

Learning in a Small/Big World^{*}

Benson Tsz Kin Leung[†]

May 4, 2026

Abstract

Complexity and limited ability affect how we learn and make decisions under uncertainty. Using finite automata to model belief formation, this paper studies the characteristics of optimal learning behavior in small and big worlds, where the complexity of the environment is low and high, respectively, relative to the cognitive ability of the decision maker. The optimal long-run average payoff and distribution of actions are well approximated by the Bayesian benchmark in very small worlds, but the gap widens as the worlds get bigger. In big worlds, it can be optimal for the decision maker to ignore some states entirely, concentrating limited cognitive resources on a subset of states. This “optimal ignorance” is reminiscent of well-documented non-Bayesian learning patterns such as the use of heuristics, correlation neglect, and inattentive learning. The results establish a relationship between complexity, cognitive ability, and the scope for non-Bayesian learning behavior.

Keywords: Learning, Bounded Memory, Bayesian, Complexity, Cognitive Ability

JEL codes: D83, D91

^{*}I am grateful to Tilman Börgers, Jacques Crémer, Matthew Elliott, Renato Gomes, Philippe Jehiel, Hamid Sabourian, Mikhael Safronov, Larry Samuelson, Tak-Yuen Wong, anonymous referees, and the audiences at various seminars and conferences for their insightful discussions and comments.

[†]Hong Kong Baptist University. Email: btkleung@hkbu.edu.hk, ltkbenson@gmail.com.

1 Introduction

Savage (1972) introduces the distinction between “small” and “big” worlds, arguing that Bayesian decision theory applies only to small world problems but not big world problems, where the latter refers to scenarios where it is difficult to form prior beliefs on states and signal structures or construct the state space.¹ This paper borrows the terminology from Savage (1972) but offers an alternative distinction between “small” and “big” worlds based on the complexity of the inference problem relative to the cognitive ability of individuals. “Small” worlds refers to inference problems that are simple relative to the cognitive ability of individuals. By comparing learning behaviors in the “small” and “big” worlds, this paper sheds light on how complexity and cognitive ability affect learning behaviors. Analyzing a theoretical model with finite automata, I show that optimal learning behavior in big worlds exhibits ignorance which is reminiscent of well-documented non-Bayesian patterns, including the use of heuristics (Kahneman et al. (1982)), correlation neglect (Enke and Zimmermann (2019)), persistent over-confidence (Hoffman and Burks (2017)), and inattentive learning (Graeber (2023)), but not in small worlds. Thus, these “biased” learning behaviors are more prominent in more complex problems and among individuals with lower cognitive ability.

More specifically, I consider a decision maker (DM) who tries to learn the true state of the world from a finite state space, where the number of possible states N measures the complexity of the inference problem. In each period $t = 1, \dots, \infty$, the DM guesses the true state of the world and gets a higher utility if he is correct. After making a guess, he receives a signal and updates his belief. As $t \rightarrow \infty$, a Bayesian would learn perfectly the true state. In contrast, to model limited cognitive ability, I assume that the DM’s belief is confined to an M -sized automaton that captures limited cognitive ability, as in the seminal work of Hellman and Cover (1970) and Wilson (2014).² In each period, he starts in one of the M memory states, makes his guess given his memory state based on a decision rule, receives a signal and transits to another memory state based on a transition rule, and finally enters the next period. In contrast to the Bayesian model, the DM has a coarser idea of the likelihood of different states of the world, and the coarseness decreases in M . Thus M measures the cognitive ability of the DM. I define small worlds as cases where $\frac{N}{M}$ is small, otherwise the decision problem is a big world: Whether a problem is a small or big world depends on the relative complexity of the world with respect to the individual’s cognitive ability.³

It is, in general, not tractable to explicitly characterize the optimal updating mechanisms, i.e., the starting memory state, transition rule, and decision rule, that maximize the asymp-

¹In Savage’s words: “[A person] is free to . . . look before he leaps . . . when the world is sufficiently small” (Savage, 1972, p. 16).

²See also Compte and Postlewaite (2012), Monte and Said (2014), Basu and Chatterjee (2015), Chauvin (2019), Chatterjee and Sabourian (2020), Chatterjee and Hu (2023) in the economic literature that model belief updating and the aversion of complexity with finite automata. See also Oprea (2020) and Banovetz and Oprea (2023) for experimental evidence.

³Note that if the individual tracks his belief not with a finite automaton but with a real number statistic, the cardinality of the belief statistics is much larger than N , and the model collapses to a Bayesian model.

	Small Worlds: low complexity relative to cognitive ability $\frac{N}{M}$	Big Worlds: high complexity relative to cognitive ability $\frac{N}{M}$
Are actions close to Bayesian?	Yes	No
Could ignorance in learning be “optimal”?	No	Yes
Could disagreement be persistent?	No	Yes

Table 1: Differences in learning behaviors in small/big worlds

totic utility of the DM: the design space is a product of $M \times |S|$ stochastic transition matrices and M action assignments, and the objective, while linear in the stationary distributions, is a non-convex and non-continuous function of the transition matrices. However, without showing the characterization, I am able to compare the characteristics of the optimal updating mechanisms in small and big worlds, and hence shed light on how perceived complexity affects learning. The results are summarized in Table 1. First, I analyze how the individual’s actions differ from the Bayesian benchmark in small and big worlds. This sheds light on under what circumstances the Bayesian model serves as a good approximation of decision-making under uncertainty. I show that in small worlds, the DM’s long-run average payoff and distribution of actions approach the Bayesian benchmark. In particular, as $\frac{N}{M} \rightarrow 0$, the DM almost always makes the same guess as a Bayesian individual. In contrast, when the world is bigger, i.e., when $\frac{N}{M}$ increases, the DM makes more mistakes and his payoff diverges further from the Bayesian benchmark.

The second, which is the main result of this paper, shows that in big worlds, as the DM faces a trade-off in allocating his scarce cognitive resources, i.e., the M memory states, it is optimal to ignore some states and focus learning on a subset of states in some decision environments. In contrast, such “ignorant” behavior is never optimal in small worlds. This shows a relationship between complexity and ignorance in learning. The structure of optimal ignorance is reminiscent of several well-documented learning biases — including the use of heuristics, correlation neglect, persistent over-confidence, and inattentive learning.

Before I briefly describe the intuition of the result, to see how “ignorance” captures different biases, consider the example of persistent over-confidence (Hoffman and Burks (2017), Heidhues et al. (2018)). Suppose that the state of the world comprises the DM’s and his teammate’s ability, where both could be high or low, and the DM observes team performances as signals. In the current setting, the DM is observationally indistinguishable from exhibiting persistent over-confidence when he never guesses the states where his ability is low, behaves as if he always believes he has high ability, and only updates his belief about his teammate’s ability, even after observing a sequence of bad team performance. Similarly, for correlation neglect (Enke and Zimmermann (2019)), consider that the state of the world comprises a strong or weak stock market and positive, no, or negative correlation among data. The DM is observationally indistinguishable from exhibiting correlation neglect when

he ignores the possible correlation among data and only updates his belief about the stock market as if he always believes there is no correlation.

As mentioned above, this paper shows that such ignorant behavior is optimal only when relative complexity is large (big worlds). When relative complexity is low (small worlds), the DM faces little trade-off between learning about different states of the world and therefore does not ignore any states of the world. The trade-off kicks in when relative complexity is high, which makes it optimal for the DM to ignore some states of the world to focus learning on other states, especially when the ignored state is a priori unlikely or when information supporting that state is weak. For example, if the DM is confident about his ability, he focuses on learning his teammate's ability; if it is difficult to distinguish states with and without correlation, the DM ignores the possibility that data are correlated. More surprisingly, I show that even if the states and information structures are symmetric, ignorance is optimal in environments where it is difficult to learn, e.g., when signals are noisy or when the state space is large. It echoes that complexity, or difficulty in learning, drives ignorance, and shows that optimal ignorance is not a knife-edge result.

Lastly, I analyze whether disagreements persist (with an infinite sequence of i.i.d. signals) in both small and large worlds. As the asymptotic distribution of actions is close to the Bayesian benchmark in small worlds, intuitively, individuals would eventually agree and make the same guesses. In contrast, in big worlds, because individuals with different prior beliefs and/or cognitive ability adopt different optimal learning mechanisms and could ignore different states, they could disagree with each other with probability 1 even after receiving the same infinite sequence of public information. For example, after observing a large sequence of bad team performance, two individuals with persistent over-confidence would disagree on the assessment of their abilities: they ignore the states where their ability is low and attribute the bad performance to the other person. Moreover, I show a novel driving force of disagreement: even when two individuals have the same prior beliefs and observe the same infinite sequence of public signals with no uncertainties in signal structures, they could eventually disagree with each other when they have different levels of cognitive ability M .

This paper is organized as follows. In the next section, I briefly discuss how this paper relates to the literature. Section 3 presents the model. Section 4 shows a series of illustrative examples. I analyze the optimal learning behavior in small and big worlds in Section 5, including limit results as $N, M \rightarrow \infty$. In Section 6, I discuss both research and managerial implications. Lastly, I conclude. The proofs are presented in the Appendix, and additional extensions are presented in the online Appendix.⁴

2 Literature

In this section, I discuss the existing literature and the contribution of this paper.

⁴The online Appendix could be found on <https://sites.google.com/site/ltkbenson/research>.

First, this paper is obviously related to the literature using finite automata to model learning with aversion to complexity (Hellman and Cover (1970), Börgers and Morales (2004), Compte and Postlewaite (2012), Wilson (2014), Chatterjee and Hu (2023), etc.).⁵ While these papers discuss the impact of complexity on learning, they do not analyze how *different levels* of complexity and cognitive ability affect learning and the prominence of learning biases. Moreover, the studies above focus on a binary state space because of technical difficulties. While this paper does not “fully” characterize the optimal automaton, the results shed light on its characteristics when $N > 2$. In particular, the results about ignorance suggest that for larger state spaces, unlike when $N = 2$, ignorance is a crucial feature even when the states are a priori the same.

This paper also contributes to a growing literature that explains behavioral anomalies as optimal/efficient strategies in light of limited cognitive ability. Sims (2003) and Matějka and McKay (2015) study the implications of rational inattention and show that it explains sticky prices in the market and micro-founds the multinomial logit choice model. Steiner and Stewart (2016) shows that an optimal response to noises in perceiving the details of lotteries leads to probability weighting in prospect theory. Jehiel and Steiner (2020) and Leung (2020) show that a capacity constraint on the number of signals that individuals could update their beliefs with drives confirmation bias. In contrast, this paper connects to a larger class of biases under the same framework and illustrates a relationship between their prominence and the level of perceived complexity.

Similar to Section 5.2 in this paper, Caplin et al. (2019) present conditions where the DM would ignore some actions in a rational inattention setting. Unlike this paper, they only present conditions depending on the prior belief and the utility matrix, but not the level of complexity nor the informativeness of signals. Moreover, in contrast to Caplin et al. (2019), in this paper, I show that it could be optimal to ignore some states even in symmetric environments where prior belief is uniform and the utility matrix is symmetric.

Last, this paper’s results on asymptotic disagreement contribute to the large literature that explains the phenomenon. In the existing literature, asymptotic disagreement is driven by differences in signal distributions across states or differences in learning mechanisms (Morris (1994), Mailath and Samuelson (2020), Gilboa, Samuelson and Schmeidler (2022)), the lack of identification or uncertainty in signal distributions (Acemoglu, Chernozhukov and Yildiz (2016)), confirmation bias (Rabin and Schrag (1999)), or model misspecification (Freedman (1963, 1965), Berk (1966)). Differently, this paper looks into the connection between limited ability and disagreement, and shows when asymptotic disagreement could arise and when it will not occur, depending on the relative complexity of the inference problem. Moreover, I show a novel mechanism that disagreement could arise solely because of differences in cognitive abilities.

⁵Also see Chatterjee and Sabourian (2020) for a review, and Oprea (2020) and Banovetz and Oprea (2023) for empirical evidence.

3 Model

I consider a world with $N \geq 2$ possible true states, i.e., $\omega \in \Omega = \{1, 2, \dots, N\}$, and a decision-maker (DM). In each period $t = 1, \dots, \infty$, the DM receives a signal and tries to guess what the true state of the world is. Formally, the unknown state of the world ω is realized before period 1 and is unchanged across all periods. In each period t , the DM takes an action $a_t \in \mathcal{A} = \Omega$ and gets utility $u(a, \omega) \in \mathcal{R}$ where $a = \omega$ is the unique maximizer of $u(a, \omega)$.⁶ The DM does not observe his utility after taking action; thus, $u(a, \omega)$ is best interpreted as an intrinsic utility of being correct. Otherwise, the problem becomes trivial as the DM will experiment and learn perfectly the true state after observing the utility. The prior belief of the DM is $(p^\omega)_{\omega=1}^N$ where $\sum_{\omega=1}^N p^\omega = 1$ and $p^\omega > 0$ for all $\omega \in \Omega$.

In each period after taking an action, the DM receives a signal s_t that is independently drawn from a set S across different periods.⁷ For the ease of exposition, I assume that S is finite and the probability mass function in state ω is denoted as f^ω .⁸ I assume that no signal perfectly rules out any state of the world, i.e., $f^\omega(s) > 0$ for all (s, ω) , such that asymptotic learning is non-trivial. We denote

$$\varsigma \equiv \min_{\omega' \neq \omega, s \in S} \frac{f^\omega(s)}{f^{\omega'}(s)}.$$

As S and Ω are finite, $\varsigma > 0$. Intuitively, ς measures (inverse) signal informativeness: small ς means some signal sharply rules out one state in favor of another, while ς close to 1 means no signal carries strong information. Moreover, I assume no pairs of signal structures are the same, i.e., there are no ω and $\omega' \neq \omega$ such that $f^\omega(s) = f^{\omega'}(s)$ for all $s \in S$. This implies that states are identifiable, and a Bayesian learns the true state perfectly as $t \rightarrow \infty$.⁹ In contrast, I focus on the bounded memory setting I now describe.

The DM is subject to a memory constraint such that he can only update his belief using an M -memory-state automaton. In each period, his belief is represented by a memory state $m_t \in \{1, 2, \dots, M\}$. With some abuse of notations, I sometimes use M to denote the set of memory states. An updating mechanism specifies an initial state $m_1 \in \Delta M$, a (stationary) transition function between the M memory states given a signal $s \in S$, which is denoted as $\mathcal{T} : M \times S \rightarrow \Delta M$, and a (stationary) decision rule $d : M \rightarrow A$.¹⁰ For the ease of exposition,

⁶This assumption rules out “safe” actions that are not optimal under complete information, i.e., actions that are not maximizers in any state but yield good payoffs in multiple states. However, actions that are maximizers in some state and also yield good payoffs in other states are not ruled out.

⁷The order, i.e., whether the DM receives a signal before or after taking an action in each period, does not affect the result.

⁸To ease exposition, I assume that signals follow a discrete distribution, but all results hold with more general probability measures including non-absolutely continuous distributions.

⁹More specifically, [Doob \(1949\)](#) shows that Bayesian posterior converges to a mass point at the true state of the world with probability 1.

¹⁰Restricting attention to deterministic decision rules is without loss of generality: given any stochastic decision rule, as the long-run average payoff is linear in the action probabilities at each memory state and additively separable across memory states, one can replace each memory state’s randomized action with the action that maximizes the expected per-period payoff at that memory state, weakly increasing the long-run

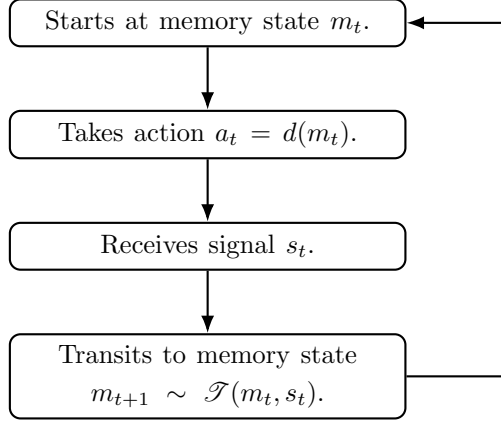


Figure 1: Timeline at period t given an updating mechanism $(\mathcal{T}, d)$.

I denote the set of memory states where the DM chooses action ω as M^ω .

In each period $t = 1, 2, \dots$, as illustrated in Figure 1, the DM starts with some memory state m_t , takes action $a_t = d(m_t)$, receives a signal s_t , and transits to memory state $m_{t+1} \sim \mathcal{T}(m_t, s_t)$. This paper analyzes the asymptotic learning of the DM, i.e., before period 1, the DM designs and commits to an updating mechanism that maximizes his expected long-run per-period utility:¹¹

$$\lim_{T \rightarrow \infty} E_{\omega, m_1, \mathcal{T}, d} \left[\frac{1}{T} \sum_{t=1}^T u(a_t, \omega) \right].$$

An alternative setup is that the DM maximizes the discounted sum of utility as in Wilson (2014). As shown in Appendix B, the results in this paper hold qualitatively when the discount factor is close to 1.

