Selective Memory Equilibrium*

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

September 6, 2023

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. We show that if the agent's behavior converges, their limit strategy is a *selective memory equilibrium*, and we provide a sufficient condition for behavior to converge. We illustrate how this new equilibrium concept can be used to understand the long-run effects of several well-documented memory biases.

JEL codes: D83, D90

*We thank Chiara Aina, Ian Ball, Abhijit Banerjee, Douglas Bernheim, Pierpaolo Battigalli, Roland Benabou, Dirk Bergemann, Jonathan Brownrigg, Simone Cerreia-Vioglio, Gary Charness, Roberto Corrao, Martijn de Vries, Glenn Ellison, Ignacio Esponda, Erik Eyster, Thomas Graeber, Jerry Green, Philippe Jehiel, Mats Koster, Roger Lagunoff, Shengwu Li, Stephen Morris, Ryan Oprea, Parag Pathak, Jêrome Renault, 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. We thank Neil He for excellent research assistance.

[†]Department of Economics, MIT
[‡]Department of Economics, Harvard University
[§]Department of Economics, Yale University

1 Introduction

We provide a new conceptual framework for the study of agents who have selective memory in 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.[1] 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.[2] These assumptions fit evidence from both experimental and real-world settings. 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 memory formation and retrieval process. Instead, we develop a tractable model of how selective memory shapes behavior and beliefs in decision problems. This 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 concentrate on the *memory-weighted likelihood maximizers*, i.e., distributions that maximize the likelihood of a distorted version of the true outcome distribution that gives more weight to realizations that are more likely to be remembered. We also provide conditions on the agent's payoff function and the support of their prior that imply their behavior does converge. Whether or not these conditions are satisfied, when behavior converges, it converges to a *selective memory equilibrium*, which is a strategy that myopically maximizes their expected payoff against a probability distribution over these maximizers. 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 lets us analyze the long-run consequences of important and widely documented forms of selective memory such as pleasant memory bias (Mischel, Ebbe-

[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 than others.

[2]Reder [2014], Zimmermann [2020], Gödker, Jiao, and Smeets [2022] provide evidence of partial or complete unawareness of memory biases. Appendix (A.3) shows that if agents are aware they sometimes forget but are not aware that their memory is selective the main results extend as stated.

sen, and Zeiss [1976], Adler and Pansky [2020], Chew, Huang, and Zhao [2020] and the related ego-boosting bias, Zimmermann [2020]), 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 [2022], Goetzmann, Watanabe, and Watanabe [2022]), confirmatory bias (Hastie and Park [1986]), and the relative memorability 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 induce 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] finds that investors are more likely to 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 dogmatic 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 make the negative feedback be recalled more often, which leads to less bias in long-run beliefs.[3]

Agents with *associative memory* are more likely to recall situations similar to the current decision problem, for example, when they had a similar mood. In general, this can lead the agent to underweight data relative to its true informativeness. However, the simplest version of associativeness, similarity weighting (Kahana [2012]), does not alter the possible long-run outcomes for a correctly specified agent: With similarity weighting, all selective memory equilibria are (unitary) self-confirming equilibria (Battigalli [1987] and Fudenberg and Levine [1993a]).

We also study *extreme experience bias*, which makes experiences with more extreme payoffs more memorable. We 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

[3]This is suggested in the management literature by e.g., Procházka, Ovcari, and Durinik [2020].

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 [2021]) is equivalent to a uniformly strict selective memory equilibrium for some memory function and a full-support prior, and 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 equilibria that are not uniformly strict.[4] In addition, unlike Berk-Nash equilibria, selective memory equilibria generally do not reduce to self-confirming equilibria 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. To illustrate this, we show that combining positive and negative feedback has qualitatively different effects on agents with ego-boosting memory than on dogmatically overconfident agents.

Related Theoretical Work Mullainathan [2002] studies selective memory where the probability of recalling an observation is the sum of a base rate, an "associativeness" term that measures the experience's similarity to the current observation, and a "rehearsal" term that indicates 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 Gaussian and are not influenced by the agent's actions. Afrouzi, Kwon, Landier, Ma, and Thesmar [2020] also studies an agent 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, Coffman, Gennaioli, Schwerter, and Shleifer [2021] shows how memory depends on the relative frequency of various characteristics, and can be manipulated by making some observations stand out more. None of these papers addresses our question of determining the agent's long-run beliefs and actions.

There is also a set of papers that study long-run behavior with selective attention

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

or recording, where whether an experience is recorded determines whether it will be recalled in every future period, as in the model in Online Appendix B.5. Compte and Postlewaite [2004] considers a myopic agent with the choice between a safe action with a known payoff and a risky action whose outcome distribution is unknown. It assumes that taking the risky action is sometimes a dominant strategy so that the agent will eventually take it infinitely often, and that periods with good performance are more likely to be recorded. This leads to overconfidence, as in our example in Section 4.2. 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 they did not pay attention to some relevant aspects of what they observed. As with selective memory and misspecified beliefs, this can lead the agent to make systematically biased forecasts. Relatedly, Schweizer and De Vries [2022] assumes that for exogenous reasons, the agent weights outcomes differently depending on how extreme they were (compared to other outcomes) at the time they realized and shows this can lead to probability distortion analogous to that of cumulative prospect theory, similar to a selective memory with rare experience bias (see Section 4.3).

Wilson [2014] and Jehiel and Steiner [2020] study the optimal use of a finite memory in 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' objective information and memory in games. Bénabou and Tirole [2002] considers a two-period model where a time-inconsistent agent receives either a null signal or a bad signal in the first period, and at a cost can change the probability that the second-period self recalls the bad signal. The resulting game can have either a unique equilibrium or three of them; in some cases, the first-period agent distorts their memory to become overconfident. 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, and Nagel and Xu [2022] analyzes an asset pricing model where the representative agent has fading memory.

2 Setup

We study a sequence of choices made by a single agent. In every period $t \in \mathbb{N}$ 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 $p^*_{s,a} \in \Delta(Y)$ over the finite set of possible outcomes Y.[5] A (pure) strategy is a map $\sigma : S \to A$, and the agent's flow payoff is given by the utility function $u : S \times A \times Y \to \mathbb{R}$.

We assume the agent knows the fixed and i.i.d. full-support distribution $\zeta \in \Delta(S)$ over signals.[6] 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 are uncertain about the outcome distributions each signal-action pair induces. To model this uncertainty, we suppose that the agent has a prior μ over data-generating processes $p \in \Delta(Y)^{S\times A}$, where $p_{s,a}(y)$ denotes the probability of outcome $y \in Y$ when signal s is observed and action a is played. The support of μ 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*. To simplify the exposition, we will assume throughout the paper that selective-memory agents are correctly specified, but this is not essential; all results except for Proposition 1 are true as stated without that assumption.

Assumption 1 (Maintained Assumption)**.**

(i) The agent is correctly specified.

(ii) There is $p \in \Theta$ such that for all $y \in Y$, $a \in A$, and $s \in S$, $p^*_{s,a}(y) > 0$ only if $p_{s,a}(y) > 0$.

We maintain this assumption throughout the paper. The second part of the assumption implies that the agent never observes an event that has zero probability under their prior.

Objective Histories and Recalled Histories We assume that the agent always recalls the signal they just observed. The agent's memory of the outcomes corresponding to past signal-action pairs is distorted by a collection of signal-dependent *memory*

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

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

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*.

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 objective history $h_t = (s_\tau, a_\tau, y_\tau)_{\tau=1}^t$ and signal s_{t+1}, the *recalled periods* R are a random subset of $\{1, ..., t\}$. Period τ is remembered with probability $m_{s_{t+1}}(s_\tau, a_\tau, y_\tau)$, independently of which other periods are remembered. Thus, the distribution over subsets $R \subseteq \{1, .., t\}$ is given by

$$\mathbb{P}\left[R | (s_\tau, a_\tau, y_\tau)_{\tau=1}^t, s_{t+1}\right] = \prod_{\tau \in R} m_{s_{t+1}}(s_\tau, a_\tau, y_\tau) \prod_{\tau \in \{1,...,t\} \setminus R} (1 - m_{s_{t+1}}(s_\tau, a_\tau, y_\tau)).$$

For every objective history h_t and set of recalled periods R, the *recalled history* $h_t(R) \in H_{|R|}$ is the subsequence of recalled experiences listed in the order in which they realized.[7]

Beliefs We assume the agent recomputes their beliefs each period based on all of their recollections, as opposed to simply updating their period-t beliefs on the basis of their period-t observation, and 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,[8] so that the posterior probability of every (measurable) $C \subseteq \Theta$ after recalled history $h_t = (s_\tau, a_\tau, y_\tau)_{\tau=1}^t$ is

$$\mu(C|h_t) = \frac{\int_C \prod_{\tau=1}^t p_{s_\tau, a_\tau}(y_\tau) d\mu(p)}{\int_\Theta \prod_{\tau=1}^t p_{s_\tau, a_\tau}(y_\tau) d\mu(p)}. \tag{1}$$

We briefly discuss partial naïveté in the conclusion and analyze it in Appendix A.3, where we show that if agents recognize that their memory is faulty but believe it is not selective, the main results extend as stated.[9]

[7] Because the environment is i.i.d., the fact that the order of the consequences is correctly recalled is irrelevant and only made for convenience. Appendix A.1 formally describes the map from objective histories and recalled periods to recalled histories.

[8] See, e.g., d'Acremont, Schultz, and Bossaerts [2013] for fMRI evidence that agents access their accumulated evidence each period when updating beliefs, and Reder [2014] for evidence that agents are often näive about their selective memory and do not make inferences about their forgotten observations from the actions they remember taking.

