p'\,(1,0)\end{array}} \\
= &\frac{(1-\alpha)0.9^2 + (0.8\alpha + (1-\alpha))\, 0.9 * 0.1}{(1-\alpha)0.9^2 + (0.8\alpha + (1-\alpha))\, 0.9 * 0.1 + (1-\alpha)0.1^2 + (0.8\alpha + (1-\alpha))\, 0.9 * 0.1}.
\end{aligned}$$

In particular, a completely sophisticated agent ($\alpha = 1$) ends up with a posterior equal to the prior, as they understand that the fact that they do not recall the second component means it was a failure, and success in one dimension and failure in the other leaves the prior unchanged. A completely naïve agent ($\alpha = 0$) instead ends up with a posterior probability of 0.9 for the optimistic distribution p.

B.10 Permanent Memories

Suppose that the memory function m determines the probability that a particular experience is recalled in the period just after it occurs. If is is recalled it is never forgotten, and if it is not recalled it is never remembered. Then the belief process has the recursive formula:

$$\mu_{t+1}(C) = \begin{cases} \frac{\int_{p \in C} p_{a_t}(y_t) d\mu_t(p)}{\int_{p \in \Theta} p_{a_t}(y_t) d\mu_t(p)} & \text{with probability } m(a_t, y_t) \\ \mu_t(C) & \text{with probability } (1 - m(a_t, y_t)). \end{cases}$$

It is easy to see that if the actions in this dynamic system converge, they converge to a selective memory equilibrium. However, just as with in Example 9 on underfinference, the fact that memory here is "less stochastic" allows some selective memory equilibria to be limit actions that are not there limit actions under our main baseline selective memory process.

B.11 Examples

Example 11. Let $S = \{s_{-1}, s_1\}$, $A = \{-1, 1\}$, $Y = \{-1, 1\}$, $u(a, y) = ay$, $p^*_{s_{-1},a}(-1) = 3/4$, $\zeta(s_{-1}) = \zeta(s_1)$, $p^*_{s_1,a}(-1) = 1/5$, and $\Theta = \{p : p_{s,a} = p_{s',a'}$ for all $s, s' \in S, a, a' \in A\} \in \Theta$. Under perfect memory the unique equilibrium is to always play -1. If the agent has similarity weighted memory with $f(d(s_{-1}, s_1)) = 1/10$, $f(0) = 1$, then the unique selective memory equilibrium is the objectively optimal $\sigma(s_{-1}) = -1$, $\sigma(s_1) = 1$. ▲

Example 12. In the setting of Proposition 4, let $Y = \{0, 2.5, 4, 8\}$ with $y_0 = 2.5$, $p^*_1(0) = p^*_1(4) = p^*_1(8) = 1/3$, and $v(y) = \sqrt{y}$. Then the unique selective memory equilibrium with perfect memory is to play the risky lottery. However, under the extreme event bias where $m(0) = m(8) = 1$, $m(3) = 1/2$, $m(4) = 1/10$ the unique selective memory equilibrium is to play the safe action. ▲