Given a state $\omega \in \Omega$, the sequence m_t , together with some specified initial memory state m_1 , forms a Markov chain. Denote μ_m^ω as the long-run proportion of time that the DM is in memory state m when the true state of the world is ω , and $\mathbf{Q}^\omega$ as the matrix of transition probabilities, i.e., $\mathbf{Q}^\omega = [\sum_{s \in S} \Pr\{\mathcal{T}(m, s) = m'\} f^\omega(s)]_{mm'}$. By the Birkhoff–Khinchin theorem, the distribution $\boldsymbol{\mu}^\omega = (\mu_1^\omega, \mu_2^\omega, \dots, \mu_M^\omega)^T$ solves the following equation:

$$\boldsymbol{\mu}^\omega = (\boldsymbol{\mu}^\omega)^T \mathbf{Q}^\omega. \quad (1)$$

By the Brouwer fixed-point theorem, a solution always exists. When there is a unique solution, e.g., when the transition matrix is irreducible, $\boldsymbol{\mu}^\omega$ depends only on the transition function $\mathcal{T}$ but not on the initial memory state; when there are multiple solutions, it implies that the Markov chain is reducible, and the long-run distribution is one of the solutions of Equation (1) and is uniquely pinned down by the initial memory state. For simplicity of exposition, I omit the arguments in $\boldsymbol{\mu}^\omega$ throughout the paper.

The asymptotic utility, or the long-run per-period utility, of an updating mechanism average payoff. Thus, for ease of exposition, I restrict attention to deterministic decision rules unless stated otherwise.

¹¹Switching between multiple M -memory-state automata requires more than M memory states, and the current setting allows switching between smaller automata, as illustrated in Online Appendix A.

$(m_1, \mathcal{T}, d)$ is equal to:

$$U(m_1, \mathcal{T}, d) \equiv \lim_{T \rightarrow \infty} E_{\omega, m_1, \mathcal{T}, d} \left[\frac{1}{T} \sum_{t=1}^T u(a_t, \omega) \right] = \sum_{\omega=1}^N \left[p^\omega \left(\sum_{m=1}^M u(d(m), \omega) \mu_m^\omega \right) \right] \quad (2)$$

and the asymptotic utility loss is equal to:

$$L(m_1, \mathcal{T}, d) = \sum_{\omega=1}^N [p^\omega u(\omega, \omega)] - U(m_1, \mathcal{T}, d). \quad (3)$$

The DM maximizes the asymptotic utility or, equivalently, minimizes the asymptotic utility loss. I mostly refer to the optimal design of the updating mechanism as the minimization of L . In general, following similar arguments as in [Hellman and Cover \(1970\)](#), an optimal mechanism may not exist.¹² Therefore, the rest of the paper focuses on ϵ -optimal updating mechanisms (for small ϵ) that are defined as follows. Define

$$L_M^* \equiv \inf_{m_1, \mathcal{T}, d} L(m_1, \mathcal{T}, d).$$

An updating mechanism $(m_1, \mathcal{T}, d)$ is ϵ -optimal if and only if $L(m_1, \mathcal{T}, d) \leq L_M^* + \epsilon$. Throughout the paper, I focus on the case where

$$L_M^* < \min_a \sum_{\omega=1}^N p^\omega [u(\omega, \omega) - u(a, \omega)], \quad (4)$$

where the right-hand side is the minimum long-run utility loss from always taking the same fixed action regardless of signals. This inequality says that the DM can strictly improve his utility by learning, i.e., the memory constraint is not so severe as to make learning worthless.

In [Section 5](#), I compare the characteristics of the ϵ -optimal updating mechanisms in small and big worlds. Roughly speaking, N represents the complexity of the world, and M represents the cognitive ability of the DM. I call $\frac{N}{M}$ the relative, or the perceived, complexity. This gives a natural definition of small and big worlds: an inference problem is a small world when $\frac{N}{M}$ is small, and is a big world otherwise. Throughout the paper, unless stated otherwise, I assume $M \geq N$, so that the DM has enough memory states to potentially choose every action.¹³

¹²Although the set of transition probabilities is compact, an optimal mechanism may not exist because the long-run distribution μ^ω is not continuous in the transition probabilities. To illustrate, consider $M = 2$ and $\mathcal{T}(m, s) = (1 - \delta) \times \{m\} + \delta \times \{m'\}$ for $m' \neq m$ and $m = 1, 2$. That is, the DM transits from memory state 1 to 2 (and 2 to 1) with probability δ . For all $\delta \in (0, 1]$, the long-run distribution is $(\frac{1}{2}, \frac{1}{2})$; however, at $\delta = 0$ the long-run distribution jumps to $(1, 0)$ or $(0, 1)$ depending on the starting memory state. The optimal loss L_M^* is therefore defined as an infimum that may not be attained.

¹³When $M < N$, the DM is forced to ignore at least $N - M$ states (never take those actions) simply because he lacks enough memory states for them, making ignoring trivially necessary rather than an optimal allocation of memory states.

4 Illustrative Examples

Before presenting the main results, I go through a series of examples to illustrate how the design of an optimal updating mechanism changes as N and M vary. The examples build on each other and collectively motivate some of the research questions and main results of the paper.

Example 1: $N = 2, M = 2$

First consider $N = 2$ states with uniform prior $p^1 = p^2 = \frac{1}{2}$, $u(a, \omega) = 1$ if $a = \omega$ and 0 otherwise. The signal space is binary, i.e., $S = \{s^1, s^2\}$, and the DM has $M = 2$ memory states. When $N = 2$, each signal can be (deterministically) defined as supporting one of the two states by comparing the likelihood of drawing that signal in state 1 versus state 2: signal s supports state ω if and only if $f^\omega(s) > f^{-\omega}(s)$. Without loss of generality, suppose s^1 supports state 1. Because $f^\omega(s^2) = 1 - f^\omega(s^1)$, s^2 must support state 2. For simplicity, in this example, I also assume that $f^1(s^1) = f^2(s^2)$, so that the information structure is symmetric. Denote $\mathcal{J} \equiv \frac{f^1(s^1)}{f^2(s^1)} = \frac{f^2(s^2)}{f^1(s^2)}$, which measures the informativeness of the signals.

An optimal updating mechanism is illustrated in Figure 2: the DM chooses action ω in memory state ω , and switches memory states if and only if he receives a signal supporting the other state. The optimal updating mechanism is symmetric: the transition and decision rules are invariant to the permutation of states and memory states. In this optimal updating mechanism, the long-run distributions satisfy:

$$\begin{aligned}\mu_1^1 f^1(s^2) &= \mu_2^1 f^1(s^1) \\ \mu_1^2 f^2(s^2) &= \mu_2^2 f^2(s^1).\end{aligned}$$

That is, in both states 1 and 2, the long-run probability of transiting out of memory state 1 equals the long-run probability of transiting into memory state 1. Rearranging gives us:

$$\frac{\mu_1^1}{\mu_2^1} = \frac{\mu_2^2}{\mu_1^2} = \mathcal{J}$$

and

$$U = \frac{\mathcal{J}}{1 + \mathcal{J}}; \quad L = \frac{1}{1 + \mathcal{J}}.$$

Intuitively, the DM transits to memory state 1 after receiving s^1 and to memory state 2 after receiving s^2 . Combined with the fact that the DM receives s^ω more often in state ω than in state $-\omega$, this implies that the DM is more likely to be in memory state ω in state ω , matching his action with the state. Moreover, the more informative the signals are, the lower the utility loss.

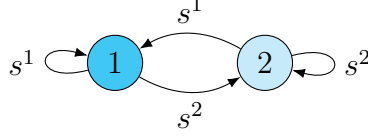


Figure 2: Optimal updating mechanism with $N = 2$, $M = 2$, and $S = \{s^1, s^2\}$, where $f^1(s^1) = f^2(s^2) > f^1(s^2) = f^2(s^1)$. Decision rule follows: $d(1) = 1$, $d(2) = 2$. The DM transits to memory state ω upon receiving signal s^ω , and otherwise stays in the current memory state.

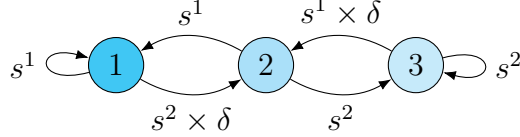


Figure 3: Optimal updating mechanism with $N = 2$, $M = 3$, and $S = \{s^1, s^2\}$. The DM chooses action 1 in both memory states 1 and 2, and action 2 in memory state 3. Memory state 1 acts as a “high-confidence” state for action 1: an s^2 signal only demotes confidence to memory state 2, not all the way to action 2.

Example 2: $N = 2$, $M = 3$

Now suppose the DM gains one extra memory state, so $M = 3$, while keeping all other model parameters the same. Consider the updating mechanism illustrated in Figure 3. Similar to the example with $N = M = 2$ illustrated in Figure 2, the DM transits to a higher memory state only if he receives a signal supporting state 2, and to a lower memory state only if he receives a signal supporting state 1. The decision rule is $d(1) = d(2) = 1$ and $d(3) = 2$.

Crucially, although the DM chooses action 1 in both memory state 1 and 2, he is more confident about action 1 in memory state 1 than in memory state 2. To see this, note that the long-run distribution follows:

$$\begin{aligned}\mu_1^1 \delta f^1(s^2) &= \mu_2^1 f^1(s^1); & \mu_2^1 f^1(s^2) &= \mu_3^1 \delta f^1(s^1) \\ \mu_1^2 \delta f^2(s^2) &= \mu_2^2 f^2(s^1); & \mu_2^2 f^2(s^2) &= \mu_3^2 \delta f^2(s^1)\end{aligned}$$

Rearranging gives us:

$$\begin{aligned}\mu_1^1 &= \frac{\mathcal{J}^2}{1 + \delta \mathcal{J} + \mathcal{J}^2}; & \mu_2^1 &= \frac{\delta \mathcal{J}}{1 + \delta \mathcal{J} + \mathcal{J}^2}; & \mu_3^1 &= \frac{1}{1 + \delta \mathcal{J} + \mathcal{J}^2} \\ \mu_1^2 &= \frac{1}{1 + \delta \mathcal{J} + \mathcal{J}^2}; & \mu_2^2 &= \frac{\delta \mathcal{J}}{1 + \delta \mathcal{J} + \mathcal{J}^2}; & \mu_3^2 &= \frac{\mathcal{J}^2}{1 + \delta \mathcal{J} + \mathcal{J}^2}\end{aligned}$$

and the likelihood ratio of each memory state is:

$$\frac{\mu_1^1}{\mu_1^2} = \mathcal{J}^2; \quad \frac{\mu_2^1}{\mu_2^2} = 1; \quad \frac{\mu_3^1}{\mu_3^2} = \mathcal{J}^{-2}.$$

Therefore, in memory state 2, the DM is unsure about the true state. The ϵ -optimal updating mechanism involves choosing a strictly positive but small enough δ . First, by choosing

a strictly positive δ , the updating mechanism remains sensitive to information. Second, by choosing a small enough δ , it ensures enough “exploitation” such that the DM predominantly remains in the two “extreme” memory states 1 and 3.¹⁴ Lastly, the chain structure ensures maximum separability of the two extreme nodes, i.e., the ratio of the long-run probability ratios on memory state 1 in state 1 vis-à-vis state 2 over that on memory state 3 is maximized:

$$\frac{\mu_1^1/\mu_3^1}{\mu_1^2/\mu_3^2} = \mathcal{J}^4.$$

As in the example of $N = M = 2$, the utility loss decreases in the signal informativeness, i.e., $L^* = \frac{1}{1+\mathcal{J}^2}$ decreases in $\mathcal{J}$.

Example 3: $N = 3$, $M = 3$

I now increase the state space to $N = 3$ with $M = 3$ memory states, continuing to assume uniform prior $p^1 = p^2 = p^3 = \frac{1}{3}$, and $u(a, \omega) = 1$ if and only if $a = \omega$ and 0 otherwise. The signal space is $S = \{s^1, s^2, s^3\}$ with a symmetric distribution, i.e., $f^\omega(s^{\omega'}) = f^{\omega'}(s^\omega)$. In the following, I discuss how generalizing to $N > 2$ introduces both mathematical complications and novel questions that are not possible under $N = 2$.

Which signal supports which states? A natural starting point is to ask whether we can extend the mechanisms and logic from Examples 1 and 2. However, for $N > 2$, even defining what constitutes a supporting signal for each state becomes more complicated. As argued above, when $N = 2$, a signal supports state ω if and only if $f^\omega(s) > f^{-\omega}(s)$. With this definition, the sets of signals supporting each state are non-empty. However, when $N > 2$, there might exist no signals supporting some state, which makes the design of a good updating mechanism considerably more subtle.

To see why, consider the following symmetric signal structure with $S = \{s^1, s^2, s^3\}$:

	s^1	s^2	s^3
f^1	0.50	0.40	0.10
f^2	0.40	0.20	0.40
f^3	0.10	0.40	0.50

If we extend the definition of supporting signals from the example with $N = 2$, a signal s supports state ω if and only if $\omega = \arg \max_{\omega'} f^{\omega'}(s)$. Under this definition, s^1 supports state 1, s^3 supports state 3, and s^2 supports state 1 or state 3. Thus, no signal supports state 2. In fact, one can show that no deterministic definition of supporting signals can give state 2 any confirmatory signal. Note that the optimal mechanisms when $N = 2$, as in [Hellman and Cover \(1970\)](#) and [Wilson \(2014\)](#), depend on the existence of confirmatory signals for both

¹⁴This example also illustrates why optimal updating mechanisms might not exist. Here the DM wants to choose a small but strictly positive δ . The infimum utility loss is achieved by taking $\delta \rightarrow 0$.

states. Therefore, generalizing the optimal mechanisms from $N = 2$ to $N > 2$ is non-trivial.

As I will discuss in the next section, the fix is to stochastically “label” each signal s as confirmatory for state ω with probability proportional to $f^\omega(s)$, normalizing so that probabilities sum to at most 1 across states (the residual probability corresponds to the case where the signal s is labeled as non-confirmatory for any states). In the example, s^2 is labeled as confirmatory for state 1 with probability 0.4, for state 2 with probability 0.2, and for state 3 with probability 0.4. Under this stochastic labeling, there is a strictly positive probability that the DM receives a signal labeled as supporting ω for all ω . Moreover, the probability of receiving a confirmatory signal for the *correct* state strictly exceeds the probability for any wrong state, in every state of the world:

$$\Pr[\text{signals support } \omega \mid \omega] = \sum_{\omega=1}^3 [f^\omega(s^\omega)]^2 > \sum_{\omega=1}^3 f^\omega(s^\omega) f^{\omega'}(s^\omega) = \Pr[\text{signals support } \omega' \mid \omega],$$

where the inequality follows from the Cauchy–Schwarz inequality. As I will discuss, this stochastic definition of supporting signals and its property—that the DM is more likely to receive a signal supporting the true state—is key to generalizing the updating mechanisms from [Hellman and Cover \(1970\)](#) and [Wilson \(2014\)](#) to $N > 2$.

Ignoring states? For now, I sidestep the problem of defining supportive signals and assume that $\frac{f^\omega(s^\omega)}{f^{\omega'}(s^\omega)} = \mathcal{J} > 1$ for all $\omega \neq \omega'$. That is,

$$f^\omega(s^\omega) = \frac{\mathcal{J}}{\mathcal{J} + 2}, \quad f^\omega(s^{\omega'}) = \frac{1}{\mathcal{J} + 2} \quad \text{for } \omega' \neq \omega,$$

where $\mathcal{J}$ measures the informativeness of the signal as in the examples with $N = 2$.

When $N > 2$, the DM additionally has to decide which states to learn. When $N = 2$, the DM decides how to distinguish between states 1 and 2; when $N = 3$, the DM can either not ignore any state and try to distinguish each state ω from $\Omega \setminus \{\omega\}$, or ignore one state ω'' to focus on distinguishing states ω and ω' .

Non-ignorant mechanism. First, consider a non-ignorant mechanism in which the DM allocates one memory state to each action: $d(m) = m$ for $m \in \{1, 2, 3\}$. Following the symmetric construction with $N = M = 2$, the transition rule is $\mathcal{T}(m, s^\omega) = \omega$ for all m . The stationary distribution under state $\omega = 1$ assigns probability $\mu_1^1 = \frac{\mathcal{J}}{\mathcal{J} + 2}$ to memory state 1 and $\frac{1}{\mathcal{J} + 2}$ to each of memory states 2 and 3. The asymptotic utility and asymptotic utility loss under this mechanism are

$$U^{\text{NI}} = \frac{\mathcal{J}}{\mathcal{J} + 2}; \quad L^{\text{NI}} = \frac{2}{\mathcal{J} + 2}. \quad (5)$$

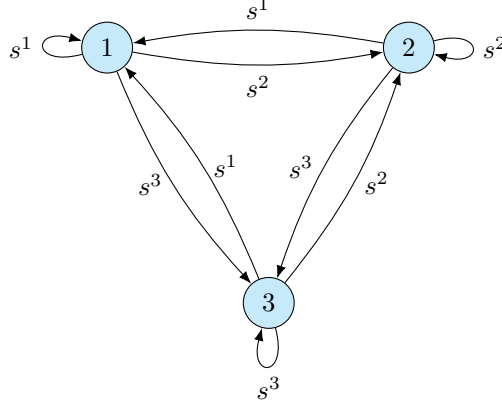


Figure 4: Non-ignorant updating mechanism with $N = M = 3$. The number in each node denotes the action the DM takes in that memory state: $d(m) = m$. Upon receiving signal s^ω , the DM transitions to memory state ω regardless of his current state.