[9] Appendix A.3 maintains our assumption that the agent either remembers an experience perfectly or not at all. We relax this in Online Appendix B.4, where the agent may remember some but not all aspects of a past experience, such as one or two components of a multi-dimensional outcome. That

Best Responses and Optimal Policies The agent's belief determines the subjective expected utility of each action. Denote by $BR(s, \nu)$ the actions that maximize expected utility when signal s is observed and the agent's belief is $\nu \in \Delta(\Theta)$:[10]

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

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 every $s \in S$ and recalled history $h_t \in H$, $\pi(h_t)(s) \in BR(s, \mu(\cdot|h_t))$.

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 $s, s' \in S, y, y' \in Y$, and $a \in A$.

Example 1 (Utility-Dependent Memory)**.** In some cases, the probability of remembering an experience depends on its associated utility, so that $m_{s'}(s, a, y) = \Phi(u(s, a, y))$ for some $\Phi : \mathbb{R} \to \mathbb{R}_+$. 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.[11] ▲

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 $y_1 \in \mathbb{R}$ 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 $\Phi : \mathbb{R} \to \mathbb{R}_+$. ▲

Example 3 (Cognitive Dissonance and Ex-post Regret)**.** *Cognitive dissonance* is a memory bias where the probability of recalling an experience depends on how well the

model assumes the agent is not fully naïve, because remembering that some experience occurred but not all of its details might lead the agent to question their ability to perfectly recollect the past.

[10]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.

[11]Because agents never make choices before the signal realizations, there is no way to pin down the relationship between the utilities of two experiences that differ in their signal component. Therefore, both here and in Example 3, the definitions of the biases should be interpreted as saying that there are a u and a Φ such that the conditions are satisfied.

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. If the outcome includes the payoff that would have been obtained with each action, the probability of remembering an outcome is decreasing in what Loomes and Sugden [1982] called "regret" (see Lanzani [2022] for the version without a state space that formally corresponds to the case we have here). ▲

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]). 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].[12]

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 $d : S^2 \to \mathbb{R}_+$, and $m_{s'}(s,a,y) = \Phi(d(s,s'))$ for some strictly decreasing function $\Phi : \mathbb{R}_+ \to [0,1]$. ▲

Example 5 (Confirmatory Memory Bias)**.** The agent has *confirmatory memory bias* (see Hastie and Park [1986] and Esponda, Vespa, and Yuksel [2023] 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 hypotheses, as in Lord, Ross, and Lepper [1979] and Rabin and Schrag [1999], so that $\Theta = \{p^0, p^1\}$,

[12]Jehiel [2018] studies investors who make their decisions based only on the outcomes of projects that were implemented after the same signal and ignore periods when the signal was different, and Bordalo, Gennaioli, and Shleifer [2020] shows how similarity weighting can lead to the attribution and projection biases.

with $\mu(p^0) > \mu(p^1)$. Then, confirmatory memory bias corresponds to

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

3 Long-Run Outcomes

Let $\mathbb{P}_\pi$ denote the probability measure on sequences of experiences $(S \times A \times Y)^{\mathbb{N}}$ induced by the objective action-contingent outcome probability, the agent's memory m, and policy π.[13]

Definition 1. 1. A strategy σ is a *limit strategy* if there is an optimal policy π such that

$$\mathbb{P}_\pi\left[\sup\{t\colon a_t \neq \sigma(s_t)\} < \infty\right] > 0.$$

2. σ is a *global attractor* if for every optimal policy π

$$\mathbb{P}_\pi\left[\sup\{t\colon a_t \neq \sigma(s_t)\} < \infty\right] = 1.$$

In words, strategy σ is a limit strategy if there is positive probability that it will be played in every period after some random but finite time, and it is a global attractor if it is a limit with probability 1. This section gives some general results about limit strategies. Section 4 then discusses the consequences of some 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 likelihood maximizers* after signal s':

$$ML_{s'}(\sigma)\colon = \operatorname*{argmax}_{p\in\Theta}\left(\sum_{s\in S}\zeta(s)\sum_{y\in Y} m_{s'}(s,\sigma(s),y)p^*_{s,\sigma(s)}(y)\log p_{s,\sigma(s)}(y)\right). \tag{2}$$

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

[13] This is the measure obtained as the unique extension from the probabilities of the finite histories $(S \times A \times Y)^t, t \in \mathbb{N}$.

Definition 2. A strategy σ is a

(i) *Selective memory equilibrium* if for all $s \in S$ there is $\nu_s \in \Delta(ML_s(\sigma))$ such that $\sigma(s) \in BR(s, \nu_s)$.

(ii) *Uniformly strict selective memory equilibrium* if for all $s \in S$ and all $\nu \in \Delta(ML_s(\sigma))$, $\{\sigma(s)\} = BR(s, \nu)$.

In a selective memory equilibrium σ, the action played after each signal s is a best reply to some belief over memory-weighted likelihood maximizers given σ. The uniformly strict version adds the restriction that there is the same unique best reply for each of these maximizers. 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 of outcomes conditional on each signal and the distribution of recalled experiences after each signal s converge to deterministic limits, and the best-fitting models are the elements of $ML_s(\sigma)$.

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 likelihood maximizers. A compactness-continuity lemma guarantees that this also holds for beliefs concentrated on a neighborhood of the maximizers, so it is enough to show that beliefs concentrate on these neighborhoods. A preliminary lemma shows that if σ is a limit strategy, then for some time t, there is an action sequence a^t such that if the agent plays a^t and then σ afterward, there is positive probability that the induced sequence of beliefs makes σ optimal at all periods $\tau \geqslant t+1$. Under a policy that converges to σ, 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 long and representative. With this, we can extend Berk [1966]'s concentration result to the product of actual and recalled experiences to show that distributions that don't maximize the memory-weighted likelihood have vanishing posterior probability under representative recalled histories.

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.

We also provide a foundation for Berk-Nash equilibrium 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. Finally, Section 5.1 shows that in our setting, the long-run action of an agent with underinference (Phillips and Edwards [1966]) must be a selective memory equilibrium.

3.2 Global Convergence to Equilibrium

We now give a sufficient condition for the agent's strategy to globally converge to a selective memory equilibrium.

Assumption 2.

(i) For all $(s, s', a) \in S^2 \times A$, $\sum_{y \in Y} p^*_{s,a}(y) m_{s'}(s, a, y) > 0$.

(ii) There is a $\hat{p} \in \Theta$ such that for every $(s, a) \in S \times A$, $ML_s(a) = \{\hat{p}\}$.

The substantial assumption here is Assumption 2 (ii), which implies that there exists a unique memory-weighted likelihood maximizer $\hat{p}$ that is independent of the action and the signal. This assumption is always satisfied if the agent correctly believes their actions have no influence on the distribution of outcomes and has the same memory function for each action, as in the examples in Section 4. Beyond that, the result needs there to be an outcome with positive probability of being remembered for each (s, a) pair. The first part of our next result considers closed balls around the data-generating process $\hat{p}$, where the distance between two data-generating processes is the maximum of the total variation distance between their signal-action contingent distributions:

$$B_\epsilon(\hat{p}) = \left\{ q \in \Delta(Y)^{A\times S} \colon \max_{s\in S, a\in A} ||q_{s,a} - \hat{p}_{s,a}||_{TV} < \epsilon \right\}.$$

Theorem 2. *Under Assumption 2, for every optimal strategy π, and every $\epsilon > 0$*

$$\mathbb{P}_\pi \left[\lim_{t\to\infty} \mu(B_\epsilon(\hat{p})|h_t(R_t)) = 1 \right] = 1.$$

If in addition $BR(s, \delta_{\hat{p}})$ is a singleton for all s, then $\hat{\sigma}(s) = BR(s, \delta_{\hat{p}})$ is a global attractor, and so it is a selective memory equilibrium.

The proof starts by using a mixingale law of large numbers to conclude that the outcome frequency converges to the one predicted by the true data-generating process and the agent's actions. We then address the complication posed by the fact that because memory is stochastic, even when the agent has played many times, their beliefs can be very different from one period to the next, unlike in learning models with perfect memory. We first use the Chernoff inequality to provide an upper bound on the probability that the recalled empirical frequency significantly diverges from the memory-distorted version of the actual empirical frequency. This upper bound is then combined with the Borel-Cantelli lemma to show that there is a $\gamma \in (0, 1)$ and a random but a.s. finite time after which any signal-action pair with frequency at least γ doesn't have a large deviation from the memory-distorted empirical distribution of its induced consequences and has recalled frequency at least $\gamma/2$.

Since for every signal-action pair, $\hat{p}$ is the unique model providing this best fit, the fact that most of the recalled memories are representative combined with exponential convergence of beliefs to the data established in Fudenberg, Lanzani, and Strack [2022] together imply that beliefs concentrate on $\hat{p}$. By the uniqueness of the best reply, this implies that the behavior converges as well.

Remark 1. 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 fits the evidence on memory retrieval, see, e.g., Kahana [2012]. Online Appendix B.5 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; if not, it is never remembered. Because experiences recalled at later dates include all those 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.

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 for all $s \in S$ there is $\nu_s \in \Delta(\Theta)$ such that for all $p \in \operatorname{supp} \nu_s$ and $s' \in S$, $p_{s',\sigma(s')} = p^*_{s',\sigma(s')}$ and $\sigma(s) \in BR(s, \nu_s)$.

Self-confirming equilibrium requires that agents have correct beliefs about the consequences of their equilibrium strategy but allows them to have incorrect beliefs about strategies they do not use. Fudenberg and Levine [1993b] shows that these equilibria correspond to the steady states of a learning model with long-lived but myopic agents.[14]

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

More generally, this conclusion holds whenever $m_{s'}(s, a, y)$ does not depend on a or y, as 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.[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 leads 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.

[14]More generally, if the agents have some discount factor that is strictly between 0 and 1, a self-confirming equilibrium that is not a Nash equilibrium can be a steady state if the priors are sufficiently concentrated in the neighborhood of that equilibrium.

[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].