Ignorant mechanism. Now consider an ignorant mechanism: the DM ignores state 3 entirely (so $d(m) \in \{1, 2\}$ for all m) and dedicates $M = 3$ memory states to learning between states 1 and 2. Using the $N = 2$, $M = 3$ construction from Example 2, now illustrated in Figure 5, the utility and utility loss of such an ignorant mechanism when $\delta \rightarrow 0$ satisfy:

$$U^I = \frac{2}{3} \left(\frac{\mathcal{J}^2}{1 + \mathcal{J}^2} \right); \quad L^I = 1 - \frac{2}{3} \left(\frac{\mathcal{J}^2}{1 + \mathcal{J}^2} \right). \quad (6)$$

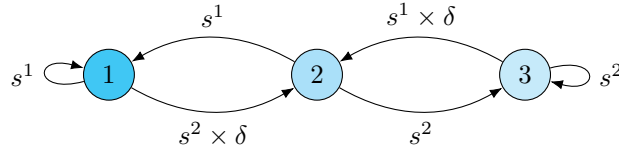


Figure 5: Ignorant updating mechanism with $N = 3$, $M = 3$, ignoring state 3. The DM chooses action 1 in memory states 1 and 2 (shaded), and action 2 in memory state 3. Memory state 1 serves as a “high-confidence” state for action 1: receiving signal s^2 only demotes the DM to low-confidence state 2, not all the way to action 2. Signal s^3 is uninformative about states 1 vs. 2 and triggers no transition.

Comparing the ignorant and non-ignorant mechanisms, the ignorant mechanism obviously yields a lower utility in state 3. However, as the ignorant mechanism allocates three memory states to distinguishing states 1 and 2, utility in those two states is higher. The comparison of utilities across states is shown in Table 2.

Note that utilities in states 1 and 2 are higher under the ignorant mechanism for two reasons. First, by ignoring state 3, the DM can better distinguish states 1 and 2. For example, under the ignorant mechanism, the DM’s probability ratio of choosing action 1 vis-à-vis action 2 in state 1 equals $\mathcal{J}^2$, while the ratio equals $\mathcal{J}$ under the non-ignorant mechanism. This effect contributes to the $\mathcal{J}^2$ terms versus $\mathcal{J}$ terms in Table 2. Second, in $N = 3$, the probability of receiving a “wrong” signal, i.e., $s^{\omega'}$ with $\omega' \neq \omega$ in state ω , equals $\frac{2}{2+\mathcal{J}}$. By ignoring one of the three states, the effective probability of receiving a “wrong”

	Non-Ignorant		Ignorant
$\omega = 1$	$\frac{\mathcal{J}}{2 + \mathcal{J}}$	$<$	$\frac{\mathcal{J}^2}{1 + \mathcal{J}^2}$
$\omega = 2$	$\frac{\mathcal{J}}{2 + \mathcal{J}}$	$<$	$\frac{\mathcal{J}^2}{1 + \mathcal{J}^2}$
$\omega = 3$	$\frac{\mathcal{J}}{2 + \mathcal{J}}$	$>$	0

Table 2: Comparison of asymptotic utility by state under the non-ignorant and ignorant mechanisms with $N = M = 3$.

signal (conditional on non-ignored signals) drops to $\frac{1}{1+\mathcal{J}}$. This contributes to the “2” term versus the “1” term in Table 2. The following observation compares the utility loss of the ignorant and non-ignorant mechanisms.

Observation 1. *There exists a threshold $\bar{\mathcal{J}} > 1$ such that for all $\mathcal{J} \in (1, \bar{\mathcal{J}})$, the ignorant mechanism in Figure 5 achieves strictly lower asymptotic utility loss than the non-ignorant mechanism in Figure 4.*

Intuitively, when signals are noisy ($\mathcal{J}$ close to 1), distinguishing any pair of states is difficult. As the probability of receiving “wrong” signals is high, the non-ignorant mechanism is too sensitive to signals and thus noisy. In contrast, by ignoring one of the three states, the DM can allocate more memory states to distinguishing the other two states and improve utility. Proposition 4 formalizes this logic for general conditions.

How utility changes when both N, M increase? Extending to $N > 2$ also allows us to ask how the DM’s utility changes with the size of the state space and the size of the automaton. How do the DM’s utility and optimal updating mechanism change when the world becomes more complex (N increases) but the DM also becomes more capable (M increases)? I answer that question with limit results where both N, M go to infinity but at different rates.

5 Results

5.1 Are actions close to Bayesian?

In the current setup, a Bayesian individual learns perfectly the true state of the world asymptotically: by Doob (1949), the Bayesian posterior converges almost surely to a mass point at the true state, so a Bayesian chooses $a = \omega$ with probability 1 as $t \rightarrow \infty$ and achieves asymptotic utility loss equal to 0. Thus, we interpret L_M^* as a measure of the distance between the DM’s payoff and distribution of actions and that of a Bayesian individual: the higher L_M^* is, the more likely the DM chooses $a \neq \omega$, differing from the decision of a Bayesian individual.¹⁵

¹⁵More precisely, L_M^* measures the gap in long-run average payoff between the DM and a Bayesian. Under the additional condition $u(\omega, \omega) - u(a, \omega) = u(a, a) - u(\omega, a)$ for all a, ω — which holds, e.g., for the 0-1 loss

The following Proposition presents the formal results on how relative complexity affects L_M^* . For ease of exposition, I denote $\mathbf{u}^\omega \equiv (u(a, \omega))_{a=1}^N$.

Proposition 1. *We have the following results regarding L_M^* :*

- (i) L_M^* is strictly decreasing in M and converges to 0 as $M \rightarrow 0$. Moreover, convergence to the Bayesian benchmark is exponential in $\frac{M}{N}$: for each N and $(\mathbf{u}^\omega, p^\omega, f^\omega)_{\omega=1}^N$, there exist constants $K > 0$ and $r < 1$ such that $L_M^* < Kr^{\lfloor \frac{M-1}{N} \rfloor}$;
- (ii) For each N and $(f^\omega)_{\omega=1}^N$, there exists a single sequence of “prior-free” updating mechanisms $(m_1, \mathcal{T}_M, d_M)_{M=1}^\infty$ — depending only on the signal structures $(f^\omega)_{\omega=1}^N$ but not on the prior beliefs $(p^\omega)_{\omega=1}^N$ or the utility function $(\mathbf{u}^\omega)_{\omega=1}^N$ — such that $\lim_{M \rightarrow \infty} L(m_1, \mathcal{T}_M, d_M) = 0$ for all $(\mathbf{u}^\omega, p^\omega)_{\omega=1}^N$. That is, the same mechanism achieves a distribution of actions that approximates the Bayesian benchmark regardless of the prior and payoffs.

To briefly sum up, since the utility loss L is a weighted average of the probabilities of taking incorrect actions across states, $L_M^* \rightarrow 0$ implies that the long-run fraction of each action converges to the Bayesian benchmark. Thus, the Proposition shows that the DM’s distribution of actions is closer to, and converges to, Bayesian as the DM’s cognitive ability increases. Moreover, the convergence speed depends on the relative complexity of the world, i.e., $\frac{N}{M}$. Proposition 1 presents results of comparative statics w.r.t. M , thus one natural question is whether $\frac{N}{M}$ truly affects the payoff and distribution of actions. When N changes, naturally, the optimal updating mechanism depends hugely on how the prior belief p^ω , utility matrix u and signals structure f^ω change. Later in this subsection, we analyze the case where both $N, M \rightarrow \infty$ and impose extra assumptions at the limit, and show the optimal mechanism depends crucially on the ratio $\frac{N}{M}$ in the limit. In the following, we first briefly present the proof of Proposition 1.

Proving the Proposition is tricky for two reasons. First, characterizing the optimal updating mechanism is intractable because the space of feasible updating mechanisms is large. Moreover, the asymptotic utility loss is neither continuous nor convex in the transition probabilities. As will be described in the next several paragraphs, I prove the Proposition by construction. Second, as mentioned in the illustrative example, it is non-trivial to define the set of confirmatory signals for each state ω when $N > 2$. When $N = 2$, the confirmatory signals for state ω are $S_\omega = \{s \in S : f^\omega(s) > f^{-\omega}(s)\}$, and the optimal updating mechanism involves transiting to a different memory state only if the DM receives a strong enough confirmatory signal for some state (signals with either high enough $\frac{f^1(s)}{f^2(s)}$ or high enough $\frac{f^2(s)}{f^1(s)}$). The set of confirmatory signals for both states defined in this way is non-empty, and because of the binary nature, it is also more likely that the DM receives confirmatory signals for the correct state than for the wrong state, i.e., $\sum_{s \in S_1} f^1(s) ds > \sum_{s \in S_2} f^1(s) ds$. However, these characteristics do not generalize to $N > 2$. As illustrated in Example 3 in Section 4, the

$u(a, \omega) = \mathbb{1}[a = \omega]$ — L_M^* equals the expected total variation distance between the DM’s action and the Bayesian’s action ω .

fix is to define whether a signal s is confirmatory for state ω stochastically. In particular, a signal s confirms state ω' with probability proportional to $f^{\omega'}(s)$, and does not confirm any state with complementary probability. Such a definition ensures that in any state ω , the probability of receiving a confirmatory signal for state ω is higher than that of receiving a confirmatory signal for any other state ω' . In the following, when I say the DM transits from m to m' upon receiving a signal supporting state ω , I mean that the transition probability from m to m' given signal s is proportional to $f^{\omega}(s)$.

Now I briefly describe the proof that L_M^* strictly decreases in M . The strategy is to show that any mechanism with M memory states can be strictly improved by a mechanism with $M + 1$ memory states, and that improvement is bounded below. To illustrate, consider the simple case $N = 2$. Take an ϵ -optimal mechanism $(m_1, \mathcal{T}_M, d_M)$ with M memory states, and let $M^1 = \{m : d(m) = 1\}$ and $M^2 = \{m : d(m) = 2\}$ denote the sets of memory states where the DM chooses action 1 and action 2, respectively; both are non-empty by Equation (4). I construct a new mechanism with $M + 1$ memory states that strictly improves the asymptotic utility, in three steps:

- *Decision rule.* Assign $d(M + 1) = 1$, i.e., the DM chooses action 1 in the new memory state. Keep d unchanged for memory states $1, \dots, M$.
- *Entry into $M + 1$.* When the DM is at some $m \in M^1$ and receives a signal supporting state 1, he transitions to $M + 1$ with probability δ_1 .
- *Exit from $M + 1$.* When the DM is at $M + 1$ and receives a signal supporting state 2, he transitions back to some state in M^1 with probability δ_2 .

All other transitions are unchanged up to renormalization. Figure 6 illustrates the construction. The key insight is that, with the stochastic definition of confirmatory signal, in state 1 the DM receives signals supporting state 1 more often than signals supporting state 2, so entries into $M + 1$ are more frequent than exits — the DM spends more time at $M + 1$ and chooses action 1 more often. Conversely, in state 2, exits are more frequent than entries, so the DM spends less time at $M + 1$ and chooses action 1 less often. By choosing δ_1 and δ_2 appropriately, the net effect is that the DM chooses the correct action more often in state 2 without choosing it less often in state 1, strictly reducing L . Crucially, I show that the improvement is bounded away from 0. The construction extends to $N > 2$.

Now to prove the rest of Proposition 1, I construct a simple updating mechanism, illustrated in Figures 7 and 8. The M memory states are segmented into N branches of equal size and a central memory state, reminiscent of a star network. For ease of exposition, the remaining memory states are discarded, and in the following, I rename the memory states as $\{0\} \cup \{k1, \dots, k\lfloor \frac{M-1}{N} \rfloor\}_{k=1}^N$ such that memory state 0 is the central memory state, and $(k1, \dots, k\lfloor \frac{M-1}{N} \rfloor)$ are memory states in the k -th branch. In all memory states in the k -th branch, the DM chooses action k ; otherwise, e.g., when the DM is in the central memory

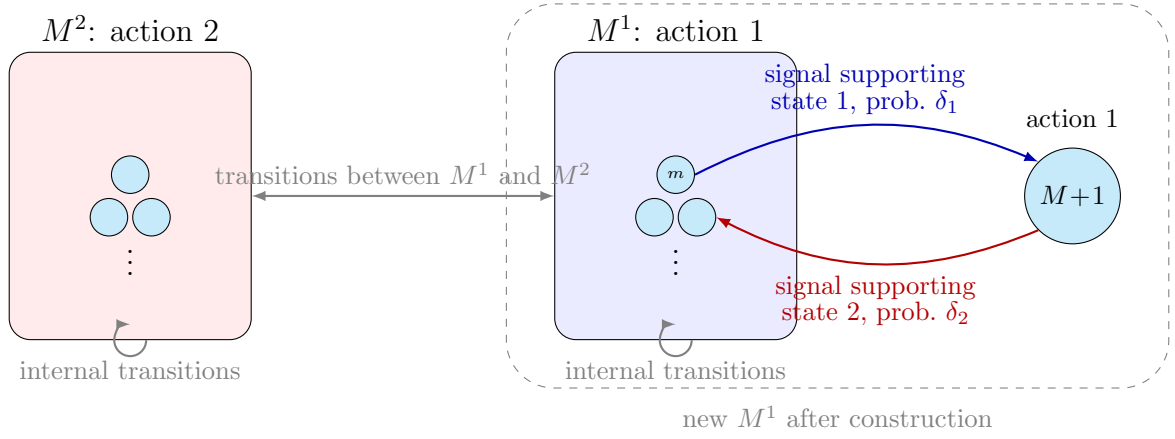


Figure 6: Construction of the $(M + 1)$ -state mechanism from an M -state mechanism (with $N = 2$). The new memory state $M + 1$ is assigned action 1. The DM enters $M + 1$ from M^1 upon receiving a signal supporting state 1 and exits back to M^1 upon receiving a signal supporting state 2. In state 1, entries are more frequent than exits; in state 2, the reverse holds, improving asymptotic utility.

state, he takes action 1.¹⁶ The transition rule is described as follows.

First, the DM starts in the central memory state 0.¹⁷ If he receives a confirmatory signal for state ω , he transits to memory state $\omega 1$, the first state in branch ω ; if he receives a signal that does not confirm any state, he stays in memory state 0. Second, suppose the DM is in memory state ωk , the k -th state in branch ω . If he receives a confirmatory signal for state ω , he moves one state further from the center to $\omega(k+1)$, unless he is already at the extreme of the branch ($k = \lfloor \frac{M-1}{N} \rfloor$), in which case he stays. Third, if he receives a confirmatory signal for some other state $\omega' \neq \omega$, then with probability $\frac{1}{\delta} < 1$ he moves one state closer to the center: to $\omega(k-1)$ if $k \geq 2$, or back to the central state 0 if $k = 1$; with probability $1 - \frac{1}{\delta}$ he stays in his current memory state. Lastly, if he receives a signal that does not confirm any state, he stays in his current memory state. Intuitively, this simple updating mechanism tracks only one action at a time — the action associated with the branch the DM currently occupies — and a more extreme position within the branch reflects greater certainty that this action is optimal. The non-transition upon receiving disconfirmatory signals (captured by $\frac{1}{\delta} < 1$) ensures that the DM does not switch between branches too readily.

The proof involves choosing a sufficiently large δ such that the DM is more likely to move towards the extreme of a branch than towards the center. This ensures enough “exploitation” that the DM does not switch between actions too often. Crucially, when $\frac{M}{N}$ increases, the maximum number of confirmatory signals the DM can “store” in each branch increases, and the more likely the DM is to be in the correct branch choosing the correct action, i.e., matching the state of the world. When $\frac{M}{N} \rightarrow \infty$, the DM almost surely learns perfectly the true state as $t \rightarrow \infty$. Lastly, when N increases, the DM needs more memory states to record confirmatory information for every state, and thus an increase in M induces a smaller

¹⁶The decision rules in memory states not in $\{k1, \dots, k\lfloor \frac{M-1}{N} \rfloor\}_{k=1}^N$ do not affect the proof and result.

¹⁷The starting memory state has no impact on the long-run distribution over the memory states and does not affect the asymptotic payoff.

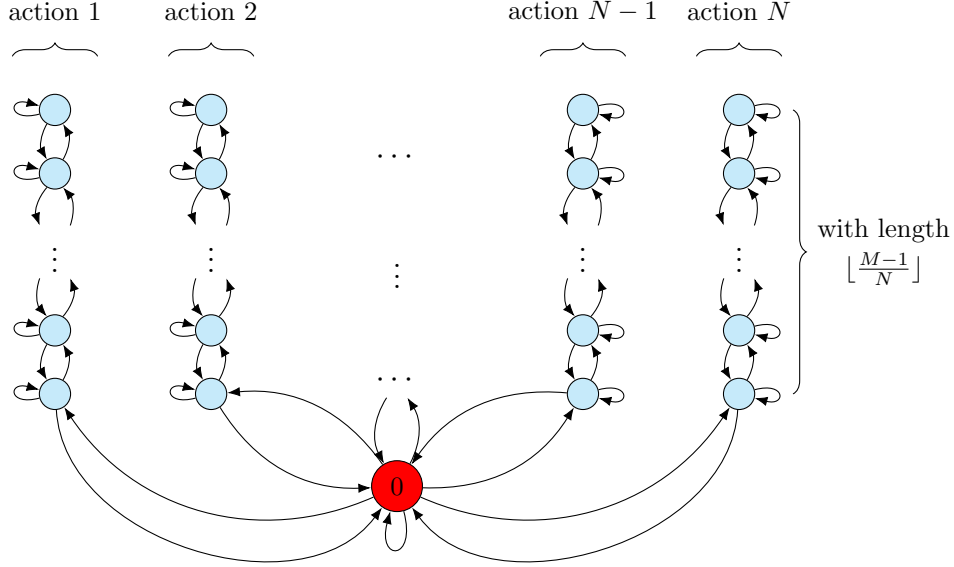


Figure 7: A simple updating mechanism that achieves perfect learning in small worlds for all N , $(p^\omega)_{\omega=1}^N$, and with convergence rate $\frac{1}{N}$.

improvement in the DM’s decisions, explaining why a larger N slows down the convergence rate of L_M^* .

Discussion Here I point out several noteworthy implications of the analysis. Several features of the simple updating mechanism reflect a fundamental difficulty that arises when $N \geq 3$: each signal carries information about multiple pairwise comparisons simultaneously, and the mechanism must use this information efficiently within the memory constraint.

First, as argued previously, unlike [Cover \(1969\)](#) and [Wilson \(2014\)](#) with $N = 2$, when $N > 2$, the definition of confirmatory signals for each state must in general be stochastic in order for the simple updating mechanism to achieve near-optimal performance. This means that upon receiving the same signal, the individual will sometimes regard it as supporting one state, and sometimes regard it as supporting a different state. With $N = 2$, there is a natural deterministic partition of signals into those supporting each state. With $N \geq 3$, however, a signal that is most likely under state ω may also be informative for distinguishing other pairs of states; a deterministic labeling would discard this information, while stochastic labeling allows the mechanism to exploit it. This resembles empirical evidence that documents heterogeneous interpretations of the same pieces of information, especially when the piece of information is imprecise ([Gaines et al. \(2007\)](#)).

Second, the same difficulty explains why it is worth comparing the simple updating mechanism in [Figure 7](#) to a naïve alternative in which the DM directly memorizes past signal realizations. Under such a mechanism, each memory state encodes a specific sequence of recent signals, so the number of memory states required to approximate Bayesian is $|S|^t$, which increases with $|S|$, the size of the signal space. With large state space and signal space, the space of sequences of signals grows combinatorially. In contrast, the simple updating mecha-

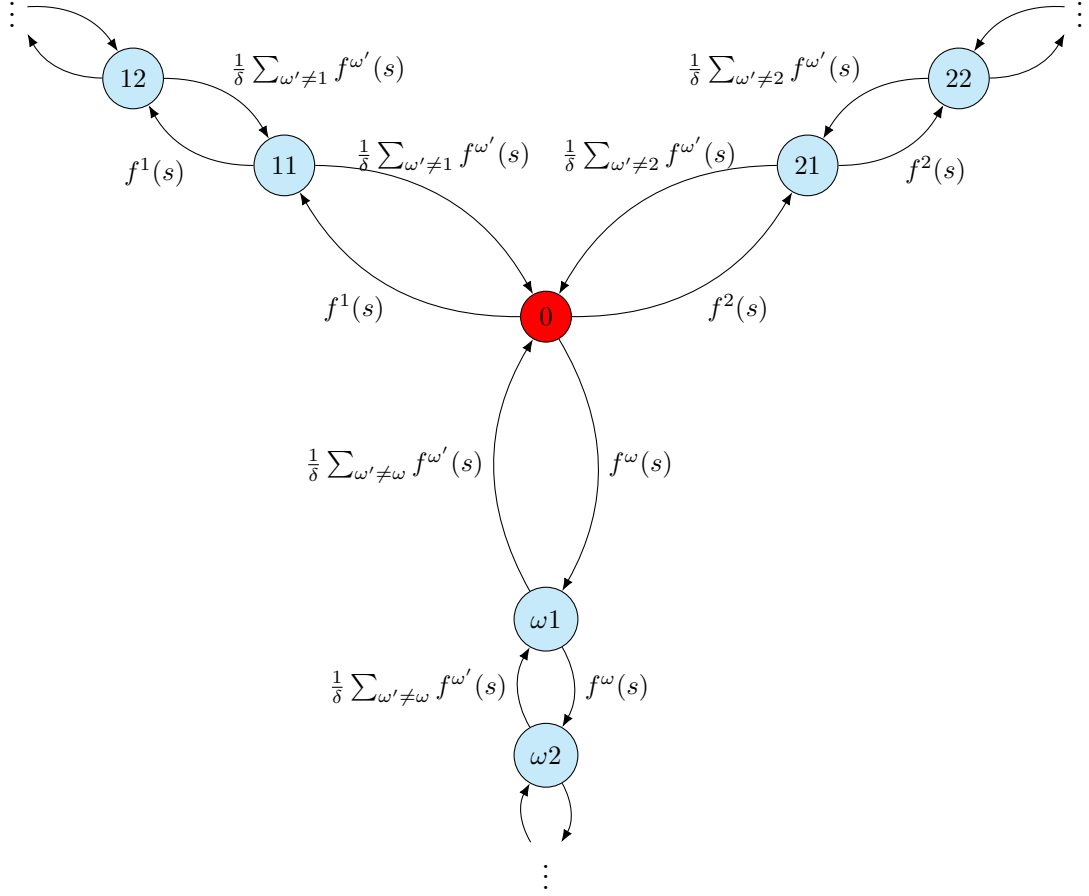


Figure 8: Zoom-in of Figure 7, showing the transitions between memory state 0 and the first two memory states of three of the N branches. Memory states are labeled ωk (k -th state in branch ω). Upon receiving signal s , the DM in state ωk moves one state further from the center to $\omega(k+1)$ with probability $f^{\omega}(s)$ (a confirmatory signal for state ω), one state closer to the center to $\omega(k-1)$ with probability $\frac{1}{\delta} \sum_{\omega' \neq \omega} f^{\omega'}(s)$ (a disconfirmatory signal, underweighted by $\frac{1}{\delta}$), and stays with complementary probability. At the extreme ($k = \lfloor \frac{M-1}{N} \rfloor$), the DM stays instead of moving further; at the bottom ($\omega 1$), a disconfirmatory signal sends the DM back to the central state 0.

nism processes signals through the stochastic confirmatory-signal labeling and only tracks a confidence level for one state at a time. Its memory requirement for achieving a given level of accuracy depends on N (the number of states) but not directly on $|S|$, and thus converges to Bayesian in a faster rate. This distinction is important because the signal space can be much larger than the state space — for instance, when the DM observes high-dimensional data but the underlying uncertainty is low-dimensional. The simple mechanism’s independence from $|S|$ demonstrates that efficient learning under bounded memory requires compressing signals into decision-relevant statistics rather than directly storing observations.

Third, as shown in Proposition 1(ii), the simple updating mechanism described in this section approximates perfect learning for all N and $(\mathbf{u}^\omega, p^\omega)_{\omega=1}^N$ in small worlds when M is large. Therefore, no knowledge of prior belief $(p^\omega)_{\omega=1}^N$ and utility matrices $(\mathbf{u}^\omega)_{\omega=1}^N$ is required. This is particularly consistent with the result that the distribution of actions is close to Bayesian in small worlds, as the simple updating mechanism is parsimonious and easy to implement.

Fourth, perfect learning with the simple updating mechanism in very small worlds is robust to “implementation errors”, as shown in the online Appendix D. Roughly speaking, I assume that the DM mistakenly transits to a neighboring memory state with some probability γ in each period regardless of the signal realization s . Such local mistakes could be induced by mistakes in the perception of signals or imperfect tracking (local fluctuation) of memory states. Online Appendix D shows that Proposition 1(ii) holds for all $\gamma \in [0, 1)$, further strengthening the result of perfect learning in very small worlds.

Lastly, the Proposition is a comparative statics w.r.t. M . Therefore, it remains to see whether the ratio N/M is important. However, as N changes, the decision problem would be highly sensitive to the parameters related to the newly introduced states. For example, adding a state with p^ω close to 0 would be very different from adding a state with a larger prior likelihood. Similarly, adding a state where there are signals easily ruling out that state relative to other states would be different from the situation where signals are close to uninformative. In the following, to side-step these issues, I analyze the limit where N, M both go to ∞ and show that the limit payoff and distributional results depend on the ratio $\frac{N}{M}$ at the limit.

5.1.1 $N, M \rightarrow \infty$: Assumptions and main result

In this subsection, I present the analysis for $N, M \rightarrow \infty$, showing that the qualitative distinction between small and big worlds is preserved at the limit. Formally, I consider a sequence of inference problems $\left\{ (\mathbf{u}_{N(k)}^\omega, p_{N(k)}^\omega, f_{N(k)}^\omega)_{\omega=1}^{N(k)}, M(k) \right\}_{k=1}^\infty$ where $\lim_{k \rightarrow \infty} N(k), M(k) \rightarrow \infty$ and $N(k) = \Theta(M(k)^h)$. With $h < 1$, $\frac{N(k)}{M(k)} \rightarrow 0$, corresponding to small worlds; with $h \geq 1$, $\lim_{k \rightarrow \infty} \frac{N(k)}{M(k)} > 0$, corresponding to big worlds. I impose the following assumptions on the limit environment.

Assumption (Limit-FS). *For any sequence of subsets of states $\Omega_{N(k)} \subseteq \{1, \dots, N(k)\}$ with*

cardinality $|\Omega_{N(k)}|$ and a well-defined limit $\lim_{N(k) \rightarrow \infty} \frac{|\Omega_{N(k)}|}{N(k)}$,

$$\lim_{k \rightarrow \infty} \sum_{\omega \in \Omega_{N(k)}} p_{N(k)}^\omega > 0 \quad \text{if} \quad \lim_{k \rightarrow \infty} \frac{|\Omega_{N(k)}|}{N(k)} > 0. \quad (7)$$

That is, any sequence of subsets of states that constitutes a non-negligible fraction of the state space at the limit must also carry non-negligible prior probability. This is the natural analogue of the full-support assumption $p^\omega > 0$ in the baseline model. Define

$$\bar{p}_0 := \sup \left\{ \lim_{k \rightarrow \infty} \sum_{\omega \in \Omega_k} p_{N(k)}^\omega : \Omega_{N(k)} \subseteq \{1, \dots, N(k)\}, \lim_{k \rightarrow \infty} \frac{|\Omega_{N(k)}|}{M(k)} = 0 \right\}, \quad (8)$$

the maximum prior probability that can concentrate on a set of states whose cardinality is vanishing relative to the memory size $M(k)$. Note that when $h < 1$ (small worlds), $M(k)$ grows faster than $N(k)$, so even $\Omega_{N(k)} = \{1, \dots, N(k)\}$ satisfies $|\Omega_{N(k)}|/M(k) \rightarrow 0$; hence $\bar{p}_0 = 1$. When $h \geq 1$ (big worlds), $M(k) \leq O(N(k))$, so the constraint $|\Omega_{N(k)}|/M(k) \rightarrow 0$ is binding and $\bar{p}_0 < 1$ in general.

A stronger version of [\(Limit-FS\)](#) replaces the “if” with “if and only if”:

Assumption (Limit-FS⁺). *For any sequence of subsets of states $\Omega_N \subseteq \{1, \dots, N\}$ with cardinality $|\Omega_{N(k)}|$ and a well-defined limit $\lim_{k \rightarrow \infty} \frac{|\Omega_{N(k)}|}{N(k)}$,*

$$\lim_{k \rightarrow \infty} \sum_{\omega \in \Omega_{N(k)}} p_{N(k)}^\omega > 0 \quad \text{if and only if} \quad \lim_{k \rightarrow \infty} \frac{|\Omega_{N(k)}|}{N(k)} > 0. \quad (9)$$

Assumption [\(Limit-FS⁺\)](#) strengthens [\(Limit-FS\)](#) by additionally requiring that any sequence of subsets constituting a *negligible* fraction of the state space must also carry negligible prior probability. This implies $\bar{p}_0 = 0$ in big worlds. In other words, at the limit, no state is infinitely more likely than any other state: prior probability cannot concentrate on a vanishing fraction of the state space. I also impose the following two assumptions on the information structures.

Assumption (Limit-NRO). *There exists $\varsigma > 0$ such that $\inf_{s \in S} \frac{f_N^\omega(s)}{f_N^{\omega'}(s)} > \varsigma$ for all $\omega, \omega' \in \Omega$ and for all N .*

Assumption (Limit-ID). *For all $\varepsilon > 0$, there exist $\xi > 0$ and a sequence of subsets $N_\xi \subseteq \{1, \dots, N\}$ such that $\lim_{N \rightarrow \infty} \sum_{\omega \in N_\xi} p_N^\omega > 1 - \varepsilon$ and*

$$\lim_{N \rightarrow \infty} \inf_{\omega, \omega' \in N_\xi; \omega' \neq \omega} \left\{ -\log \frac{\int f_N^\omega(s) f_N^{\omega'}(s) ds}{\sqrt{\int (f_N^\omega(s))^2 ds} \sqrt{\int (f_N^{\omega'}(s))^2 ds}} \right\} > \xi.$$

Assumption [\(Limit-ID\)](#) ensures that almost all states remain identifiable at the limit via their Cauchy–Schwarz distance, analogous to the assumption $f^\omega \neq f^{\omega'}$ in the baseline model.

For ease of exposition, I also assume $u_{N(k)}(\omega, \omega) = 1$ and $u_{N(k)}(\omega', \omega) = 0$ for all $\omega \neq \omega'$ and all k .¹⁸ Denote L_k^* as the infimum asymptotic utility loss with $N(k)$ states and $M(k)$ memory states.

Proposition 2. *Fix a sequence $\left\{(\mathbf{u}_{N(k)}^\omega, p_{N(k)}^\omega, f_{N(k)}^\omega)_{\omega=1}^{N(k)}, M(k)\right\}_{k=1}^\infty$ satisfying Assumptions [\(Limit-FS\)](#), [\(Limit-NRO\)](#), and [\(Limit-ID\)](#), where $N(k) = \Theta(M(k)^h)$. Then*

$$\lim_{k \rightarrow \infty} L_k^* = \begin{cases} 0 & \text{if } h < 1; \\ 1 - \bar{p}_0 & \text{if } h \geq 1. \end{cases}$$

Additionally, if Assumption [\(Limit-FS⁺\)](#) holds, $\lim_{k \rightarrow \infty} L_k^* = 1$ when $h \geq 1$; Moreover, for large enough k , L_k^* increases in k if $h \geq 1$ and decreases in k if $h < 1$.

The threshold $h = 1$ has a clean economic interpretation: it separates environments in which cognitive resources scale faster than complexity ($h < 1$) from those in which they scale at most as fast ($h \geq 1$). When $h < 1$ (small worlds), the proof constructs the same simple updating mechanism as in Proposition 1: as $\frac{M}{N} \rightarrow \infty$, the DM can allocate arbitrarily many memory states to each action, and learning approaches the Bayesian benchmark. When $h \geq 1$ (big worlds), the DM can attend to at most a vanishing fraction of states, so the utility loss is bounded below by the prior probability of the states he ignores. Under [\(Limit-FS\)](#) alone, the prior can put non-vanishing mass on a sequence of subsets Ω_k with $|\Omega_k|/M(k) \rightarrow 0$ — i.e., $\bar{p}_0$ can be strictly positive. In this case, the DM optimally focuses his memory on such a subset and recovers utility close to $\bar{p}_0$, yielding a limit loss of $1 - \bar{p}_0 < 1$. Under [\(Limit-FS⁺\)](#), in contrast, every subset with vanishing fraction in $N(k)$ — and *a fortiori* every subset with vanishing fraction in $M(k)$, since $M(k) \leq O(N(k))$ in big worlds — carries vanishing prior. Hence $\bar{p}_0 = 0$ and the loss converges to 1: the DM learns essentially nothing at the limit.

The monotonicity statement at the end of the proposition reflects the same dichotomy at the level of comparative statics. In small worlds, L_k^* decreases in k : scaling both N and M up along the sequence strictly improves the DM's optimal performance, since the additional memory more than compensates for the additional states. In big worlds, L_k^* increases in k : the DM's effective per-state memory shrinks as k grows, and the learning problem becomes strictly harder despite the larger absolute resources.

5.2 Is ignorance optimal?

5.2.1 General Environments

In this subsection, I present the main results on whether ignorant behavior is optimal in small and big worlds.

¹⁸Online Appendix C presents a information structure that satisfies Assumption [\(Limit-NRO\)](#) and [\(Limit-ID\)](#).

Definition 1. An updating mechanism $(m_1, \mathcal{T}, d)$ ignores state ω if there does not exist any memory state m such that $d(m) = \omega$. An updating mechanism is ignorant if it ignores some state, and non-ignorant otherwise.

As argued in the introduction, the structure of optimal ignorance is reminiscent of a large set of behavioral biases that depart from the Bayesian model. Note that given the assumption that information strictly improves utility, i.e., $L_M^* < \min_a \sum_{\omega=1}^N p^\omega [u(\omega, \omega) - u(a, \omega)]$, when $N = 2$, no ignorant updating mechanism is ϵ -optimal when ϵ is sufficiently small. I present the formal results on ignorance in the following Proposition when $N > 2$, and offer interpretations in the next paragraph.

Corollary 1. Ignorance could be optimal in big worlds but not in small worlds:

(i) For all $N > 2$ and M , there exists some environment $(\mathbf{u}^{\omega'}, p^{\omega'}, f^{\omega'})_{\omega'=1}^N$ and a state ω such that:

(a) every mechanism with optimal decision rule, i.e.,

$$d(m) \in \arg \max_a \sum_{\omega''} p^{\omega''} \mu_m^{\omega''} u(a, \omega'') \text{ for all } m,$$

ignores state ω ; and

(b) more generally, without any restriction on the decision rule, every ϵ -optimal mechanism picks action ω in any state ω' with probability at most $O(\epsilon)$, which vanishes as $\epsilon \rightarrow 0$.

(ii) In contrast, fix $N > 2$ and some $(\mathbf{u}^{\omega'}, p^{\omega'}, f^{\omega'})_{\omega'=1}^N$; when M is large enough, for all ϵ smaller than some threshold $\bar{\epsilon}$, every ϵ -optimal updating mechanism ignores no state.