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. There are no signals, A is endowed with a linear order, and the agent (correctly) believes their action does not affect the realized outcome. The next proposition shows that a larger bias leads to a more positive limit belief and higher limit action. This provides a selective memory foundation for the positive correlation between an agent's 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 is supermodular, and that $\Theta = \Delta(Y)$. The agent's long-run belief with memory m' concentrates on a distribution of outcomes weakly 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 weakly higher than the limit action with memory m.*

Intuitively, because the prior has full support, the memory-weighted likelihood maximizer 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 weakly higher than the objective optimum.

Example 6. 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

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

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 5/9. Consequently, they will behave like an exogenously misspecified agent who dogmatically believes their ability to pass is at least 5/9. 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 at least in this experiment selective memory is a better explanation than selective attention for long-run overconfidence.

Example 6 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 memory 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 7. Suppose that, besides taking an IQ test, the agent works on a project with a coworker. The outcome distributions $(p, q) \in [0, 1]^2$ 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^*, and the project succeeds with probability $\alpha p^* + (1 - \alpha)q^*$, where α is the share of the work done by the coworker. The agent

[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.

always remembers experiences with positive IQ test results and remembers experiences with negative test results with probability $\phi \in (0, 1)$. Thus, beliefs concentrate on

$$p = p^* + \frac{p^*(1-p^*)}{\phi/(1-\phi)+p^*} \qquad q = q^* - \frac{\alpha}{1-\alpha}\frac{p^*(1-p^*)}{\phi/(1-\phi)+p^*}:$$

The agent 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^*. The second component y_2 is drawn independently with probability $\alpha p^*(y_2) + (1-\alpha)q^*(y_2)$ for some $\alpha \in (0, 1)$. The agent knows that the outcomes are independently drawn each period according to these conditions, but does not know p^* or q^*, and their prior belief assigns positive probability to each of these distributions.[18]

Proposition 3. *If 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 concentrates on a distribution $\hat{p}$ that is weakly higher in first-order stochastic dominance than p^*, and the agent's long-run belief about q concentrates on a distribution that is weakly lower than q^*.*

"Perhaps the most robust finding in the psychology of judgment is that people are overconfident." (DeBondt and Thaler, 1995, p. 389). The proposition provides an explanation for two commonly found forms of overconfidence: (i) overestimation of one's own absolute level of performance and (ii) overestimate of performance relative to others (see, e.g., Svenson, 1981; Merkle and Weber, 2011). For example, Gilovich [2008] finding that 94% of the college professors thought they were better than their average colleague.[19]

Reinforcement through Actions Actions can play an important role in amplifying the misconceptions caused by selective memory. For example, suppose that in the setting of Example 7 the agent starts out with an unbiased belief about their coworker's ability, and each period t chooses the fraction α_t of work to delegate to them. Because actions here do influence the agent's observations and inference, Theorem 2 does

[18] Formally, $\Theta = \{r \in \Delta(Z \times Z) : r(y_1, y_2) = p(y_1)[\alpha p(y_2) + (1-\alpha)q(y_2)]$ for some $p, q \in \Delta(Z)\}$,

[19] Benoıt and Dubra [2011] shows how this "I'm better-than-average effect" can be explained within a purely Bayesian framework; Benoıt, Dubra, and Moore [2015] provides more direct evidence for relative overconfidence that rules out the purely rational explanation.

not apply, but as in Heidhues, Kőszegi, and Strack [2018]'s analysis of exogenously overconfident agents, there is a global attractor: 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 their coworker's. The agent thus delegates delegate less work to their coworker, whose ability then has a smaller effect on output. To explain the disappointingly low output, the agent becomes even more pessimistic about the coworker's ability, leading to even less delegation in the unique limit strategy.

More generally, Section 5.2 shows that the long-run belief induced by selective memory equilibria can be replicated by exogenous misspecification in any fixed environment, and vice versa. However, selective memory and exogenous misspecification can lead to very different predictions about the effect of changes in information. 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 makes a "feedback sandwich," which the management and psychology literatures suggest strengthens the effect of feedback.[20] 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, so their long-run belief would move closer to their actual ability, and they would be less biased about their coworker's ability. In contrast, with exogenous misspecification, 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 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

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

value and does not depend on s or a:

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

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

Our next result shows that the risky lottery is a selective memory equilibrium with extreme experience bias only if it is a selective memory equilibrium with perfect memory. Moreover, Example 11 in the Online Appendix shows that extreme 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.[21]

Proposition 4. *Suppose p_1^* is symmetric and that the agent thinks all outcome distributions are possible under the risky action.*[22] *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. By making the tail realizations relatively more memorable, extreme experience bias makes 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 when the stock market crashed might be more easily remembered than many days of average returns, leading 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]).

[21]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.

[22]That is, $p_1^*(\bar{y} + c) = p_1^*(\bar{y} - c)$ for all $c \in \mathbb{R}$, and $\Theta = \{p \in \Delta(Y)^A : p_0(y_0) = 1\}$.

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

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

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

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

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

for every measurable $C \subseteq \Theta$, 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 memory-weighted likelihood maximizers, and as the next result shows the underinference distortion maps directly to a selective memory function.[23]

[23] We identify the underinference distortion with the vector of memory functions that do not depend on the current signal.

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.[24] Gervais and Odean [2001] studies a different underinference bias, where traders overweight successful trades when learning about their ability can lead to overconfidence in a similar way as selective memory did in Section 4.2. It seems difficult to distinguish selective memory from underinference bias using data about beliefs alone, and none of the data we have found includes information on which histories the subjects recall (see Benjamin [2019] for a survey).

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 with selective memory as with uniform underinference, so their limit action will be the same. If the data-generating process is endogenous, random memory realizations can induce switches in actions, reducing the set of actions that can be long-run limits for a given memory function. The following example illustrates this possibility.

Example 8. There are no signals, $A = \{a', a''\}$, $Y = \{0, 1\}$, $u(\cdot, y) = y$, and the agent knows 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 a''. Their initial belief is that it is larger than that of action a', so $BR(\mu) = b$, although there is $p' \in \Theta$ with $p'_{a''}(1) < c$. The truth is that $1 > p^*_{a''}(1) > c$, so action a'' is optimal, but if m is strictly positive, both a' and a'' are selective memory equilibria. In the underinference model a' is a limit action for any such m, and if $a_t = a'$ then

[24] Frick, Iijima, and Ishii [2021] shows that uniform underinference leads to the same speed of belief convergence as correct updating in a setting with a fixed outcome distribution.

$a_\tau = a'$ almost surely for all $\tau > t$,. Instead, with the selective memory model a' is not a limit action because if a'' is played only a finite number of times, there is a positive probability of forgetting all such experiences and only using the prior to choose the action, which favors action a''. ▲

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 likelihood maximizer 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 now 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 and the agent remembers all past observations. It will be convenient here to make the dependence of the memory-weighted likelihood maximizers on Θ and m explicit, so let

$$ML^{\Theta,m}(\sigma) = \operatorname*{argmax}_{p \in \Theta} \left(\sum_{s \in S} \zeta(s) \sum_{y \in Y} p^*_{s,\sigma(s)}(y) \log p_{s,\sigma(s)}(y) \right) \tag{5}$$

be the set of memory-weighted likelihood maximizers for Θ and memory function m. The case studied in the misspecification literature has perfect memory, so there $m = \mathbf{1}$.

Definition 4. A strategy σ is a

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

2. *Uniformly strict Berk-Nash equilibrium* if for all $\nu \in \Delta(ML^{\Theta,\mathbf{1}}(\sigma))$ 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 [2021] shows that uniformly strict Berk-Nash equilibrium is a necessary condition for a strategy to have probability near 1 of being a long-run outcome.

There is a close relationship between the uniformly strict versions of Berk-Nash equilibrium and selective memory equilibrium: 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 equilibrium is equivalent to a Berk-Nash equilibrium for some support. To formalize this idea, we say that two equilibria are belief equivalent if they prescribe the same strategies, and the behavior after each signal can be justified by the same belief.

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

Proposition 6.

1. *Every uniformly strict Berk-Nash equilibrium with support Θ is belief equivalent to a selective memory equilibrium with support $\Theta^{'} = \Delta(Y)^{S\times A}$ 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 maximizer p with perfect memory but incomplete support, we can choose a memory function that rescales the probability of each (s, a, y) by some constant times $p_{s,a}(y)/p^*_{s,a}(y)$. This makes the recalled frequency equal to p, so p is a weighted-memory likelihood maximizer, and σ is the best reply.[25] 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 belief in the specified selective memory equilibrium.

Remark 2. As we prove in Online Appendix B.2, the uniform strictness conditions of Proposition 6 are needed:

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

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.

Moreover, selective memory equilibria need not be objectively optimal when the agent knows that the distribution of outcomes is independent of their action ($p^*_{s,a} = p^*_{s,a'}$ and $p_{s,a} = p_{s,a'}$ for every $p \in \Theta, a, a' \in A, s \in S$).

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 $x \in X \subseteq \mathbb{R}$ is the value for the seller, $v \in V \subseteq \mathbb{R}$ measures the gains from trade, and ε is a noise term, 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, Section 4.2 shows that the two models have different comparative statics with respect to changes in the environment.[27] Thus empirical work might be able to distinguish between the two models based on variations in information.

Persistence While agents undoubtedly are sometimes misspecified, some recent papers have theoretically questioned how likely these misperceptions are to persist and proposed various mechanisms by which agents might realize that some model not in the support of their initial beliefs better fits the data (Schwartzstein, 2014; Fudenberg and Lanzani, 2023; He and Libgober, 2023; Lanzani, 2023). In contrast, an agent with a selective memory and a full support prior will be able to explain their recollections

[26] The corresponding memory function is $m(a,(x,v)) = k\big[\sum_{v'\in V} p^*(x,v')\sum_{x'\in X} p^*(x',v)\big]/p^*(x,v)$ for sufficiently small k.