First, (i) shows that generally, there exist decision environments where ignorance is optimal in two related senses. By (i)(a), every mechanism with the optimal decision rule strictly ignores some state ω ; by (i)(b), every ϵ -optimal mechanism (without restriction on the decision rule) picks ω with vanishing probability as $\epsilon \rightarrow 0$. As $\frac{M}{N}$ is finite, the DM cannot allocate infinite cognitive resources to every state of the world. The DM is bound to make mistakes and faces trade-offs between the probability of mistakes in different states of the world. (i) implies that there exist model parameters such that it is optimal for the DM to ignore some states altogether to improve learning in other states. On the other hand, (ii) shows that the set of decision environments where ignorance is optimal vanishes as M grows large, or equivalently as the world gets smaller. As M increases, the DM can allocate many more memory states to each action. The trade-off between learning in different states is less important. In particular, when $\frac{N}{M}$ converges to 0, the DM learns almost perfectly for all states of the world, and has no incentive to ignore any of the states. Corollary 1 also implies that when $\frac{N}{M}$ is small, the DM will not choose actions that are “safe” but are sub-optimal in every state of the world.

A natural follow-up question is: *which* states does the DM ignore? The following Proposition characterizes three properties of states that make them candidates for optimal ignorance.

Proposition 3. *Fix any $N > 2$, M , and $(f^\omega)_{\omega=1}^N$, and suppose $u(\omega', \omega) = 0$ for all $\omega' \neq \omega$. There exist thresholds $\xi_p, \xi_u > 0$ and $\bar{F} > 1$ — depending only on the model primitives — such that if*

- (i) $p^\omega < \xi_p$, i.e., state ω is a priori unlikely, or
- (ii) $\frac{u(\omega, \omega)}{\min_{\omega'' \neq \omega} u(\omega'', \omega'')} < \xi_u$, i.e., guessing state ω correctly is unimportant, or
- (iii) $\sup_s \frac{f^\omega(s)}{f^{\omega'}(s)} \times \sup_s \frac{f^{\omega'}(s)}{f^\omega(s)} \leq \bar{F}$ for some $\omega' \neq \omega$, i.e., distinguishing states ω and ω' is difficult,

then:

- (a) every mechanism with optimal decision rule, i.e.,

$$d(m) \in \arg \max_a \sum_{\omega''} p^{\omega''} \mu_m^{\omega''} u(a, \omega'') \text{ for all } m,$$

ignores state ω in cases (i) and (ii), and ignores at least one of ω and ω' in case (iii);

- (b) more generally, without any restriction on the decision rule, every ϵ -optimal mechanism picks action ω in any state ω'' with probability at most $O(\epsilon)$ in cases (i) and (ii), and in case (iii), either action ω or action ω' is picked with probability at most $O(\epsilon)$ in any state ω'' ; these probabilities vanish as $\epsilon \rightarrow 0$.

In summary, the DM ignores states that are a priori unlikely, relatively unimportant, or hard to distinguish from other states. More specifically, (i) shows that when the prior probability of a state is low, the DM would rather ignore that state and allocate his limited memory states to learn other states of the world. For example, an individual who is confident about his ability would not update his belief in his own ability but only in his teammates' ability, resembling persistent overconfidence. In contrast, a Bayesian who observes a series of bad team performances would eventually assign positive probability to the DM's own ability being low. Similarly, (ii) shows that when a state is relatively unimportant, the utility of guessing correctly that state is relatively small, the DM would ignore that state to focus on learning about other states. Lastly, (iii) shows that when it is difficult to distinguish states ω and ω' , the DM ignores one of the two states. That is, the DM ignores states that are difficult to identify. To see the intuition, imagine two states ω and ω' where $f^\omega(s) \approx f^{\omega'}(s)$ for all s , and without loss of generality that state ω is a priori less favorable than state ω' . The DM has to receive a large number of signals that support ω against ω' such that he prefers to take action ω instead of ω' . To use many memory states to record signals supporting ω and still be (almost) unsure about choosing ω over ω' is an inefficient use of memory states. Instead, saving those numerous cognitive resources to learn other states of the world, for example, by

recording signals supporting ω' against some ω'' and vice versa, he can improve his utility in the other states (ω and ω'') and will be better off. Condition (iii) is particularly relevant for behavior that resembles correlation neglect: when it is difficult to distinguish positive or no correlation in the data, and if positive correlation is slightly less likely, the DM ignores the possibility of positive correlation and focuses on learning other states of the world.¹⁹

5.2.2 $N = M$, symmetric environments and symmetric mechanisms

Corollary 1 and Proposition 3 show that ignorance is optimal in some asymmetric environments when $\frac{M}{N}$ is small. In the following, I show that ignorance is also optimal in symmetric environments where learning is “difficult”. It implies that optimal ignorance is not a knife-edge result, but rather a general phenomenon when the individual has limited resources for learning.

I consider the following simple symmetric environment where states and actions are ex-ante identical: There are N possible signal realizations, i.e., $S = \{s^1, \dots, s^N\}$, and p^ω , $u(\omega, \omega)$, $u(\omega', \omega)$, $\frac{f^\omega(s^\omega)}{f^\omega(s^{\omega'})} > 1$ are invariant across all ω and $\omega' \neq \omega$. Moreover, I focus on settings where $M = N$ and is an even number.

As mentioned earlier, it is generally intractable to determine the optimal updating mechanism when $N > 2$. In the following, for simplicity, I restrict attention to the class of *symmetric* mechanisms and ask whether the optimal mechanism *within this class* is ignorant. Formally, denote the set of states that are not ignored by an updating mechanism as $\Omega^{NI} = \{\omega : \exists m \text{ s.t. } d(m) = \omega\}$. An updating mechanism is *symmetric* if for any permutation $\sigma : \Omega^{NI} \rightarrow \Omega^{NI}$, there exists a one-to-one mapping $\sigma^M : M \rightarrow M$ such that for all $\omega \in \Omega^{NI}$,

$$d(\sigma^M(m)) = \sigma(d(m))$$

$$\Pr\{\mathcal{T}(\sigma^M(m), s^{\sigma(\omega)}) = \sigma^M(m')\} = \Pr\{\mathcal{T}(m, s^\omega) = m'\}.$$

That is, after any permutation of the non-ignored states, there exists a relabeling of memory states such that the updating mechanism remains unchanged. Note that when $N = 2$, the optimal updating mechanism is symmetric, so the restriction is without loss of generality in that case.

Proposition 4. *In symmetric environments, consider optimality within the class of symmetric mechanisms. All ϵ -optimal symmetric updating mechanisms (with small enough ϵ)*

¹⁹For example, consider an investor who wants to learn the state of the stock market — strong, moderately strong, moderately weak, or weak — and whether analyst forecasts are positively correlated or independent. A signal is a profile of predictions from market analysts. If the predictions are predominantly favorable, this could indicate either that the market is genuinely strong with independent forecasts, or that the market is only moderately strong but the analysts are drawing from the same source and echoing each other (positive correlation). Because these two scenarios produce similar signal profiles, they are difficult to distinguish. According to condition (iii), the DM would optimally ignore one of them — say, the possibility of positive correlation — and focus his limited cognitive resources on learning whether the market is strong, moderately strong, moderately weak, or weak, behaving as if analyst forecasts are always independent.

are ignorant when the state space is large enough or when signals are noisy enough. That is, when N is large enough or $\frac{f^\omega(s^\omega)}{f^\omega(s^{\omega'})}$ (with $\omega \neq \omega'$) is close enough to 1, for all ϵ smaller than some threshold, every ϵ -optimal symmetric mechanism ignores some state.

In the proof, I first show that among all symmetric non-ignorant mechanisms, the one illustrated in Figure 9 is optimal. In this mechanism, the decision rule $d(m)$ is a one-to-one mapping from m to ω , i.e., the DM chooses each action in one and only one unique memory state, and the DM transits to the memory state corresponding to action ω if and only if he receives signal s^ω (the signal that supports state ω). Second, I compare the asymptotic utility of this best symmetric non-ignorant mechanism with that of a simple symmetric ignorant mechanism that ignores half of the states, as illustrated in Figure 10.

The result is illustrated in Figure 11; and Figure 12 shows the comparison of the non-ignorant and ignorant mechanisms where $N = M \in \{4, 6, 8\}$. The y-axis is the asymptotic utility of the updating mechanism that ignores half of the states minus that of the best non-ignorant updating mechanism. Thus, if the y-axis is above 0, all ϵ -optimal mechanisms must be ignorant for small enough ϵ . The x-axis is the informativeness of the signal structure, i.e., $\frac{f^\omega(s^\omega)}{f^\omega(s^{\omega'})}$ where $\omega' \neq \omega$. Both Figure 11 and Figure 12 show that ignorance is optimal when the informativeness is smaller than some threshold, or when the state space N is large enough. In other words, ignorance is optimal when it is difficult to learn, which echoes the results in Corollary 1.²⁰ The result also contrasts with the setting when $N = 2$, in which the optimal updating mechanism is symmetric and non-ignorant in such symmetric environments.

The intuition of the result is as follows. To consider all states, the DM allocates one memory state to each action. It is thus easy for the DM to alternate between different actions and unavoidably make mistakes. Put differently, the updating mechanism is “noisy”. This is especially true when signals are very noisy or when N is large: even when the likelihood ratio $\frac{f^\omega(s^\omega)}{f^\omega(s^{\omega'})}$ is fixed, $f^\omega(s^\omega)$ vanishes as $N \rightarrow \infty$ because probability mass is spread across $N - 1$ other signal realizations. Formally, note that in the symmetric environment,

$$f^\omega(s^\omega) = \frac{f^\omega(s^\omega)/f^\omega(s^{\omega'})}{f^\omega(s^\omega)/f^\omega(s^{\omega'}) + N - 1},$$

which vanishes as N grows. Each signal becomes almost uninformative, yet with only one memory state per action, the DM cannot accumulate enough evidence for any single state.

If, in contrast, the DM ignores half of the actions, he allocates two memory states to each of the actions that he considers and can therefore “records” two consecutive signals supporting the same state. As it is exponentially unlikely to receive consecutive signals supporting the wrong state, the ignorant mechanism improves the individual’s decision-making among the smaller set of states that he considers. When N is large or when $\frac{f^\omega(s^\omega)}{f^\omega(s^{\omega'})}$ is small, the improvement outweighs the loss he incurs among the states that he ignores, because the loss is small to begin with, i.e., the asymptotic utility of a “noisy” updating mechanism

²⁰The magnitude of the thresholds also suggests that ignorance is not an extreme event. For example, when $N = 4$, ignorance is better when the likelihood ratio of the signals is smaller than (around) 3.

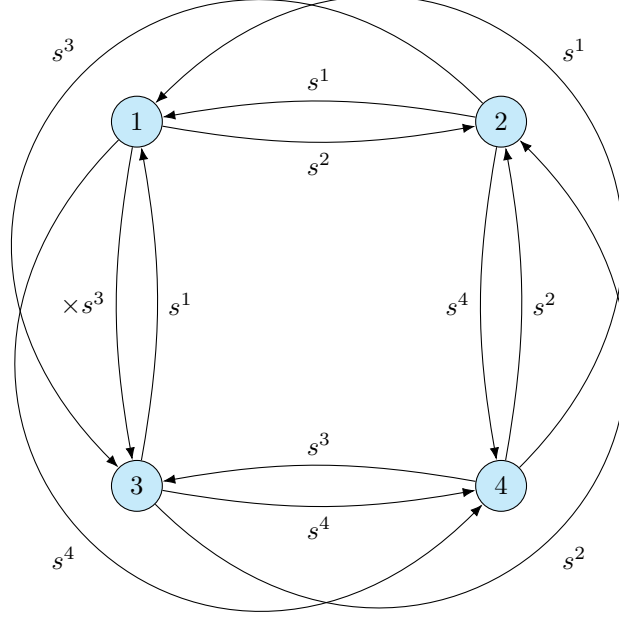


Figure 9: The optimal non-ignorant updating mechanism that considers all states, with $N = M = 4$. The number in the node denotes the action that the DM takes when he is in this memory state. Moreover, in memory state ω , and upon receiving a signal that supports state $\omega' \neq \omega$, the DM transits to memory state ω' with probability 1 and stays in his current memory state otherwise.

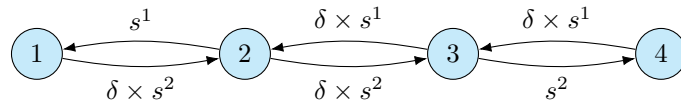


Figure 10: An example of an updating mechanism that ignores two states, with $N = M = 4$. The DM takes action 1 in memory states 1 and 2, and takes action 2 in memory states 3 and 4. In memory state 3, if the DM receives a signal supporting state 2, he transits to memory state 4; if the DM receives a signal supporting state 1, he transits to memory state 2 with probability δ and stays in his current memory state otherwise. In memory state 4, if the DM receives a signal supporting state 1, he transits to memory state 3 with probability δ and stays in his current memory state otherwise. The transition functions in memory state 1 and 2 are defined accordingly. δ is chosen to be close to 0 to maximize the asymptotic utility.

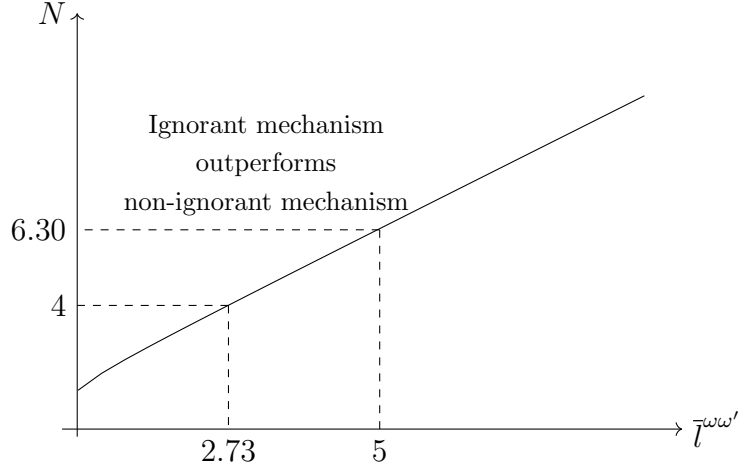


Figure 11: The ignorant mechanism outperforms the non-ignorant mechanism when information is noisy enough or when N is big enough.

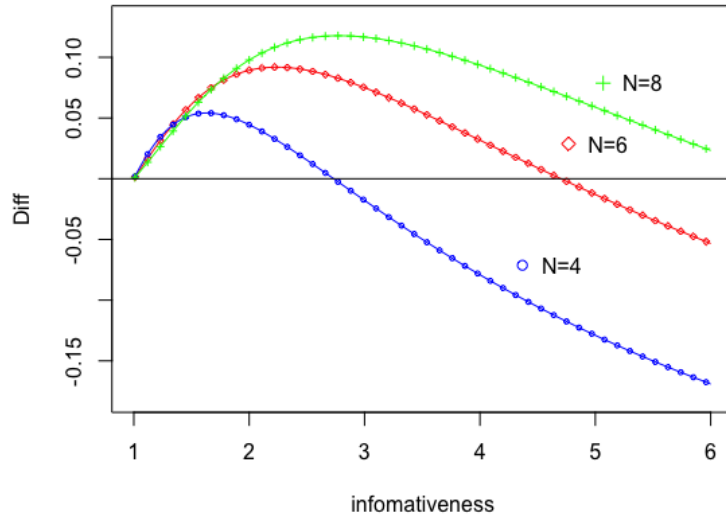


Figure 12: The y-axis is the (estimated) asymptotic utility of the ignorant mechanism illustrated in Figure 10 minus that of the non-ignorant mechanism illustrated in Figure 9, while the x-axis is $\frac{f^\omega(s^\omega)}{f^\omega(s^{\omega'})}$. The ignorant mechanism outperforms the non-ignorant mechanism when $\frac{f^\omega(s^\omega)}{f^\omega(s^{\omega'})}$ is smaller than some threshold. Moreover, the threshold is larger as N increases from 4 to 6, and from 6 to 8.

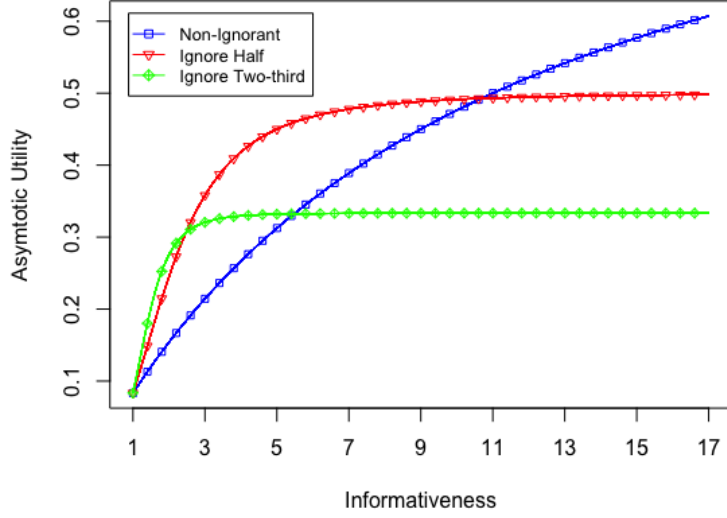


Figure 13: Consider $N = M = 12$. The y-axis is the asymptotic utility of the ignorant mechanism ignoring half of the states illustrated in Figure 10, a symmetric ignorant mechanism that ignores 2/3 of the states, and the non-ignorant mechanism illustrated in Figure 9, while the x-axis is $\frac{f^\omega(s^\omega)}{f^\omega(s^{\omega'})}$. As the informativeness $\frac{f^\omega(s^\omega)}{f^\omega(s^{\omega'})}$ decreases, it becomes more beneficial to ignore more states.

that considers all states is small. In particular, Figure 13 further extends this intuition by analyzing an ignorant mechanism that ignores 2/3 of the states of the world. It shows that when the signals are very noisy, i.e., when $\frac{f^\omega(s^\omega)}{f^\omega(s^{\omega'})}$ is close to 1, ignoring 2/3 of the states of the world outperforms the mechanism that ignores half of the states, which in turns outperforms the non-ignorant mechanism.

5.2.3 $N, M \rightarrow \infty$

The results of Section 5.2.1 and 5.2.2 are results for fixed finite N . The following results show that their qualitative conclusions extend to the limit $N, M \rightarrow \infty$, further confirming that the ratio $\frac{N}{M}$ is the key determinant of ignorance.

When $M \rightarrow \infty$, the finite-case definition of ignorance (Definition 1) is no longer adequate: a mechanism could assign a single memory state to action ω without affecting long-run utility, since the cost of “wasting” one memory state out of infinitely many vanishes in the limit. Accordingly, I adopt the following long-run definition.

Definition 2. A sequence of updating mechanisms $\{m_{1k}, \mathcal{T}_k, d_k\}_{k=1}^\infty$ ignores state ω at the limit if, for every state ω' ,

$$\lim_{k \rightarrow \infty} \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \Pr^{\omega'}(a_t = \omega) = 0,$$

where $\Pr^{\omega'}$ denotes the probability when the true state is ω' . A sequence is ignorant at the limit if it ignores some state at the limit, and non-ignorant at the limit otherwise.

That is, a sequence of mechanisms ignores state ω at the limit if the long-run fraction of choosing action ω converges to zero regardless of the true state. Definition 2 thus replaces the previous definition of ignorance — “state ω is not in the range of d ” — with one defined in terms of observable actions: “action ω is essentially never chosen, regardless of the true state.” This definition has two advantages. First, it is robust to redundant memory allocation that does not affect the actions taken — a key concern when $M \rightarrow \infty$. Second, it is directly observable from the DM’s chosen actions, paralleling the action-frequency formulation of optimal ignorance in Corollary 1(i)(b).

In small worlds ($h < 1$), Proposition 2 established that $L_k^* \rightarrow 0$. The following corollary shows that non-ignorance at the limit (Definition 2) follows immediately:

Corollary 2. *Under the same conditions as Proposition 2 with $h < 1$, a sequence of updating mechanisms $\{m_{1k}, \mathcal{T}_k, d_k\}_{k=1}^\infty$ that is ϵ -optimal at the limit $k \rightarrow \infty$ ignores a set of states with total prior probability at most ϵ at the limit, which vanishes as $\epsilon \rightarrow 0$.*

For big worlds ($h \geq 1$), the following result is considerably stronger than Corollary 1: it holds for *all* environments (not just some), and shows that any updating mechanism must ignore almost all states at the limit in the sense of Definition 2.

Proposition 5. *Fix a sequence $\left\{(\mathbf{u}_{N(k)}^\omega, p_{N(k)}^\omega, f_{N(k)}^\omega)_{\omega=1}^{N(k)}, M(k)\right\}_{k=1}^\infty$ satisfying Assumptions (Limit-FS), (Limit-NRO), and (Limit-ID), where $N(k) = \Theta(M(k)^h)$ with $h \geq 1$. Then for any sequence of updating mechanisms $\{m_{1k}, \mathcal{T}_k, d_k\}_{k=1}^\infty$,*

$$\lim_{k \rightarrow \infty} \frac{\left| \left\{ \omega : \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \Pr^{\omega'}(a_t = \omega) > 0 \text{ for some } \omega' \right\} \right|}{M(k)} = 0.$$

That is, the set of actions taken with strictly positive long-run probability has cardinality $o(M(k))$ — vanishing relative to the memory size and, a fortiori, a vanishing fraction of the $N(k)$ states.

Proposition 5 provides the structural reason behind the big-world part of Proposition 2: by the memory constraint alone, the DM is forced to focus on a vanishing fraction of states regardless of the environment, and therefore captures at most the $\bar{p}_0$ -measure of states he attends to — yielding the loss $1 - \bar{p}_0$. The result sharpens the contrast between small and big worlds: in the former the DM attends to essentially all states, while in the latter he attends to vanishingly few. Under (Limit-FS⁺), $\bar{p}_0 = 0$, so the DM attends to a set of states with vanishing prior probability and the loss converges to 1.

The intuition is that to ever choose action ω in the long run, the DM must accumulate enough evidence to make ω stand out among the alternatives, which requires devoting a chain of memory states to tracking signals favoring ω . As we have established in Section 5.2.2, when

$N(k)$ increases, signals favoring ω become less likely even when state ω is true, since there are many alternative states. Tracking signals for many states simultaneously consumes memory roughly linearly in the number of states. With memory $M(k)$ scaling no faster than $N(k)$ in big worlds, the DM's budget cannot accommodate tracking for more than a vanishing fraction of states, forcing him to ignore almost all states regardless of the prior.

5.3 Is disagreement persistent?

Lastly, I turn to the question of whether disagreement could persist asymptotically in small and big worlds. Consider two individuals A and B who have different utility matrices $(\mathbf{u}_A^\omega)_{\omega=1}^N$ and $(\mathbf{u}_B^\omega)_{\omega=1}^N$, and/or different prior beliefs $(p_A^\omega)_{\omega=1}^N$ and $(p_B^\omega)_{\omega=1}^N$, and/or different abilities of information acquisition captured by $(f_A^\omega)_{\omega=1}^N$ and $(f_B^\omega)_{\omega=1}^N$.²¹ Their updating mechanisms, $(m_{A1}, \mathcal{T}_A, d_A)$ and $(m_{B1}, \mathcal{T}_B, d_B)$, induce a (random) sequence of actions over time. To define disagreement, I sample one action from each individual A and B and define the “disagreement between $(m_{A1}, \mathcal{T}_A, d_A)$ and $(m_{B1}, \mathcal{T}_B, d_B)$ in state ω ” as the probability that the two sampled actions are different.²² For a given ϵ , the disagreement between individual A and B in state ω is the supremum of disagreement between all pairs of ϵ -optimal updating mechanisms in state ω . Lastly, the two individuals almost always disagree (resp. agree) with each other for a given ϵ if their disagreement equals 1 (resp. 0) in all states.

Corollary 3. *In big worlds, disagreement can persist; in small worlds, individuals cannot disagree forever.*

- (i) *For all $N > 2$, M_A and M_B , there exists some $(\mathbf{u}_A^\omega, p_A^\omega, f_A^\omega)_{\omega=1}^N$ and $(\mathbf{u}_B^\omega, p_B^\omega, f_B^\omega)_{\omega=1}^N$ such that individuals A and B almost always disagree when both adopt ϵ -optimal updating mechanisms, when $\epsilon \rightarrow 0$.*
- (ii) *In contrast, take $(\mathbf{u}_A^\omega, p_A^\omega, f_A^\omega)_{\omega=1}^N$ and $(\mathbf{u}_B^\omega, p_B^\omega, f_B^\omega)_{\omega=1}^N$, there exist $K_1, K_2 > 0$ and $r < 1$ such that when both individuals adopt ϵ -optimal updating mechanisms, their disagreement is bounded above by $\left(K_1 r^{\lfloor \frac{\min\{M_A, M_B\}-1}{N} \rfloor} + K_2 \epsilon\right)$ in all states.*

(i) shows that in general, there exist decision environments such that disagreement almost always happens and is persistent. However, as implied by (ii), when the world gets smaller ($\frac{N}{M}$ goes to 0), such set of decision environments with disagreement vanishes: different individuals with different prior beliefs and/or information acquisition abilities who adopt (almost) optimal updating mechanisms are bound to agree with each other.

²¹For example, individual A could receive noisier signals than individual B , i.e., $f_A^\omega = \gamma + (1 - \gamma)f_B^\omega$ for some $\gamma \in (0, 1)$; or individual A could have different learning advantages in identifying some states better but other states worse than individual B , i.e., $\sup_s f_A^\omega(s)/f_A^{\omega'}(s) > \sup_s f_B^\omega(s)/f_B^{\omega'}(s)$ but $\sup_s f_A^{\omega''}(s)/f_A^{\omega'''}(s) < \sup_s f_B^{\omega''}(s)/f_B^{\omega'''}(s)$ for some $\omega, \omega', \omega'', \omega'''$.

²²For example, if both individuals alternate between action 1 and 2, their disagreement is $\frac{1}{2}$. An alternative definition is to measure the proportion of time t where $a_t^A \neq a_t^B$: If both individuals alternate between action 1 and 2, their disagreement is 0 if they both start with the same action and is 1 otherwise. Corollary 3 (i) and the limit result of (ii), i.e., disagreement approaches 0 as M goes to infinity, hold with this alternative definition.

The intuition of the result is as follows: as different individuals with different prior beliefs and utilities would adopt different updating mechanisms, they could focus their learning on different subsets of states of the world when they have limited memory. In particular, consider an example with $N = 4$. If individual A ignores states 1 and 2 and individual B ignores states 3 and 4, they will never choose the same action and thus disagree with certainty.²³ However, as $\frac{N}{M}$ goes to 0, decisions get closer to Bayesian, and the two individuals almost always choose the same actions.

The following result further shows that disagreement could be driven solely by differences in cognitive ability. For simplicity, I allow M to be smaller than N in the following result.

Corollary 4. *Disagreement can be driven solely by differences in cognitive ability even when individuals have the same prior beliefs and preferences. There exists $M_A \neq M_B$ and $(\mathbf{u}_A^\omega, p_A^\omega, f_A^\omega)_{\omega=1}^N = (\mathbf{u}_B^\omega, p_B^\omega, f_B^\omega)_{\omega=1}^N$ such that individual A and B almost always disagree when both adopt ϵ -optimal updating mechanisms as $\epsilon \rightarrow 0$.*

I prove the Corollary with an example where $M_A = 1 < M_B = 2$, $N = 3$, and $u(\omega, \omega) = 1 > u(\omega, \omega') = 0$ for all ω and $\omega' \neq \omega$. Importantly, state 1 is a priori more likely than state 2 and 3, while it is easier to distinguish state 2 and 3 than to distinguish state 1 and state 2, or state 1 and state 3.²⁴ In this example, individual A always chooses action 1 as he does not have sufficient cognitive resources to learn. On the other hand, if p^1 is not too large compared to p^2 and p^3 , individual B ignores state 1. He takes advantage of the informative information structure to distinguish state 2 and 3, such that he can be confident that he doesn't take action 2 in state 3 and vice versa. Thus, as long as state 1 is not a priori too likely, ignoring it would yield a higher utility.²⁵ To sum up, as individual A has a lower cognitive ability, his learning and actions would depend a lot on the prior belief, while individual B has the ability to take advantage of the information structure.²⁶ As a result, they adopt different updating mechanisms, which leads to disagreement.

5.3.1 $N, M \rightarrow \infty$

The following results extend Corollary 3 to the limit $N, M \rightarrow \infty$. In small worlds, Proposition 2 implies that individuals are bound to agree:

²³Note that the result of disagreement continues to hold in a framework where individuals observe each other's actions. More specifically, one could re-define the signal structures in the current setting to incorporate the information conveyed by the actions taken by the two individuals.

²⁴Mathematically, $\sup_s \frac{f^3(s)}{f^2(s)} \sup_s \frac{f^2(s)}{f^3(s)} > \sup_s \frac{f^1(s)}{f^2(s)} \sup_s \frac{f^2(s)}{f^1(s)}$.

²⁵One may argue that after seeing individual B choosing action 2 or 3, individual A should change his action. However, this is not possible, as he has only one unit of memory capacity ($M = 1$) and thus has to effectively commit to one action. In particular, one can generalize this framework so that the two individuals also see each other's actions as signals, and Proposition 4 would still hold.

²⁶Note that although this example imposes strong assumptions in particular on the size of the automaton of individual A , it generates a strong form of disagreement in which the two individuals disagree asymptotically with certainty. Similar intuition implies that even when the assumption is relaxed, the difference in M would lead to asymptotic disagreement at least probabilistically.

Corollary 5. *Under the same conditions as Proposition 2 for individuals $i = A, B$ (each with their own sequence $M_i(k)$ satisfying $h_i < 1$), suppose the two individuals have the same limit support, i.e., $\lim_{k \rightarrow \infty} p_{N(k)A}^\omega > 0$ if and only if $\lim_{k \rightarrow \infty} p_{N(k)B}^\omega > 0$ for all ω . Then individuals A and B must almost always agree with each other (measured in prior probability $p_{N(k)A}^\omega$ or $p_{N(k)B}^\omega$) at the limit as $\epsilon \rightarrow 0$.*

The intuition is direct. In a small world, the optimal loss vanishes ($L_k^* \rightarrow 0$), so each individual's ϵ -optimal mechanism picks the correct action in (almost measured in prior probability) every true state at the limit. The same-limit-support condition ensures that both individuals consider the same set of states as a priori plausible: without it, an individual with vanishing prior on a state would not bother to learn that state (the contribution to loss is also vanishing) and would almost never play the corresponding action, while the other individual (with non-vanishing prior) would, generating trivial disagreement. With shared support, both individuals optimally pick the same correct action in every true state, and disagreement vanishes. The corollary thus establishes the substantive content of the small-world result: persistent disagreement does not occur when complexity is low compared to cognitive ability.

In big worlds, however, disagreement is possible.

Corollary 6. *Fix sequences $\left\{(\mathbf{u}_{N(k)i}^\omega, p_{N(k)i}^\omega, f_{N(k)i}^\omega)_{\omega=1}^{N(k)}, M_i(k)\right\}_{k=1}^\infty$ for $i = A, B$, satisfying Assumptions (Limit-FS), (Limit-NRO), and (Limit-ID), where $N(k) = \Theta(M_i(k)^{h_i})$ with $h_i \geq 1$. Suppose the two individuals have the same limit support, i.e., $\lim_{k \rightarrow \infty} p_{N(k)A}^\omega > 0$ if and only if $\lim_{k \rightarrow \infty} p_{N(k)B}^\omega > 0$ for all ω . There exist environments $\left\{(\mathbf{u}_{N(k)i}^\omega, p_{N(k)i}^\omega, f_{N(k)i}^\omega)_{\omega=1}^{N(k)}\right\}_{k=1}^\infty$ for $i = A, B$ such that for all $\epsilon > 0$, there exist sequences of ϵ -optimal updating mechanisms for individuals A and B under which they disagree in all states.*

The intuition is as follows. By Proposition 5, each individual in a big world is forced to focus on a vanishing fraction of states — almost all states are ignored. The driving force of the result is most starkly visible under (Limit-FS⁺), where any updating mechanism is ϵ -optimal in the limit. This freedom explains why Corollary 6 is slightly weaker than Corollary 3(i), in that it asserts *some* (rather than all) ϵ -optimal mechanisms lead to disagreement: a large set of mechanisms are ϵ -optimal in big worlds, leaving ample room for divergent sets of actions chosen by the individuals. Nonetheless, combining Corollaries 5 and 6 confirms the same qualitative picture as in the baseline model with fixed N : disagreement does not arise in small worlds but can arise in big worlds.

6 Discussion and Conclusion

6.1 Research Implications

Heterogeneity of learning and heuristics under different environments or across different individuals This paper shows that optimal ignorance under limited cognitive

resources produces behavior that is observationally indistinguishable from a wide range of behavioral anomalies, through a single mechanism: efficient allocation of limited cognitive resources in light of complexity. Importantly, the comparison of small and big worlds illustrates a link between the degree of (relative) complexity of the inference problem and the aforementioned non-Bayesian learning behaviors, which is supported by the experimental results in [Enke and Zimmermann \(2019\)](#) and [Graeber \(2023\)](#). [Enke and Zimmermann \(2019\)](#) shows that correlation neglect negatively correlates with the cognitive ability of subjects and “an extreme reduction in the environment’s complexity eliminates the bias”, while [Graeber \(2023\)](#) shows that a reduction in the complexity of the problem by removing a deciphering stage of signals reduces inattentive learning behavior. This paper theoretically demonstrates that a relationship exists between biases, complexity, and cognitive ability, extending to other behavioral biases. It illustrates that heterogeneity across both individual and decision environments leads to varying prominence of behavioral biases.

6.2 Managerial Implications

Encourage Strategic Ignorance and Focused Inquiry One of the important insights of this paper is that ignorance could be an optimal strategy, especially when cognitive resources are limited. Moreover, perhaps counterintuitively, as shown in Proposition 4, strategic ignorance is optimal especially when uncertainty is high or when information is noisy. This result complements a recent study [Yoon et al. \(2021\)](#), which found that consuming excessive news information during the COVID-19 pandemic decreases employees’ goal progress and creativity due to anxiety. The results of this paper demonstrate that even without psychological channels, reacting to excessive information can compromise the quality of inference and decision-making. For example, as Alan Mulally took over as CEO of Ford in 2006, he streamlined Ford’s product lines, prioritized core brands, and sold off premier automobile brands like Jaguar and Land Rover. He inherited 97 models and reduced Ford’s product lineup to as few as 20 models. While Ford might lose out on a potential boom in the premier automobile market, the streamlining of product lines allowed Ford to conserve resources and return to profitability by 2010, outpacing competitors who pursued broader, less focused strategies ([Edmondson and Jung \(2021\)](#)). Alan Mulally was included in the 2009 Time 100 list due to his successful strategies at Ford.²⁷ Similarly, Steve Jobs has emphasized Apple’s corporate culture of “saying no to hundreds of good ideas”.²⁸

Summarizing Information to Save Cognitive Resources and Reduce Ignorance

On the other hand, the results of this paper also shed light on how to discourage ignorance if one aims to do so. Despite the optimality of ignorance, a manager might want to discourage

²⁷See https://content.time.com/time/specials/packages/article/0,28804,1894410_1893837_1894182,00.html.