[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.

with one of their conjectured models and so have less reason to learn of their errors.

6 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 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.[28] 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.

Convergence to Equilibria Theorem 2 gives sufficient conditions for there to be a global attractor. Even when no such strategy exists, one could hope that there is probability 1 of converging to some limit strategy, with which strategy occurs depending both on the agent's prior and on the realized outcomes. We hope to find sufficient conditions for that in future work, along with (presumably weaker) conditions that ensure a positive probability of converging to a limit strategy.

[28] However, see Danz, Vesterlund, and Wilson [2022] for practical challenges in belief elicitation.

Partial naïveté We have assumed that agents treat the experiences they remember as if these were the only ones that happened. Appendix A.3 considers agents who are partially aware of their memory limitations. 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.[29] For an agent who is aware of their own forgetfulness but unaware 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, any action that is optimal for the true data-generating process is always a selective memory equilibrium.

Finite Memory In our model, the number of recalled experiences converges to infinity, as if the agent had perfect memory. In ongoing work, we modify the model to make the expected number of recalled periods bounded. Here the agent's beliefs need not converge to a deterministic limit even when the strategy is fixed, which can make the limit behavior stochastic. Thus instead of characterizing the possible limit strategies, we show that if the frequency with which strategies are used converges, the limit strategy distribution is generated by a best response to the distribution of memories it generates. We hope to use this to model the effect of "rehearsal," where an experience recalled in one period is more likely to be recalled again.

Other Possible Extensions It would be relatively easy to extend our analysis to agents who "misremember" and access false memories as opposed to simply forgetting things that happened. If what is misremembered depends only on the particular experience, this would be equivalent to selective memory with memory weights that reflect both the selective memory and the error function. A more substantive generalization would be from an agent who believes that outcomes are i.i.d. to an agent who believes that outcomes follow a Markov process. This would let us capture, the gambler's fallacy (see Rabin and Vayanos [2010] and He [2022]) if an outcome is more memorable when it is different than the outcome in the previous period.[30] Or it might be much easier for agents to recall whether an experience happened at all than whether it happened

[29] However, as Example 9 in the Appendix shows, less naïve agents can take worse actions and get lower payoffs.

[30] This extension could make use of the analysis of belief concentration for misspecified agents with Markov models developed in Fudenberg, Lanzani, and Strack [2022].

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

In the Appendix, we keep the dependence of the memory-weighted likelihood maximizers on Θ and m explicit, as in equation (5). For every $t \in \mathbb{N}$, we first explicitly describe the map

$$\begin{array}{ll} H_t \times 2^{\{1,\dots,t\}} & \to H \\ \left(h_t = (s_i, a_i, y_i)_{i=1}^{t}, R\right) & \mapsto h_t(R) \end{array}$$

that transforms an objective history and a set of recalled periods into the recalled history. Let $n(k, R) = \tau$ if τ is the k-th smallest number in R, i.e., $n(1, R) = \tau$ if $\tau = \min R$, $n(2, R) = \tau$ if $\tau = \min(R \backslash \{\min R\})$, so on and so forth. Then

$$h_t(R) = \left(s_{n(k,R)}, a_{n(k,R)}, y_{n(k,R)}\right)_{k=1}^{|R|}. \tag{6}$$

Combining equations (1) and (6) we have that the posterior probability of every measurable $C \subseteq \Theta$ after objective history h_t when the recalled periods are $\varnothing \neq R \subseteq \{1, ..., t\}$ is

$$\frac{\int_C \prod_{\tau \in R} p_{s_\tau, a_\tau}(y_\tau) d\mu(p)}{\int_\Theta \prod_{\tau \in R} p_{s_\tau, a_\tau}(y_\tau) d\mu(p)}. \tag{7}$$

We now state 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 $h_t = (s_i, a_i, y_i)_{i=1}^{t}$, and let

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

denote the recalled empirical distribution in objective history h_t when the recalled periods are $\varnothing \neq R \subseteq \{1, .., t\}$. Also, for every $\gamma \in \Delta(S \times A \times Y)$ and $p \in \Delta(Y)^{S \times A}$ let

$D(\gamma||p) = -\sum_{(s,a,y)} \gamma(s,a,y) \log(p_{s,a}(y))$.

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

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

Let $ML_s^{\Theta,m}(\sigma,\varepsilon) = \{p \in \Theta : \exists q \in ML_s^{\Theta,m}(\sigma), ||p-q||_\infty \leqslant \varepsilon\}$. The next lemma shows that posteriors that assign a sufficiently high probability to balls around the maximizers induce the expected behavior: actions that are not selective equilibria are not best replies, and make 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 for all $\nu \in \Delta(\Theta)$,*

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

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

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

For any $t \in \mathbb{N}$, $\sigma \in A^S$ and $a^t \in A^t$ let $\pi^{\sigma,a^t} \in A^H$ be the policy that prescribes action a_τ at period $\tau \leqslant t$ and action $\sigma(s_\tau)$ at all periods $\tau > t$ and let $\mathbb{P}_{\sigma,a^t}$ be the probability distribution induced by π^{σ,a^t}.

The next lemma shows that if σ is a limit strategy, then for some 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 afterward, there is positive probability that the induced sequence of beliefs makes σ optimal at all periods $\tau \geqslant t+1$.

Lemma A.3. *Let $\sigma \in A^S$. If for every $t \in \mathbb{N}$, every $a^t \in A^t$, and every optimal policy $\tilde{\pi}$, $\mathbb{P}_{\sigma,a^t}[\sigma(s_{\tau+1}) = \tilde{\pi}(h_\tau(R))(s_{\tau+1})$ for all $\tau \geqslant t] = 0$ then σ is not a limit strategy.*

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

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

Throughout the Appendix, we let $\mathbf{R}_t$ denote the random variable corresponding to the subset of periods recalled at period t.

A.2 Proof of Theorem 1

Proof. Suppose towards a contradiction that σ is a limit strategy under the optimal policy π, but not a selective memory equilibrium. By Lemma A.2 there are $s' \in S$ and $\varepsilon, C \in \mathbb{R}_{++}$ such that for all $\nu \in \Delta(\Theta)$

$$\frac{\nu(ML_{s'}^{\Theta,m}(\sigma,\varepsilon))}{1-\nu(ML_{s'}^{\Theta,m}(\sigma,\varepsilon))} > C \implies \sigma(s') \notin BR(s',\nu). \tag{8}$$

If $\sigma(s') \notin BR(s',\nu)$ for all $\nu \in \Delta(\Theta)$, we immediately reach a contradiction by definition of optimal policy, since by Kolmogorov $0-1$ Law (see, e.g., Theorem 8.4.4 in Dudley [2018]) signal s' will realize infinitely many times $\mathbb{P}_\pi$-a.s. If not, since $\sigma(s') \notin BR(s', \Delta(ML_{s'}^{\Theta,m}))$ by equation (8), we have $ML_{s'}^{\Theta,m}(\sigma) \neq \Theta$, so there is $(\underline{s}, \underline{a}, \underline{y}) \in S \times A \times Y$ with $\underline{a} = \sigma(\underline{s})$, $p^*_{\underline{s},\underline{a}}(\underline{y}) > 0$, and $m_{s'}(\underline{s}, \underline{a}, \underline{y}) =: \ell > 0$.

Let $h_t = (s^t, a^t, y^t) \in H_t$ be an objective history that has positive probability under an optimal policy π, i.e., $\mathbb{P}_\pi[h_t] > 0$. We show that if the agent plays σ at every period after h_t, $\mathbb{P}_{\sigma,a^t}$ almost surely the belief $\mu_\tau(\cdot|h_\tau(\mathbf{R}_\tau))$ reaches a region where no optimal policy prescribes $\sigma(s')$ after signal s', i.e., $\sigma(s') \notin BR(s', \mu_\tau(\cdot|h_\tau(\mathbf{R}_\tau)))$. By Lemma *A.3*, this is enough to obtain the desired conclusion.

By the strong law of large numbers, for every $(s,a,y) \in S \times A \times Y$

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

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

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

For every two periods $\tau' > \tau$ and $R' \subseteq \{1, ..., \tau'\}$, the probability of recalling R' at time τ' conditional on the objective history $h_{\tau'}$ is independent of the recalled periods R at period τ, i.e., $\mathbb{P}_{\sigma,a^t}[R'|h_{\tau'}] = \mathbb{P}_{\sigma,a^t}[R'|h_{\tau'}, R]$. The next claim shows that for every $k \in \mathbb{N}$, $\mathbb{P}_{\sigma,a^t}$ almost surely there is a $\tau > t$ such that $s_\tau = s'$, and the number of periods recalled at period τ is larger than k.