²⁸See <https://www.forbes.com/sites/bernhardschroeder/2022/10/31/entrepreneurs-and-small-business-owners-steve-jobs-advice-on-saying-no-more-often-to-achieve-your-goals/>.

ignorance in an employee if the employee has a different utility matrix or disagrees on what possibilities (states of the world) are more probable. As shown in Proposition 3 and Proposition 4, individuals tend to ignore possibilities when information is noisy. Thus, conversely, managers can use AI to summarize information before employees process it. Different from a Bayesian setting and models with entropy-based rational inattention, summarizing multiple signals into one helps individuals process information with limited cognitive resources.

Thanks to the recent advancements in LLM technologies, summarizing information has never been easier. For example, Bloomberg integrated BloombergGPT to summarize earnings calls, news, and filings. Analysts using the tool can easily process multiple signals simultaneously, which reduces their cognitive load when processing information and reduces their incentive to ignore possibilities.²⁹ Similarly, in a recent Microsoft study, they reported that users of their Copilot system are more efficient in pulling together information from multiple sources, which reduces employees' incentive to ignore information, as suggested by the results of this paper.³⁰

7 Conclusion

Using finite automata to model belief formation under limited cognitive resources, this paper shows that ignorant behavior could be optimal when the state space relative to the size of the finite automaton is large, but is not optimal otherwise. Ignorant behavior allows individuals to improve decision-making within the states that they consider, and it is particularly beneficial when signals are noisy. The structure of optimal ignorance is reminiscent of several well-documented learning biases, and the results establish a connection between the relative complexity of the inference problem and the scope for non-Bayesian learning behavior. Fully rationalizing specific biases such as overconfidence or correlation neglect within this framework remains an important direction for future work.

²⁹See Wu et al. (2023) and <https://www.forbes.com/sites/jamielsheikh/2023/04/05/the-chatgpt-of-finance-is-here-bloomberg-is-combining-ai-and-fintech/>.

³⁰See <https://www.microsoft.com/en-us/worklab/work-trend-index/copilots-earliest-users-t-each-us-about-generative-ai-at-work>.

References

- Acemoglu, Daron, Victor Chernozhukov, and Muhamet Yildiz (2016) “Fragility of asymptotic agreement under Bayesian learning,” *Theoretical Economics*, 11 (1), 187–225.
- Banovetz, James and Ryan Oprea (2023) “Complexity and procedural choice,” *American Economic Journal: Microeconomics*, 15 (2), 384–413.
- Basu, Pathikrit and Kalyan Chatterjee (2015) “On interim rationality, belief formation and learning in decision problems with bounded memory,” *Working Paper*.
- Berk, Robert H (1966) “Limiting behavior of posterior distributions when the model is incorrect,” *The Annals of Mathematical Statistics*, 37 (1), 51–58.
- Börgers, Tilman and Antonio J Morales (2004) “Complexity constraints and adaptive learning: An example,” in *Advances in Understanding Strategic Behaviour*, 9–27: Springer.
- Caplin, Andrew, Mark Dean, and John Leahy (2019) “Rational inattention, optimal consideration sets, and stochastic choice,” *The Review of Economic Studies*, 86 (3), 1061–1094.
- Chatterjee, Kalyan and Tai-Wei Hu (2023) “Learning with limited memory: Bayesianism vs heuristics,” *Journal of Economic Theory*, 209, 105642.
- Chatterjee, Kalyan and Hamid Sabourian (2020) “Game theory and strategic complexity,” *Complex Social and Behavioral Systems: Game Theory and Agent-Based Models*, 639–658.
- Chauvin, Kyle P (2019) “Euclidean properties of Bayesian updating,” *Working paper*.
- Compte, Olivier and Andrew Postlewaite (2012) “Belief formation,” *Unpublished manuscript, University of Pennsylvania*.
- Cover, Thomas M (1969) “Hypothesis testing with finite statistics,” *The Annals of Mathematical Statistics*, 40 (3), 828–835.
- Doob, J. L. (1949) “Application of the theory of martingales,” *Actes du Colloque International Le Calcul des Probabilités et ses applications*, (Lyon, 28 Juin – 3 Juillet, 1948), pages 23–27.
- Edmondson, Amy C and Olivia Jung (2021) “The Turnaround at Ford Motor Company,” *Harvard Business School Case*, 621, 101.
- Enke, Benjamin and Florian Zimmermann (2019) “Correlation neglect in belief formation,” *The Review of Economic Studies*, 86 (1), 313–332.
- Freedman, David A (1963) “On the asymptotic behavior of Bayes’ estimates in the discrete case I,” *The Annals of Mathematical Statistics*, 34 (4), 1386–1403.

- (1965) “On the asymptotic behavior of Bayes estimates in the discrete case II,” *The Annals of Mathematical Statistics*, 36 (2), 454–456.
- Gaines, Brian J, James H Kuklinski, Paul J Quirk, Buddy Peyton, and Jay Verkuilen (2007) “Same facts, different interpretations: Partisan motivation and opinion on Iraq,” *The Journal of Politics*, 69 (4), 957–974.
- Gilboa, Itzhak, Larry Samuelson, and David Schmeidler (2022) “Learning (to disagree?) in large worlds,” *Journal of Economic Theory*, 199, 105166.
- Graeber, Thomas (2023) “Inattentive inference,” *Journal of the European Economic Association*, 21 (2), 560–592.
- Hardy, Godfrey Harold (1949) *Divergent Series*: Oxford University Press.
- Heidhues, Paul, Botond Köszegi, and Philipp Strack (2018) “Unrealistic expectations and misguided learning,” *Econometrica*, 86 (4), 1159–1214.
- Hellman, Martin E and Thomas M Cover (1970) “Learning with finite memory,” *The Annals of Mathematical Statistics*, 41 (3), 765–782.
- Hellman, Martin Edward (1969) *Learning with finite memory*, Ph.D. dissertation, Stanford University.
- Hoffman, Mitchell and Stephen V Burks (2017) “Worker overconfidence: Field evidence and implications for employee turnover and returns from training,” Technical report, National Bureau of Economic Research.
- Jehiel, Philippe and Jakub Steiner (2020) “Selective sampling with information-storage constraints,” *The Economic Journal*, 130 (630), 1753–1781, <https://doi.org/10.1093/ej/uez068>.
- Kahneman, Daniel, Paul Slovic, and Amos Tversky (1982) *Judgment under uncertainty: Heuristics and biases*: Cambridge University Press.
- Leung, Benson Tsz Kin (2020) “Limited cognitive ability and selective information processing,” *Games and Economic Behavior*, 120, 345–369.
- Mailath, George J and Larry Samuelson (2020) “Learning under Diverse World Views: Model-Based Inference,” *American Economic Review*, 110 (5), 1464–1501.
- Matějka, Filip and Alisdair McKay (2015) “Rational inattention to discrete choices: A new foundation for the multinomial logit model,” *American Economic Review*, 105 (1), 272–98.
- Monte, Daniel and Maher Said (2014) “The value of (bounded) memory in a changing world,” *Economic Theory*, 56 (1), 59–82.

- Morris, Stephen (1994) “Trade with heterogeneous prior beliefs and asymmetric information,” *Econometrica*, 1327–1347.
- Oprea, Ryan (2020) “What Makes a Rule Complex?” *American Economic Review*, forthcoming.
- Rabin, Matthew and Joel L Schrag (1999) “First impressions matter: A model of confirmatory bias,” *The Quarterly Journal of Economics*, 114 (1), 37–82.
- Savage, Leonard J (1972) *The foundations of statistics*: Courier Corporation.
- Sims, Christopher A (2003) “Implications of rational inattention,” *Journal of monetary Economics*, 50 (3), 665–690.
- Steiner, Jakub and Colin Stewart (2016) “Perceiving prospects properly,” *American Economic Review*, 106 (7), 1601–31.
- Wilson, Andrea (2014) “Bounded memory and biases in information processing,” *Econometrica*, 82 (6), 2257–2294.
- Wu, Shijie, Ozan Irsoy, Steven Lu et al. (2023) “Bloomberggpt: A large language model for finance,” *arXiv preprint arXiv:2303.17564*.
- Yoon, Seoin, Shawn T McClean, Nitya Chawla, Ji Koungh Kim, Joel Koopman, Christopher C Rosen, John P Trougakos, and Julie M McCarthy (2021) “Working through an “infodemic”: The impact of COVID-19 news consumption on employee uncertainty and work behaviors,” *Journal of applied psychology*, 106 (4), 501.

A Proofs

Before I present the proofs of the results in the main text, I first prove the following Lemma that is used throughout the proof of the different results.

Lemma A.1. *Fix $N, (f^\omega)_{\omega=1}^N$ and any updating mechanism (irreducible or reducible), for all ω, ω' ,*

$$\min_m \frac{\mu_m^{\omega'}}{\mu_m^\omega} \geq \varsigma^{2(M-1)}. \quad (\text{A.1})$$

It implies that $\mu_m^\omega = 0$ if and only if $\mu_m^{\omega'} = 0$ for all ω, ω', m ; in other words, $\mu_m^\omega > 0$ if and only if $\mu_m^{\omega'} > 0$ for all ω, ω', m .

Proof. I first focus on irreducible updating mechanisms. Consider state ω and ω' , recall that the long-run distribution in state ω and ω' are the solution of the following fixed point equations:

$$\begin{aligned} \boldsymbol{\mu}^\omega &= (\boldsymbol{\mu}^\omega)^T \mathbf{Q}^\omega; \\ \boldsymbol{\mu}^{\omega'} &= (\boldsymbol{\mu}^{\omega'})^T \mathbf{Q}^{\omega'}. \end{aligned} \quad (\text{A.2})$$

Equation (A.2) shows that, given an updating mechanism $(m_1, \mathcal{T}, d)$, f^ω and $f^{\omega'}$, the long-run distribution $\boldsymbol{\mu}^\omega, \boldsymbol{\mu}^{\omega'}$ and the ratio $(\frac{\mu_m^{\omega'}}{\mu_m^\omega})_{m=1}^M$ are invariant in $N, (p^{\omega''})_{\omega''=1}^N$ and $(f^{\omega''})_{\omega'' \neq \omega, \omega'}$. Therefore, we can apply Theorem 2 of [Hellman and Cover \(1970\)](#) where $N = 2$ such that:

$$\frac{\max_m \frac{\mu_m^{\omega'}}{\mu_m^\omega}}{\min_m \frac{\mu_m^{\omega'}}{\mu_m^\omega}} \leq (\bar{l}^{\omega\omega'} \bar{l}^{\omega'\omega})^{(M-1)}. \quad (\text{A.3})$$

where $\bar{l}^{\omega\omega'} = \sup_s \frac{f^\omega(s)}{f^{\omega'}(s)}$. First, $\min_m \frac{\mu_m^{\omega'}}{\mu_m^\omega} > 0$ for all $\omega' \neq \omega$. Suppose to the contrary that for some ω' , $\min_m \frac{\mu_m^{\omega'}}{\mu_m^\omega} = 0$, Equation (A.3) implies that $\max_m \frac{\mu_m^{\omega'}}{\mu_m^\omega} = 0$ and $\mu_m^{\omega'} = 0$ for all m . It in turn implies that $\max_m \frac{\mu_m^\omega}{\mu_m^{\omega'}}$ is unbounded and $\min_m \frac{\mu_m^\omega}{\mu_m^{\omega'}} = 0$ and it contradicts

$$\frac{\max_m \frac{\mu_m^\omega}{\mu_m^{\omega'}}}{\min_m \frac{\mu_m^\omega}{\mu_m^{\omega'}}} \leq (\bar{l}^{\omega'\omega} \bar{l}^{\omega\omega'})^{(M-1)}.$$

Now since $\min_m \frac{\mu_m^{\omega'}}{\mu_m^\omega} > 0$ for all ω' , we have

$$\begin{aligned} \min_m \frac{\mu_m^{\omega'}}{\mu_m^\omega} &\geq (\bar{l}^{\omega\omega'} \bar{l}^{\omega'\omega})^{-(M-1)} \max_m \frac{\mu_m^{\omega'}}{\mu_m^\omega} \\ \min_m \frac{\mu_m^{\omega'}}{\mu_m^\omega} &\geq \varsigma^{2(M-1)} \max_m \frac{\mu_m^{\omega'}}{\mu_m^\omega} \\ \min_m \frac{\mu_m^{\omega'}}{\mu_m^\omega} &\geq \varsigma^{2(M-1)} \end{aligned}$$

The second inequality follows $\varsigma \equiv \min_{\omega' \neq \omega, s \in S} \frac{f^\omega(s)}{f^{\omega'}(s)}$ while the third inequality is implied by the fact that $\sum_m \mu_m^{\omega'} = \sum_m \mu_m^\omega = 1$. Now, if $\mu_m^{\omega'} = 0$ for some m , we must have $\mu_m^\omega = 0$ as otherwise it violates Inequality (A.1).

The above argument handles the case where the updating mechanism induces an irreducible Markov chain, so that the long-run distribution μ^ω does not depend on the initial memory state m_1 . For a reducible chain, denote the recurrent communicating classes as $\mathcal{R}_1, \dots, \mathcal{R}_r$ and the set of transient memory states as $\mathcal{R}_0$. If $r = 1$ or $m_1 \in \mathcal{R}_j$ for some $j \geq 1$, the chain is effectively irreducible on $\mathcal{R}_j$ (with $|\mathcal{R}_j| \leq M$ states), and the bound above applies within $\mathcal{R}_j$, yielding $\min_m \mu_m^{\omega'} / \mu_m^\omega \geq \varsigma^{2(|\mathcal{R}_j|-1)} \geq \varsigma^{2(M-1)}$.

In the remaining case, $r \geq 2$ and $m_1 \in \mathcal{R}_0$. Let $\mathcal{P}^\omega(\mathcal{R}_j)$ denote the probability of absorption into $\mathcal{R}_j$ under state ω , and let $\mu^{0,\omega}$ be the stationary distribution on $\mathcal{R}_0$ for the modified transition rule that re-routes transitions out of $\mathcal{R}_0$ back to m_1 . As is standard (see Appendix 2 of Hellman (1969)),

$$\mathcal{P}^\omega(\mathcal{R}_j) = \sum_{m \in \mathcal{R}_0} \mu_m^{0,\omega} \sum_{m' \in \mathcal{R}_j} q_{mm'}^\omega,$$

where $q_{mm'}^\omega = \sum_s f^\omega(s) \Pr[\mathcal{T}(m, s) = m']$. Let $\mu_m^{j,\omega} = \mu_m^\omega / \sum_{m' \in \mathcal{R}_j} \mu_{m'}^\omega$ for $m \in \mathcal{R}_j$ (within-class stationary distribution). For $m \in \mathcal{R}_j$, $\mu_m^\omega = \mathcal{P}^\omega(\mathcal{R}_j) \cdot \mu_m^{j,\omega}$, so

$$\frac{\mu_m^{\omega'}}{\mu_m^\omega} = \frac{\mathcal{P}^{\omega'}(\mathcal{R}_j)}{\mathcal{P}^\omega(\mathcal{R}_j)} \cdot \frac{\mu_m^{j,\omega'}}{\mu_m^{j,\omega}}.$$

The within-class factor satisfies $\mu_m^{j,\omega'} / \mu_m^{j,\omega} \geq \varsigma^{2(|\mathcal{R}_j|-1)}$ by the irreducible-case bound applied to $\mathcal{R}_j$. The absorption-probability ratio satisfies

$$\frac{\mathcal{P}^{\omega'}(\mathcal{R}_j)}{\mathcal{P}^\omega(\mathcal{R}_j)} \geq \varsigma^2 \cdot \min_{m \in \mathcal{R}_0} \frac{\mu_m^{0,\omega'}}{\mu_m^{0,\omega}} \geq \varsigma^2 \cdot \varsigma^{2(|\mathcal{R}_0|-1)} = \varsigma^{2|\mathcal{R}_0|},$$

where the first inequality uses $q_{mm'}^{\omega'} \geq \varsigma \cdot q_{mm'}^\omega$ pointwise and the second applies the irreducible-case bound to the transient sub-chain $\mathcal{R}_0$. Combining,

$$\frac{\mu_m^{\omega'}}{\mu_m^\omega} \geq \varsigma^{2|\mathcal{R}_0|+2(|\mathcal{R}_j|-1)} \geq \varsigma^{2(M-1)},$$

where the last inequality uses $|\mathcal{R}_0| + |\mathcal{R}_j| \leq M$. For $m \in \mathcal{R}_0$, $\mu_m^\omega = \mu_m^{\omega'} = 0$ in the long run, so the bound holds vacuously.

The Lemma follows as Inequality (A.1) holds for all ω, ω' . $\square$

A.1 Proof of Proposition 1

I first define “confirmatory signals” for each state: A signal s supports state ω with probability $G^\omega(s)$ and supports no state with probability $1 - \sum_{\omega=1}^N G^\omega(s)$, where

$$G^\omega(s) \propto \frac{f^\omega(s)}{\sqrt{\sum_{s \in S} (f^\omega(s))^2}} \quad (\text{A.4})$$

with normalization such that $\sum_{\omega=1}^N G^\omega(s) \leq 1$. With some abuse of notations, I use $F^\omega(G^{\omega'}(s))$ to denote the probability of receiving a signal supporting state ω' , or equivalently “a signal $G^{\omega'}$ ”, in state ω , i.e., $F^\omega(G^{\omega'}) \equiv \sum_{s \in S} f^\omega(s) G^{\omega'}(s)$. Importantly, Equation A.4 implies it is more likely that the DM receives a signal supporting a correct state than a signal supporting a wrong state:

$$F^\omega(G^\omega) = \frac{\sum_{s \in S} (f^\omega(s))^2}{\sqrt{\sum_{s \in S} (f^\omega(s))^2}} = \sqrt{\sum_{s \in S} (f^\omega(s))^2} > \frac{\sum_{s \in S} f^\omega(s) f^{\omega'}(s)}{\sqrt{\sum_{s \in S} (f^{\omega'}(s))^2}} = F^\omega(G^{\omega'}),$$

implied by the Cauchy–Schwarz inequality.