Claim 1. *For all $\hat{\tau} \in \mathbb{N}$ and $k \in \mathbb{N}$, $\mathbb{P}_{\sigma,a^t}\left[|\mathbf{R}_\tau| \leqslant k, \forall \tau \geqslant \hat{\tau}\right] = 0$.*

Proof of the Claim. For every $\tau \in \mathbb{N}$ and $h = (s_i, a_i, y_i)_{i=1}^{\tau}$ let $N(h) = \sum_{i=1}^{\tau-1} \mathbb{1}_{\left(\underline{s},\underline{a},\underline{y}\right)}(s_i, a_i, y_i)$ be number of times $(\underline{s}, \underline{a}, \underline{y})$ occurs between period 1 and τ. For any $j \in \mathbb{N}$, we have

$$
\begin{aligned}
& \mathbb{P}_{\sigma,a^t}\left[|\mathbf{R}_\tau| \leqslant k, \forall \tau \in \{\hat{\tau}, \hat{\tau}+j\}\right] \\
= & \prod_{\tau=\hat{\tau}}^{\hat{\tau}+j} \sum_{h \in H_{\tau-1}} \mathbb{P}_{\sigma,a^t}(h)\left(1 - \mathbb{P}_{\sigma,a^t}\left[|\mathbf{R}_\tau| > k \middle| \left(|\mathbf{R}_{\hat{\tau}}| \leqslant k, ..., |\mathbf{R}_{\tau-1}| \leqslant k, h\right)\right]\right) \\
= & \prod_{\tau=\hat{\tau}}^{\hat{\tau}+j} \sum_{h \in H_{\tau-1}} \mathbb{P}_{\sigma,a^t}(h)\left(1 - \mathbb{P}_{\sigma,a^t}\left[|\mathbf{R}_\tau| > k | h\right]\right) \\
\leqslant & \prod_{\tau=\hat{\tau}}^{\hat{\tau}+j} \left(\mathbb{P}_{\sigma,a^t}\left(\{h \in H_{\tau-1} : N(h) \leqslant k\}\right) + \textstyle\sum_{h \in H_{\tau-1}: N(h) \geqslant k+1} \mathbb{P}_{\sigma,a^t}(h)\left(1 - \mathbb{P}_{\sigma,a^t}\left[|\mathbf{R}_\tau| > k | h\right]\right) \right) \\
\leqslant & \prod_{\tau=\hat{\tau}}^{\hat{\tau}+j} \left(\mathbb{P}_{\sigma,a^t}\left(\{h \in H_{\tau-1} : N(h) \leqslant k\}\right) + \textstyle\sum_{h \in H_{\tau-1}: N(h) \geqslant k+1} \mathbb{P}_{\sigma,a^t}(h)(1-\ell) \right) \\
= & \prod_{\tau=\hat{\tau}}^{\hat{\tau}+j} \left(\mathbb{P}_{\sigma,a^t}\left(\{h \in H_{\tau-1} : N(h) \leqslant k\}\right) + \left(1 - \mathbb{P}_{\sigma,a^t}\left(\{h \in H_{\tau-1} : N(h) \leqslant k\}\right)\right)(1-\ell) \right) \\
= & \prod_{\tau=\hat{\tau}}^{\hat{\tau}+j} \left(1 - \ell + \mathbb{P}_{\sigma,a^t}\left(\{h \in H_{\tau-1} : N(h) \leqslant k\}\right)\ell\right),
\end{aligned}
$$

where the first equality follows from the law of iterated expectations. Since $1 - x \leqslant e^{-x}$ for all $x \in \mathbb{R}$, the last term is smaller than

$$
\exp\left(\sum_{\tau=\hat{\tau}}^{\hat{\tau}+j} -\ell + \mathbb{P}_{\sigma,a^t}\left(\{h \in H_{\tau-1} : N(h) \leqslant k\}\right)\ell\right).
$$

For sufficiently large $\hat{\tau}$, and every $\tau \geqslant \hat{\tau}$, $\mathbb{P}_{\sigma,a^t}\left(\{h \in H_{\tau-1} : N(h) \leqslant k\}\right) < 1$ so that

$$
\lim_{j\to\infty} \mathbb{P}_{\sigma,a^t}\left[|\mathbf{R}_\tau| \leqslant k, \forall \tau \in \{\hat{\tau}, \hat{\tau}+j\}\right] \leqslant \lim_{j\to\infty} \exp\left(\sum_{\tau=\hat{\tau}}^{\hat{\tau}+j} -\ell + \mathbb{P}_{\sigma,a^t}\left(\{h \in H_{\tau-1} : N(h) \leqslant k\}\right)\ell\right) = 0
$$

proving the claim. □

By the previous claim, for every $\varepsilon \in \mathbb{R}_{++}$, $s' \in S$, and $k \in \mathbb{N}_{++}$, $\mathbb{P}_{\sigma,a^t}$ almost surely there is a $\tau > t$ such that $s_\tau = s'$, the number of the recalled periods R is larger than k. Moreover, by the SLLN $||\hat{f}(h_\tau, R) - \tilde{p}(\sigma, s')||_\infty < \varepsilon$, $\mathbb{P}_{\sigma,a^t}$ almost surely. We show that eventually $\frac{\nu(ML_{s'}^{\Theta,m}(\sigma,\varepsilon))}{1-\nu(ML_{s'}^{\Theta,m}(\sigma,\varepsilon))} > C$ on the histories where these two conditions are satisfied. Since they hold $\mathbb{P}_{\sigma,a^t}$ almost surely, the result follows.

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

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

and

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

So, by Lemma A.1

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

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

$$\begin{aligned}
&\lim_{\tau\to\infty} \inf_{p \notin ML_{s'}^{\Theta,m}(\sigma,\varepsilon)} - \sum_{(s,a,y)} \hat{f}(h_\tau,R)(s,a,y)\log(p_{s,a}(y)) - \sup_{p \in ML_{s'}^{\Theta,m}(\sigma,\varepsilon')} - \sum_{(s,a,y)} \hat{f}(h_\tau,R)(s,a,y)\log(p_{s,a}(y)) \\
&= - \sup_{p \in ML_{s'}^{\Theta,m}(\sigma,\varepsilon')} - \sum_{s\in S}\zeta(s)\sum_{y\in Y} m_{s'}(s,\sigma(s),y)p^*_{s,\sigma(s)}(y)\log(p_{s,\sigma(s)}(y)) \\
&\qquad + \inf_{p \notin ML_{s'}^{\Theta,m}(\sigma,\varepsilon)} - \sum_{s\in S}\zeta(s)\sum_{y\in Y} m_{s'}(s,\sigma(s),y)p^*_{s,\sigma(s)}(y)\log(p_{s,\sigma(s)}(y)) > \frac{\kappa}{2} > 0\,. \qquad \square
\end{aligned}$$

Proof of Theorem 2. First, we introduce some notation. Let

$$\underline{m}: \;= \min_{s',s,a \in S^2 \times A} \sum_{y \in Y} p^*_{s,a}(y) m_{s'}(s,a,y) > 0.$$

and

$$\bar{p}_{s,a}(y) := \frac{m(s,a,y)p^*_{s,a}(y)}{\sum_{y'} m(s,a,y')p^*_{s,a}(y')};$$

the latter is the memory-adjusted version of the data-generating process. Fix an arbitrary $\underline{y}$. Let $n_{s,a,t}$ the number of times the signal-action pair $(s,a) \in S \times A$ occurred in periods $\{1, ..., t\}$ and $g_{s,a,t}$ be the frequency of outcomes induced by signal s and action a until period t, i.e.,

$$g_{s,a,t}(y) = \frac{\sum_{i=1}^{t} \mathbb{1}_{\{(s,a,y)\}}(s_i, a_i, y_i)}{n_{s,a,t}}$$

with $g_{s,a,t}(y) = \mathbb{1}_{\{\underline{y}\}}(y)$ whenever $n_{s,a,t} = 0$. Similarly, let $\tilde{n}_{s,a,t}$ be the number of times the signal-action pair (s, a) is recalled at period t and $f_{s,a,t}$ be the frequency of outcomes induced by signal s and action a that is recalled at period t, with $f_{s,a,t}(y) = \mathbb{1}_{\{\underline{y}\}}(y)$ whenever $\tilde{n}_{s,a,t} = 0$.

Now we will prove the first part of the theorem, namely that

$$\mathbb{P}_\pi \left[\lim_{t\to\infty} \mu(B_\epsilon(\hat{p})|h_t(R_t)) = 1\right] = 1.$$

The first step is to characterize the distribution of outcomes given the realized signals and actions.

Consider the stochastic processes $(\mathbf{X}_t^{(s,a,y)})_{(s,a,y)\in S\times A\times Y, t\in\mathbb{N}}$ defined by

$$\mathbf{X}_t^{(s,a,y)} = (\mathbb{1}_{\{y\}}(y_t) - p^*_{s,a}(y))\mathbb{1}_{\{(s,a)\}}(s_t, a_t) \quad \forall t \in \mathbb{N}.$$

These stochastic processes correspond to the deviation of the number of times each y has appeared from their expected frequency given the signal realized and actions chosen. They are measurable with respect to the filtration $(\mathcal{F}_t)_{t\in\mathbb{N}}$ generated by the stochastic process of histories $(h_t)_{t\in\mathbb{N}}$. These processes are not i.i.d., as previous outcome realizations affect current period choices, but for each $(s, a, y) \in S \times A \times Y$, $\left(\mathbf{X}_t^{(s,a,y)}\right)_{t\in\mathbb{N}}$ is a martingale difference sequence. A fortiori, $\left(\mathbf{X}_t^{(s,a,y)}\right)_{t\in\mathbb{N}}$ is a mixingale difference sequence, and by the strong law of large numbers for mixingale sequences (see Theorem 2.7 in Hall and Heyde, 2014 for the version that applies here), we have

$$\lim_{n\to\infty} \frac{\sum_{t=1}^{n} \mathbf{X}_t^{(s,a,y)}}{n} = 0 \qquad \mathbb{P}_\pi\text{-a.s.}$$

This in turn implies that for every $\varepsilon > 0$, $\gamma > 0$, $\mathbb{P}_\pi$ almost surely

$$\limsup_{t\to\infty} \mathbb{1}_{[\gamma,\infty)}(n_{(s,a,t)}/t)\mathbb{1}_{[\epsilon,\infty)}(||g_{s,a,t} - p^*_{s,a}||_\infty) = 0. \tag{9}$$

Let $t \in \mathbb{N}$, $(h_t, s') \in H_t \times S$ and $\varepsilon > 0$. We next bound the distance between the recalled empirical frequency and the memory-distorted version of the realized empirical frequency. Indeed, by the Chernoff inequality (see, e.g., page 24 of Boucheron, Lugosi, and Massart [2013]), for every $(s, a, y) \in S \times A \times Y$

$$\mathbb{P}_\pi \left[\left| \frac{\tilde{n}_{s,a,t} f_{s,a,t}(y)}{n_{s,a,t}} - m_{s'}(s, a, y)\, g_{s,a,t}(y) \right| > \varepsilon \mid (h_t, s') \right] \leqslant 2 \exp^{-\varepsilon n_{s,a,t}[\log 1/2 - \frac{\log(1/2+\varepsilon)+\log(1/2-\varepsilon)}{2}]}$$

and

$$\mathbb{P}_\pi \left[\underline{m} - \frac{\tilde{n}_{s,a,t}}{n_{s,a,t}} > \varepsilon \mid (h_t, s') \right] \leqslant 2 \exp^{-\varepsilon n_{s,a,t}[\log 1/2 - \frac{\log(1/2+\varepsilon)+\log(1/2-\varepsilon)}{2}]}.$$

Now we combine these upper bounds with the Borel-Cantelli lemma to show that for any signal-action pair (s, a) that occurs a nonvanishing fraction of time, there are only finitely many periods where either only a small fraction of recalled histories have recalled signal-action pair (s, a) or the recalled frequency is a large deviation in the sense of the last display. Since

$$\sum_{i=1}^{\infty} 2 \exp^{-\varepsilon i[\log 1/2 - \frac{\log(1/2+\varepsilon)+\log(1/2-\varepsilon)}{2}]} < \infty,$$

by the Borel-Cantelli lemma,

$$\mathbb{P}_\pi \left[\left| \begin{array}{c} t : \exists\, (s, a, s') \text{ with } \frac{n_{s,a,t}}{t} > \gamma \text{ and} \\ \left| \frac{\tilde{n}_{s,a,t} f_{s,a,t}(y)}{n_{s,a,t}} - m_{s'}(s, a, y)\, g_{s,a,t}(y) \right| > \varepsilon \text{ or } \frac{\tilde{n}_{s,a,t}}{t} < \frac{\gamma}{2} \underline{m} \end{array} \right| = \infty \right] = 0. \tag{10}$$

Observe then that under data generating process $q \in \Theta$, the log-likelihood of any history $(s_i, a_i, y_i)_{i=1}^{\tau}$, $\tau \leqslant t$ that the agent might recall at time t can be rewritten as

$$\begin{aligned} \sum_{i=1}^{\tau} \log q_{s_i,a_i}(y_i) &= \sum_{(s,a)\in S\times A} \tilde{n}_{s,a,t} \sum_{y\in Y} f_{s,a,t}(y) \log q_{s,a}(y) \\ &= \sum_{(s,a)\in S\times A} \tilde{n}_{s,a,t} \left(-D_{KL}(f_{s,a,t}, q_{s,a}) + \sum_{y\in Y} f_{s,a,t}(y) \log f_{s,a,t}(y) \right). \end{aligned}$$

This implies that for every $\epsilon > 0$,

$$\frac{\mu(B_\epsilon(\hat{p}) | (s_i, a_i, y_i)_{i=1}^{\tau})}{1 - \mu(B_\epsilon(\hat{p}) | (s_i, a_i, y_i)_{i=1}^{\tau})} = \frac{\int_{B_\epsilon(\hat{p})} \exp\left(-\sum_{(s,a)\in S\times A} \tilde{n}_{s,a,t} D_{KL}(f_{s,a,t}, p_{s,a}) \right) d\mu(p)}{\int_{\Theta\setminus B_\epsilon(\hat{p})} \exp\left(-\sum_{(s,a)\in S\times A} \tilde{n}_{s,a,t} D_{KL}(f_{s,a,t}, q_{s,a}) \right) d\mu(q)}.$$

By Assumption 2, $\hat{p}$ maximizes the log-likelihood and hence minimizes the divergence from $\bar{p}$ after every signal action pair. Thus, because D_{KL} is lower semicontinuous, there is $\epsilon > 0$ such that for all $(s,a) \in S \times A$ and $q \in \Theta \backslash B_\epsilon(\hat{p})$, $D_{KL}(\bar{p}_{s,a}, q_{s,a}) > \epsilon + D_{KL}(\bar{p}_{s,a}, \hat{p}_{s,a})$.

By equations (9) and (10), $\mathbb{P}_\pi$ almost surely

$$\liminf_{t\to\infty}(1/t) \sum_{(s,a)\in S\times A} \tilde{n}_{s,a,t} D_{KL}(f_{s,a,t}, q_{s,a}) \geqslant \liminf_{t\to\infty} 2\epsilon/t \sum_{(s,a)\in S\times A} \tilde{n}_{s,a,t} \tag{11}$$

for every $q \in \Theta\backslash B_\epsilon(\hat{p})$.

Conversely, by Lemma 3 of Fudenberg, Lanzani, and Strack [2022], for all $p \in \Theta$

$$D_{KL}(f_{s,a,t}, p_{s,a}) \leqslant 2 \left(\max_y \frac{f_{s,a,t}(y)}{p_{s,a}(y)} \right) ||f_{s,a,t} - p_{s,a}||_{TV}.$$

Therefore by Assumption 1, we can choose ϵ' small enough that $||f_{s,a,t} - \bar{p}_{s,a}||_{TV} \leqslant \epsilon'$, and $p \in B_{\varepsilon'}(\hat{p})$

$$D_{KL}(f_{s,a,t}, \hat{p}_{s,a}) \leqslant \frac{\epsilon}{2}$$

almost surely. Let $K = \max_{s\in S, a\in A, p\in B_{\epsilon'}(\hat{p}), f\in\Delta(Y):\text{supp} f\subseteq p_{(s,a)}} D_{KL}(f, \hat{p}_{s,a})$.

By equation (10), we get that for all $\beta \in (0,1)$, $\mathbb{P}_\pi$-almost surely

$$\begin{aligned}
\lim_{t\to\infty} \frac{\mu(B_{\varepsilon'}(\hat{p})|h_t(\mathbf{R}_t))}{1-\mu(B_{\varepsilon'}(\hat{p})|h_t(\mathbf{R}_t))} &= \lim_{t\to\infty} \frac{\int_{B_{\varepsilon'}(\hat{p})} \exp\left(-\sum_{s\in S, a\in A} \tilde{n}_{s,a,t} D_{KL}(f_{s,a,t}, p_{s,a})\right) d\mu(p)}{\int_{\Theta\backslash B_{\varepsilon'}(\hat{p})} \exp\left(-\sum_{s,a\in A} \tilde{n}_{s,a,t} D_{KL}(f_{s,a,t}, q_{s,a})\right) d\mu(q)} \\
&\geqslant \lim_{t\to\infty} \frac{\int_{B_{\varepsilon'}(\hat{p})} \exp\left(-t\beta K - \sum_{s\in S, a\in A: \frac{\tilde{n}_{s,a,t}}{t} > \frac{\beta}{|S\times A|}} \tilde{n}_{s,a,t} D_{KL}(f_{s,a,t}, p_{s,a})\right) d\mu(p)}{\int_{\Theta_{\varepsilon'}(\hat{p})} \exp\left(-\sum_{s\in S, a\in A: \frac{\tilde{n}_{s,a,t}}{t} > \frac{\beta}{|S\times A|}} \tilde{n}_{s,a,t} D_{KL}(f_{s,a,t}, q_{s,a})\right) d\mu(q)} \\
&\geqslant \lim_{t\to\infty} \frac{\mu(B_{\varepsilon'}(\hat{p}))}{1-\mu(B_{\varepsilon'}(\hat{p}))} \exp\left(t\left[-\beta K + \frac{m}{2}(1-\beta)\left(\epsilon - \frac{\epsilon}{2}\right)\right]\right)
\end{aligned}$$

where the last inequality follows by equation (A.2). For β small enough that $-\beta K + \frac{m}{2}(1-\beta)\left(\epsilon - \frac{\epsilon}{2}\right) > 0$, the right-hand side goes to infinity as t goes to infinity, so the left-hand side must also diverge, which shows that $\mathbb{P}_\pi\big[(] \lim_{t\to\infty} \mu(B_{\varepsilon'}(\hat{p})|h_t(\mathbf{R}_t)) = 1\big] = 1$. In particular, the random variable T defined as

$$T := \sup\{t \in \mathbb{N} : \mu(B_{\varepsilon'}(\hat{p})|h_t(\mathbf{R}_t)) < 1 - \varepsilon'\} \tag{12}$$

is $\mathbb{P}_\pi$-almost surely finite.

To prove the second part of the theorem, note that because there is a unique best response to $\hat{p}$ for every signal s, $\hat{\sigma}$ is a uniformly strict selective memory equilibrium. By Lemma A.2, there is an ϵ such that $\hat{\sigma}(s)$ is the response to s for any belief ν that assigns probability at least $1-\epsilon$ to $B_\epsilon(\hat{p})$. Since by equation (12) $\mathbb{P}_\pi$-almost surely there will be a finite time T (that can depend on the sample path) with $\mu(B_\epsilon(\hat{p})|h_t(\mathbf{R}_t)) > 1-\epsilon$ for all $t > T$, the result follows. □

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

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

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

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

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

proving that $p \notin ML^{\Theta,m}_{\hat{s}}(\sigma)$. □

Proof of Proposition 2. From Theorem 2, we know that beliefs converge. 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, this long-run belief is unique and independent of a, so we suppress the dependence of p and $\tilde{m}$ on a.

Because $\Theta = \Delta(Y)$, for every σ the unique memory-weighted likelihood maximizers is the distribution

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

and by Lemma A.1 the beliefs concentrate on $p^{\tilde{m}}$. Moreover $p^{m'}(y) = w(y)p^m(y)$, 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 $z \mapsto \sum_{x \leqslant z}(p^{m'}(x) - p^m(x)) = \sum_{x \leqslant z} p^m(x)(w(x)-1)$ is quasi-convex. It equals 0 for $z < \min_{y \in Y} y$ and for $z \geqslant$

$\max_{y \in Y} y$, so it is non-positive for $z \in [\min_{y \in Y} y, \max_{y \in Y} y]$, and $p^{m'}$ 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 weakly higher under m' than under m. □

Proof of Proposition 3. From Theorem 2, we know that beliefs converge. Because (y_1, y_2) are subjectively independent conditional on the value of p, the learning problem decouples across the two dimensions. By Proposition 2, the long-run belief about p is weakly higher than the true distribution p^*. The probability with which an outcome is remembered is independent of the second component, so the agent learns $\alpha p^*(y_2) + (1-\alpha) q^*(y_2)$. They infer q to be

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

Thus $q - q^* \equiv \frac{\alpha}{1-\alpha}(p^* - p)$, 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. □

Proof of Proposition 4. If $a = 1$ is not a self-confirming equilibrium, 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 likelihood maximizer $\hat{p}$ under action 1 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 $z = \bar{y}$. And since

v is concave, Theorem 3 and Footnote 19 of Machina and Pratt [1997] implies that

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

and the risky action cannot be a selective memory equilibrium. $\square$

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 $\frac{\nu(ML_{s'}^{\Theta,m}(\sigma,c))}{1-\nu(ML_{s'}^{\Theta,m}(\sigma,c))} > C$ then $\sigma(s') \notin BR(s', \nu)$. Let $h_t = (s^t, a^t, y^t)$ be a history with positive probability under π. We show that if the agent plays the strategy π^{σ,a^t}, then almost surely the underinference belief $\mu^U(\cdot|(s^\tau, a^\tau, y^\tau))$ is asymptotically in a region where no optimal policy prescribes σ after signal s'. Since almost surely signal s' occurs infinitely many times, by Lemma *A.3* this is enough to obtain the desired conclusion.

Under strategy π^{σ,a^t}, by the Strong Law of Large Numbers,

$$\lim_{\tau \to \infty} f(h_\tau)(s, a, y) = \begin{cases} \zeta(s) p_{s,a}^*(y) & \text{if } a = \sigma(s) \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

almost surely. Now we express the posterior as a function of the observed frequencies, and show that it concentrates on the memory-weighted likelihood maximizers, so that the result follows from the upper hemicontinuity of BR.

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

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

and

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

By equation (13) and the definition of κ and c', almost surely on the cylinder h_t we have

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

By Lemma A.1,

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

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

Proof of Proposition 6. To prove part (1), let σ be a uniformly strict Berk-Nash equilibrium, and let p' be an arbitrary element of $ML^{\Theta,\mathbf{1}}$. Since σ is a uniformly strict Berk-Nash equilibrium, for all $s \in S$, $\{\sigma(s)\} = BR(s,\delta_{p'})$. Moreover, by Assumption (1), $p^*_{s,\sigma(s)}(y) = 0$ implies $p'_{s,\sigma(s)}(y) = 0$, so[31] $K := \max_{(s,a,y)\in S\times A\times Y} \frac{p'_{s,a}(y)}{p^*_{s,a}(y)} < \infty$. Define $\tilde{m}$ by $\tilde{m}_{s'}(s,a,y) = \frac{p'_{s,a}(y)}{Kp^*_{s,a}(y)}$. Then for an agent with a full-support prior and memory function $\tilde{m}$ the memory-weighted likelihood maximizers for strategy σ after signal s' are the elements of

$$
\begin{aligned}
&\operatorname*{argmax}_{p\in\Delta(Y)^{S\times A}} \sum_{s\in S}\zeta(s)\sum_{y\in Y}\tilde{m}_{s'}(s,\sigma(s),y)p^*_{s,\sigma(s)}(y)\log p_{s,\sigma(s)}(y) \\
&= \operatorname*{argmax}_{p\in\Delta(Y)^{S\times A}} \sum_{s\in S}\zeta(s)\sum_{y\in Y}\frac{p'_{s,\sigma(s)}(y)}{K}\log p_{s,\sigma(s)}(y) = \operatorname*{argmax}_{p\in\Delta(Y)^{S\times A}} \sum_{s\in S}\zeta(s)\sum_{y\in Y}p'_{s,\sigma(s)}(y)\log p_{s,\sigma(s)}(y).
\end{aligned}
$$

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

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

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

A.3 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 at least 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 actions or believes they convey no information. 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 $s \in S$ and outcome $y \in Y$ with probability $\hat{m}(s, y)$, instead of the true probability $m(s, y)$. The subjective likelihood of recalling the periods R after (h_t, s') under data-generating process p is then proportional to

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

Thus the subjective log-likelihood equals

$$(t - |R|) \log \left[\sum_{s \in S} \zeta(s) \sum_{z \in Y} p_s(z)(1 - \hat{m}_{s'}(s, z))\right] + \sum_{y \in Y, s \in S, \tau \in R} \mathbb{1}_{s,y}(s_\tau, y_\tau) \log(p_s(y) \hat{m}_{s'}(s, y)) \tag{14}$$

where $|R|$ is the number of events the agent remembers. (Note that the first term does not appear when the agent believes they remember everything.)

Because the expected value of $|R|/t$ is $1 - \sum_{y \in Y} \sum_{s \in S} \zeta(s) p_s^*(y) m_{s'}(s, y)$, (14) suggests the following generalization of the definition of ML:

$$\begin{aligned} ML_{s'}^{\Theta,m,\hat{m}}(\sigma) = \underset{p \in \Theta}{\operatorname{argmax}} & \left(1 - \sum_{s \in S} \sum_{y \in Y} m_{s'}(s, y) \zeta(s) p_s^*(y)\right) \log \left(1 - \sum_{s \in S} \sum_{y \in Y} \zeta(s) p_s(y) \hat{m}_{s'}(s, y)\right) \\ & + \sum_{y \in Y} \sum_{s \in S} m_{s'}(s, y) \zeta(s) p_s^*(s, y) \log\left(\hat{m}_{s'}(s, y) p_s(y)\right) . \end{aligned}$$

Definition 6. A *selective memory equilibrium for a partially naïve agent* is a strategy σ such that for every $s \in S$ there exists a belief $\nu \in ML_s^{\Theta,m,\hat{m}}(\sigma)$ with $\sigma(s) \in BR(s, \nu)$.

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, $ML_s^{\Theta,m,\hat{m}} = ML_s^{\Theta,m}$ and the selective memory equilibria of a partially and fully naïve agent coincide.[32] This shows that what matters for our results is that the agent is unaware that their memory is selective, not that they are unaware of their forgetfulness. At the other extreme, if agents are fully aware of their memory function then $\delta_{p^*} \in ML_{s'}^{\Theta,m,m}$ so any action that is optimal for the true data-generating process is always a selective memory equilibrium.[33]

The next result, whose proof is omitted, follows by observing that for a partially naïve agent the posterior probability of C after an objective history (h_t, s_{t+1}) when the recalled periods are R is

$$\frac{\int_C \prod_{\tau\in R} \hat{m}_{s_{t+1}}(s_\tau, y_\tau) p_{s_\tau}(y_\tau) \Big(1 - \sum_{s\in S}\sum_{y\in Y} \zeta(s) p_s(y) \hat{m}_{s_{t+1}}(s,y)\Big)^{t-|R|} d\mu(p)}{\int_\Theta \prod_{\tau\in R} \hat{m}_{s_{t+1}}(s_\tau, y_\tau) p_{s_\tau}(y_\tau) \Big(1 - \sum_{s\in S}\sum_{y\in Y} \zeta(s) p_s(y) \hat{m}_{s_{t+1}}(s,y)\Big)^{t-|R|} d\mu(p)},$$

and then using an argument analogous to the proof of Theorem 1.

Proposition 7. *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 memory limitations. For $\alpha = 1$ the agent is sophisticated and understands their memory limitations, and so has 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.

[32] This is true in particular when the agent is fully naïve and $\hat{m}$ is identically 1, even though the maximand becomes ill-defined. To see why, note that when $\hat{m}_{s'}(\cdot) = k_{s'}$ for some constants $k_{s'}$, the maximand is

$$\sum_{y\in Y}\sum_{s\in S} k_{s'}\zeta(s)p_s^*(y)\log(p_s(y)) + \left[\sum_{s\in S}\sum_{y\in Y}\zeta(s)p_s^*(y)k_{s'}\log(\hat{m}_{s'}(s,y))\right] + \left[\sum_{s\in S}\zeta(s)\sum_{y\in Y}p_s^*(y)(1-m_{s'}(s,y))\right]\log(1-\hat{m}_{s'}(s,y))$$

The last two terms do not depend on p so $ML_{s'}^{\Theta,m,\hat{m}} = ML_{s'}^{\Theta,m}$, and in particular complete naivete is reached as the limit where all $k_{s'} \to 1$.

[33] More generally, if the agents are fully aware then selective memory equilibrium reduces to self-confirming equilibrium.

Proposition 8. *Suppose m and p^* are constant in a and m is increasing in y, that $\hat{m}(y) = (1-\alpha) + \alpha m(y)$, $\Theta = \Delta(Y)$, and u 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.*

The next example shows that the amount of naïveté can have a non-monotonic effect when there are more than two actions.

Example 9. *Suppose that the agent has three alternatives. They can either "do nothing," $a = n$ with certain payoff of 0, do a quick job $a = k$ with payoff 1 if the job succeeds and -1 otherwise, or do a careful and time-consuming costly job $a = h$ at cost 0.6 that yields $1 - .6 = .4$ if the project succeeds and -1.6 otherwise. The probability of success in the quick job is some unknown $p \in [0,1]$, while the probability of success in the careful job is $\max\{1, 2p\}$. The agent's prior assigns positive probability to all $p \in [0,1]$, where p is a reflection of the agent's ability.*

The true probability $p^* = 0.2$, so $\mathbb{E}_{p*}[u(n,\cdot)] > \mathbb{E}_{p*}[u(k,\cdot)] > \mathbb{E}_{p*}[u(h,\cdot)]$. Suppose that the agent has ego-boosting bias, in that they recall successes and they recall failures with probability 0.03.

In this case, welfare is nonmonotone in the amount of partial naïveté of the agent: the worst action h is never a selective memory equilibrium for a fully sophisticated agent or for a fully naive one. For the former, the unique selective equilibrium is the objectively optimal n, while the latter has two selective equilibria, n and k, with the latter sustained by the incorrect belief that their ability is so high that k is better than h. However, if the agent believes that they recall the failures with probability 0.12, playing the worst action h is a selective memory equilibrium, because the agent ends up believing that the probability of success is 0.5, which makes h the unique best reply.

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. (1987). "Comportamento razionale ed equilibrio nei giochi e nelle situazioni sociali." *Unpublished undergraduate dissertation, Bocconi University, Milano.*

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.

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.

Benoıt, J.-P. and J. Dubra (2011). "Apparent overconfidence." *Econometrica*, 79, 1591–1625.

Benoıt, J.-P., J. Dubra, and D. A. Moore (2015). "Does the better-than-average effect show that people are overconfident?: Two experiments." *Journal of the European Economic Association*, 13, 293–329.

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

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

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

Boucheron, S., G. Lugosi, and P. Massart (2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence. Univ. Press.*

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.

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

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.

Compte, O. and A. Postlewaite (2004). "Confidence-enhanced performance." *American Economic Review*, 94, 1536–1557.

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.

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

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

DeBondt, W. F. and R. H. Thaler (1995). *Handbooks in Operations Research and Management Science, Volume 9, Chapter Financial Decision-Making in Markets and Firms: A Behavioral Perspective*.

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 (2022). "Associative memory and belief formation." *NBER WP 26664*.

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., E. Vespa, and S. Yuksel (2023). *Mental models and learning: The case of base-rate neglect*. Tech. rep. Tech. rep.

Frick, M., R. Iijima, and Y. Ishii (2021). "Welfare comparisons for biased learning."

Fudenberg, D. and G. Lanzani (2023). "Which misperceptions persist?" *Theoretical Economics, Forthcoming*.

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

— (2022). "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.

Gervais, S. and T. Odean (2001). "Learning to be overconfident." *The review of financial studies*, 14, 1–27.

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.

Hall, P. and C. C. Heyde (2014). *Martingale limit theory and its application*. Academic press.

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.

He, K. and J. Libgober (2023). "Evolutionarily Stable (Mis)specifications: Theory and Application."

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

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

Lanzani, G. (2022). "Correlation made simple: Applications to salience and regret theory." *The Quarterly Journal of Economics*, 137, 959–987.

— (2023). "Dynamic Concern for Misspecification."

Loomes, G. and R. Sugden (1982). "Regret theory: An alternative theory of rational choice under uncertainty." *The economic journal*, 92, 805–824.

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.

Machina, M. and J. Pratt (1997). "Increasing risk: some direct constructions." *Journal of Risk and Uncertainty*, 14, 103–127.

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.

Merkle, C. and M. Weber (2011). “True overconfidence: The inability of rational information processing to account for apparent overconfidence.” *Organizational Behavior and Human Decision Processes*, 116, 262–271.

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.

Schweizer, N. and M. De Vries (2022). “Limited Attention and Dynamically Distorted Beliefs.” *Available at SSRN 4177807*.

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

Svenson, O. (1981). "Are we all less risky and more skillful than our fellow drivers?" *Acta psychologica*, 47, 143–148.

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.

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

Proof of Lemma A.1. Equation (7) implies that

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

Therefore,

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

□

Proof of Lemma A.2. First we show that for every $\sigma \in A^S$, $s \in S$, and $\varepsilon > 0$, $ML_s^{\Theta,m}(\sigma)$ and $ML_s^{\Theta,m}(\sigma,\varepsilon)$ are nonempty and compact. By Assumption 1, there exists a $p' \in \Theta$ such that

$$
E\colon = \sum_{s'\in S}\zeta(s')\sum_{y\in Y}m_s(s',\sigma(s),y)p^*_{s',\sigma(s')}(y)\log p'_{s',\sigma(s')}(y) < \infty,
$$

so the function

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

is finite-valued and continuous on the nonempty and compact set

$$\Theta \cap \{p : \sum_{s' \in S} \zeta(s') \sum_{y \in Y} m_s(s', \sigma(s), y) p^*_{s', \sigma(s')}(y) \log p_{s', \sigma(s')}(y) \leqslant E\}$$

Therefore $ML_s^{\Theta,m}(\sigma)$ is nonempty and compact by Theorem 2.43 in Aliprantis and Border [2013]. The result for $ML_s^{\Theta,m}(\sigma, \varepsilon)$ 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 $\sigma(s') \notin BR(s', \Delta(ML_{s'}^{\Theta,m}(\sigma)))$. The upper hemicontinuity of the best reply map $BR(s', \cdot)$ and the compactness of $ML_{s'}^{\Theta,m}(\sigma, \varepsilon)$ imply that there are $\varepsilon, C \in \mathbb{R}_{++}$ such that if $\frac{\nu(ML_{s'}^{\Theta,m}(\sigma,\varepsilon))}{1-\nu(ML_{s'}^{\Theta,m}(\sigma,\varepsilon))} > C$ then $\sigma(s') \notin BR(s', \nu)$.

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

Proof of Lemma A.3. Fix an arbitrary optimal policy $\tilde{\pi}$, $t \in \mathbb{N}$, and a history $(s^t, a^t, y^t) \in H_t$ with $\mathbb{P}_{\tilde{\pi}}(s^t, a^t, y^t) > 0$. Let

$$\tau = \min\{t' > t \colon \sigma(s_{t'}) \neq \tilde{\pi}((s^{t'}, a^{t'}, y^{t'})(R))(s_{t'+1})\}$$

be the first time after (s^t, a^t, y^t) when $\tilde{\pi}$ does not prescribe σ. Note that since

$$\tilde{\pi}((s^{\hat{t}}, a^{\hat{t}}, y^{\hat{t}})(\hat{R}))(s_{\hat{t}+1}) = \sigma(s_{\hat{t}+1}) = \pi^{\sigma,a^t}((s^{\hat{t}}, a^{\hat{t}}, y^{\hat{t}})(\hat{R}))(s_{\hat{t}+1})$$

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 $\mathbb{P}_{\tilde{\pi}}(s^{t'+1}, a^{t'}, y^{t'}, R') > 0$ implies $\mathbb{P}_{\sigma,a^t}(s^{t'+1}, a^{t'}, y^{t'}, R') > 0$, 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. Thus, since for every optimal policy $\tilde{\pi} \in A^{\mathcal{H}}$

$$\mathbb{P}_{\tilde{\pi}}\left[\sup\{t \colon a_t \neq \sigma(s_t)\} < \infty\right] \leqslant \sum_{t=0}^{\infty} \sum_{h_t \in H_t} \mathbb{P}_{\tilde{\pi}}\left[\sigma(s_{\tau+1}) = \tilde{\pi}(h_\tau(R_\tau))(s_{\tau+1}), \forall \tau \geqslant t | h_t\right] \mathbb{P}_{\tilde{\pi}}[h_t] = 0,$$

σ is not a limit strategy. □

B.2 Remark 2

Proof of Remark 2. To prove the statements we give three examples with a singleton signal space.

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 maximizers, which cannot arise from selective memory with a full-support prior. This follows from full support and the strict concavity of the memory-weighted likelihood if m is strictly positive for all positive-probability experiences, and is immediate if $m = 0$ for some (a, y).

2. Suppose that $Y = \{-1, 0, 1\} = A$, $u(a, y) = ay + \mathbb{1}_{a=-1}/20 - \mathbb{1}_{a=1}/12$, and the probability distribution over outcomes is $(1/2, 1/4, 1/4)$ regardless of a, with $\Theta = \{(1/2, 1/6, 1/3), (1/3, 1/6, 1/2)\}$ and $m(a, y) = 1 - \frac{\mathbb{1}_{\{(a,-1):a \in A\}}(a,y)}{2}$. Then both elements of Θ are memory-weighted ML maximizers. Moreover, 0 is a selective memory equilibrium that can only be sustained with beliefs that assign a probability between 1/4 and 7/20 to the data-generating process $(1/2, 1/6, 1/3)$ and in particular must be nondegenerate. But when the agent has perfect memory, there is no Θ' for which both elements of Θ are ML maximizers. Thus 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.3 Proof of Proposition 8

Proof of Proposition 8. We first derive the long-run belief for a given subjective memory function $\hat{m}$. For every σ,

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

Taking first-order conditions of the associated Lagrangian shows there is a unique element p of $L^{\Theta,m,\hat{m}}(\sigma)$, 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 weakly higher in FOSD than that under the subjective memory function $\hat{m}_\alpha$. $\square$

B.4 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 Appendix A.3. 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 an agent remembers the $B \subseteq I$ outcome components of a past realization of experience (s,a,y) and

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

Moreover, the agent believes that they remember an occurrence of signal s and outcome y with probability $\hat{m}_{s'}(s,y,B)$. Thus the *recalled history* at time t is the of recalled experiences $(s_\tau, y_\tau, B_{\tau,t})_{\tau=1}^t$ where $B_{\tau,t}$ denotes the components of the period τ outcome recalled at time t, and for all Borel measurable $C \subseteq \Theta$

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

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 likelihood maximizers:

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

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 function:

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

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.5 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 it is recalled, it is never forgotten; if it is not, it is never remembered. Then the belief process has the following recursive formula: for all Borel measurable $C \subseteq \Theta$,

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

It is easy to see that if the strategies in this dynamic system converge, they converge to a selective memory equilibrium. However, as in Example 8 on underinference, the fact that permanent memory is "less stochastic" allows behaviors that are not limit strategies under selective memory to be limit strategies.

Example 11. 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. ▲