Proof of Proposition 1 (i, L_M^* strictly decreases in M). Take a sequence of ϵ -optimal mechanism indexed by k : $(m_{1k}, \mathcal{T}_k, d_k)_{k=1}^\infty$ such that

$$\begin{aligned} d_k &\equiv d; \\ \lim_{k \rightarrow \infty} L(m_{1k}, \mathcal{T}_k, d_k) &= L_M^*; \\ \lim_{k \rightarrow \infty} (\mu_{m,k}^\omega)_{\omega,m} &= (\mu_{m,M}^{*\omega})_{\omega,m}. \end{aligned}$$

The first equation says that the decision rules are unchanged throughout the sequence. The second equation says that the limit of the sequence achieves the infimum utility loss. The third equation says that the long-run distributions induced by the sequence of updating mechanisms converge. Such sequence exists because of the definition of L_M^* , and the continuity of the utility loss in the long-run distribution $(\mu_m^\omega)_{\omega,m}$ and the decision rule d . Together, these equations imply that

$$L_M^* = \sum_{\omega=1}^N \left[u(\omega, \omega) - \sum_{m=1}^M \mu_{m,M}^{*\omega} u(d(m), \omega) \right].$$

Now take the k -th updating mechanism for some large k . For simplicity of notation, I omit the subscript k and denote it as $(m_1, \mathcal{T}, d)$. I construct a mechanism with memory size $M+1$, denoted as $(m'_1, \mathcal{T}', d')$, that delivers strictly higher asymptotic utility than $(m_1, \mathcal{T}, d)$. Moreover, I show that this increase is bounded away from 0 uniformly in k . The result then follows.

I first select a few objects from the original mechanism $(m_1, \mathcal{T}, d)$ which are used in the construction of $(m'_1, \mathcal{T}', d')$. Let $\hat{\Omega}$ denote the set of actions taken with strictly positive

probability at the limit, i.e., $a \in \hat{\Omega}$ if $\sum_{m:d(m)=a} \mu_{m,M}^{*\omega} > 0$. By Equation (4), there exists at least two actions in $\hat{\Omega}$. Pick any action in $\hat{\Omega}$ and without loss of generality, call it action 1. Pick a memory state $\tilde{m}$ with $d(\tilde{m}) = 1$ and $\mu_{\tilde{m}}^{*\omega} > 0$ for some ω . By Lemma A.1, $\min_m \frac{\mu_m^{\omega'}}{\mu_m^{\omega}}$ is uniformly bounded away from 0 across all states and mechanisms; hence $\mu_{\tilde{m}}^{*\omega'} > 0$ for every ω' .

Next, pick another element of $\hat{\Omega}$ — which I relabel as N without loss of generality — such that $\arg \min_{a \in \hat{\Omega}} u(a, N) = 1$. In other words, in state N , shifting probability away from action 1 to other actions (while preserving their relative likelihoods) strictly raises utility in state N .

Given these selections, for constants $(\delta_i)_{i=0}^N$ specified below, $(m'_1, \mathcal{T}', d')$ is defined as follows:

$$\begin{aligned} m'_1 &= m_1. \\ d(m) &= d'(m) \text{ for all } m = 1, \dots, M. \\ d(M+1) &= 1. \\ \Pr[\mathcal{T}'(m, \cdot) = m'] &\propto \Pr[\mathcal{T}(m, \cdot) = m'] \text{ for all } m, m' \notin \{\tilde{m}, M+1\}. \\ \Pr[\mathcal{T}'(\tilde{m}, s) = m] &\propto \begin{cases} \delta_0 \Pr[\mathcal{T}(\tilde{m}, s) = m] & \text{if } m \neq M+1.; \\ \sum_{i=1}^N \delta_i G^i(s) & \text{if } m = M+1. \end{cases} \\ \Pr[\mathcal{T}'(M+1, s) = m] &\propto \begin{cases} c & \text{if } m = \tilde{m}; \\ 0 & \text{if } m \neq \tilde{m}, M+1. \end{cases} \end{aligned}$$

for some constant c and with appropriate normalization such that $\sum_{m' \neq m} \Pr[\mathcal{T}'(m, s) = m'] \leq 1$ for all m .³¹ $(\delta_i)_{i=1}^N$ is chosen such that they satisfy the following system of linear equations for some chosen $\Delta > 1$:

$$\begin{aligned} \delta_1 &\begin{bmatrix} \sum_{s \in S} G^1(s) f^1(s) - \Delta \sum_{s \in S} G^N(s) f^1(s) \\ \sum_{s \in S} G^1(s) f^2(s) - \Delta \sum_{s \in S} G^N(s) f^2(s) \\ \vdots \\ \sum_{s \in S} G^1(s) f^{N-1}(s) - \Delta \sum_{s \in S} G^N(s) f^{N-1}(s) \end{bmatrix} + \delta_2 \begin{bmatrix} \sum_{s \in S} G^2(s) f^1(s) - \Delta \sum_{s \in S} G^N(s) f^1(s) \\ \sum_{s \in S} G^2(s) f^2(s) - \Delta \sum_{s \in S} G^N(s) f^2(s) \\ \vdots \\ \sum_{s \in S} G^2(s) f^{N-1}(s) - \Delta \sum_{s \in S} G^N(s) f^{N-1}(s) \end{bmatrix} + \dots + \\ \delta_{N-1} &\begin{bmatrix} \sum_{s \in S} G^{N-1}(s) f^1(s) - \Delta \sum_{s \in S} G^N(s) f^1(s) \\ \sum_{s \in S} G^{N-1}(s) f^2(s) - \Delta \sum_{s \in S} G^N(s) f^2(s) \\ \vdots \\ \sum_{s \in S} G^{N-1}(s) f^{N-1}(s) - \Delta \sum_{s \in S} G^N(s) f^{N-1}(s) \end{bmatrix} + \delta_N \begin{bmatrix} \sum_{s \in S} G^N(s) f^1(s) - \Delta \sum_{s \in S} G^N(s) f^1(s) \\ \sum_{s \in S} G^N(s) f^2(s) - \Delta \sum_{s \in S} G^N(s) f^2(s) \\ \vdots \\ \sum_{s \in S} G^N(s) f^{N-1}(s) - \Delta \sum_{s \in S} G^N(s) f^{N-1}(s) \end{bmatrix} = 0 \end{aligned} \quad (\text{A.5})$$

Such $(\delta_i)_{i=1}^N$ exists as there are more variables than number of Equations. Now, denote $(\mu_m^{\omega'})_{m=1}^{M+1}$ as the long-run distribution of $(m'_1, \mathcal{T}', d')$ and $(\mu_m^{\omega})_{m=1}^M$ as the long-run distribution of $(m_1, \mathcal{T}, d)$. We have for all $\omega = 1, \dots, N-1$,

$$\frac{\mu_{M+1}^{\omega'}}{\mu_{\tilde{m}}^{\omega'}} / \frac{\mu_{M+1}^N}{\mu_{\tilde{m}}^N} = \frac{\sum_{i=1}^N \sum_{s \in S} G^i(s) f^{\omega}(s)}{\sum_{i=1}^N \sum_{s \in S} G^i(s) f^N(s)} = \Delta > 1$$

³¹Note that the long-run distribution is invariant if the transition matrix $\Pr[\mathcal{T}'(m, s) = m']$ is scaled up by a common factor.

where Equation (A.5) ensures that the second equality holds. Next, set $\delta_0 = 1 + \frac{\sum_{i=1}^N \sum_{s \in S} G^i(s) f^N(s)}{c} \Delta$, we have for $\omega = 1, \dots, N-1$,

$$\begin{aligned}\mu'_{\tilde{m}} &= \frac{1}{\delta_0} \mu_{\tilde{m}}^\omega, \\ \mu'_{\tilde{m}} + \mu'_{M+1} &= \mu'_{\tilde{m}} \left(1 + \frac{\sum_{i=1}^N \sum_{s \in S} G^i(s) f^N(s)}{c} \Delta \right) = \mu_{\tilde{m}}^\omega, \\ \mu'_m &= \mu_m^\omega \text{ for all } m \neq \tilde{m}, M+1.\end{aligned}$$

Thus, asymptotic utility in states 1 to $N-1$ remains unchanged. On the other hand, in state N ,

$$\begin{aligned}\mu'_m &= \frac{1}{1 - \frac{\sum_{i=1}^N \sum_{s \in S} G^i(s) f^N(s)}{c} (\Delta - 1) \mu_{\tilde{m}}^\omega} \mu_m^\omega \text{ for all } m \neq \tilde{m}, M+1, \\ \mu'_{\tilde{m}} &= \frac{1}{1 - \frac{\sum_{i=1}^N \sum_{s \in S} G^i(s) f^N(s)}{c} (\Delta - 1) \mu_{\tilde{m}}^\omega} \left(\frac{1}{\delta_0} \mu_{\tilde{m}}^\omega \right) \\ \mu'_{\tilde{m}} + \mu'_{M+1} &= \mu'_{\tilde{m}} \left(1 + \frac{\sum_{i=1}^N \sum_{s \in S} G^i(s) f^N(s)}{c} \right).\end{aligned}$$

Pick $\Delta > 1$ small enough that $1 - \frac{\sum_{i=1}^N \sum_{s \in S} G^i(s) f^N(s)}{c} (\Delta - 1) \mu_{\tilde{m}}^{*N} > 0$; since $\mu_{\tilde{m}}^{N,k} \rightarrow \mu_{\tilde{m}}^{*N} > 0$, the same Δ makes the probabilities well-defined for all sufficiently large k . With this choice, the DM chooses action 1 with a lower probability in state N , and other actions with a higher probability (by a factor of $\frac{1}{1 - \frac{\sum_{i=1}^N \sum_{s \in S} G^i(s) f^N(s)}{c} (\Delta - 1) \mu_{\tilde{m}}^{*N}}$). Moreover, the constants $(\delta_i)_{i=1}^N$ and c depend only on the primitives $(G^i, f^\omega)_{i,\omega}$, not on k . As $k \rightarrow \infty$, the change in asymptotic utility in state N converges to:

$$\begin{aligned}& \sum_{m \neq \tilde{m}, M+1} \left[\left(\frac{1}{1 - \frac{\sum_{i=1}^N \sum_{s \in S} G^i(s) f^N(s)}{c} (\Delta - 1) \mu_{\tilde{m}}^{*N}} - 1 \right) \mu_m^{*N} (u(d(m), N) - u(1, N)) \right] \\ &= \sum_{m \neq \tilde{m}, M+1} \left[\left(\frac{\frac{\sum_{i=1}^N \sum_{s \in S} G^i(s) f^N(s)}{c} (\Delta - 1) \mu_{\tilde{m}}^{*N}}{1 - \frac{\sum_{i=1}^N \sum_{s \in S} G^i(s) f^N(s)}{c} (\Delta - 1) \mu_{\tilde{m}}^{*N}} \right) \mu_m^{*N} (u(d(m), N) - u(1, N)) \right] \quad (\text{A.6}) \\ &> 0.\end{aligned}$$

The inequality follows as 1) $\Delta > 1$, 2) $\mu_{\tilde{m}}^{*N} > 0$, 3) $\arg \min_a u(a, N) = 1$ such that $u(a, N) - u(1, N) \geq 0$ for all a , and 4) $\sum_{m: d(m)=N} \mu_m^{*N} > 0$, and action N is the unique maximizer in state N such that $u(N, N) - u(1, N) > 0$.

Denote the strictly positive limit in Equation (A.6) by $2\bar{\epsilon} > 0$. Since the improvement converges to this limit, there exists K^* such that $L(m_{1k}, \mathcal{T}_k, d_k) - L(m'_{1k}, \mathcal{T}'_k, d'_k) \geq \bar{\epsilon}$ for all $k \geq K^*$ (note that $\bar{\epsilon}$ is invariant in k). Combined with $L(m_{1k}, \mathcal{T}_k, d_k) \rightarrow L_M^*$, we have

$$L_{M+1}^* \leq \liminf_{k \rightarrow \infty} L(m'_{1k}, \mathcal{T}'_k, d'_k) \leq \lim_{k \rightarrow \infty} L(m_{1k}, \mathcal{T}_k, d_k) - \bar{\epsilon} = L_M^* - \bar{\epsilon} < L_M^*,$$

which establishes the strict inequality. $\square$

Proof of Proposition 1 (i, convergence rate bound) and (ii). I formally describe the simple updating mechanism described in the main text. Denote $\lambda = \lfloor (M-1)/N \rfloor$, relabel the memory states as $0, 11, \dots, 1\lambda, 21, \dots, 2\lambda, \dots, N\lambda$ and denote the unused memory states as $N\lambda + 1, N\lambda + 2, \dots, M-1$. The decision rule is as follows:

$$\begin{aligned} d(0) &= 1; \\ d(\omega k) &= \omega \text{ for all } \omega = 1, \dots, N \text{ and } k = 1, \dots, \lambda; \\ d(m) &= 1 \text{ for all } m > N\lambda. \end{aligned}$$

The transition function between the memory states is defined below for some big enough $\delta > 1$ such that

$$\frac{\delta F^\omega(G^{\omega'})}{\sum_{\omega'' \neq \omega'} F^\omega(G^{\omega''})} > 1 \text{ for all } \omega, \omega' \in \Omega.$$

Suppose the DM receives some signal s , he follows the following transition rule:

$$\begin{aligned} \mathcal{T}(0, s) &= \begin{cases} \omega 1 & \text{with probability } G^\omega(s) \text{ for all } \omega = 1, \dots, N; \\ 0 & \text{with probability } 1 - \sum_{\omega=1}^N G^\omega(s). \end{cases} \\ \mathcal{T}(\omega 1, s) &= \begin{cases} \omega 2 & \text{with probability } G^\omega(s); \\ 0 & \text{with probability } \sum_{j \in \Omega \setminus \{\omega\}} \frac{G^j(s)}{\delta}; \\ \omega 1 & \text{with probability } 1 - \sum_{j \in \Omega \setminus \{\omega\}} \frac{G^j(s)}{\delta} - G^\omega(s). \end{cases} \\ \mathcal{T}(\omega \lambda, s) &= \begin{cases} \omega(\lambda - 1) & \text{with probability } \sum_{j \in \Omega \setminus \{\omega\}} \frac{G^j(s)}{\delta}; \\ \omega \lambda & \text{with probability } 1 - \sum_{j \in \Omega \setminus \{\omega\}} \frac{G^j(s)}{\delta}. \end{cases} \end{aligned}$$

while for $k = 2, 3, \dots, \lambda - 1$,

$$\mathcal{T}(\omega k, s) = \begin{cases} \omega(k+1) & \text{with probability } G^\omega(s); \\ \omega(k-1) & \text{with probability } \sum_{j \in \Omega \setminus \{\omega\}} \frac{G^j(s)}{\delta}; \\ \omega k & \text{with probability } 1 - \sum_{j \in \Omega \setminus \{\omega\}} \frac{G^j(s)}{\delta} - G^\omega(s). \end{cases}$$

Finally, for $m > N\lambda$, $\mathcal{T}(m, s) = m$ for all s . By restricting the initial memory state to one of $0, 11, 12, \dots, 1\lambda, 21, 22, \dots, 2\lambda, \dots, N\lambda$, the DM will never transition to memory states $m > N\lambda$.

Note that this updating mechanism does not depend on p^ω nor u^ω . Now I compute the long-run distribution μ^ω and the utility loss $L(m_1, \mathcal{T}, d)$. Fix a state ω , at the two extreme

memory states in branch ω' , i.e., memory states $\omega'\lambda$ and $\omega'(\lambda - 1)$,

$$\mu_{\omega'(\lambda-1)}^\omega F^\omega(G^{\omega'}) = \mu_{\omega'\lambda}^\omega \frac{1}{\delta} \sum_{\omega'' \neq \omega'} F^\omega(G^{\omega''})$$

$$\mu_{\omega'(\lambda-1)}^\omega = \mu_{\omega'\lambda}^\omega \left[\frac{\delta F^\omega(G^{\omega'})}{\sum_{\omega'' \neq \omega'} F^\omega(G^{\omega''})} \right]^{-1}$$

for all ω' . Next, at memory state $\omega'(\lambda - 1)$,

$$\mu_{\omega'\lambda}^\omega \frac{1}{\delta} \sum_{\omega'' \neq \omega'} F^\omega(G^{\omega''}) + \mu_{\omega'(\lambda-2)}^\omega F^\omega(G^{\omega'}) = \mu_{\omega'(\lambda-1)}^\omega \left[\frac{1}{\delta} \sum_{\omega'' \neq \omega'} F^\omega(G^{\omega''}) + F^\omega(G^{\omega'}) \right]$$

$$\mu_{\omega'(\lambda-2)}^\omega = \mu_{\omega'(\lambda-1)}^\omega \left[\frac{\delta F^\omega(G^{\omega'})}{\sum_{\omega'' \neq \omega'} F^\omega(G^{\omega''})} \right]^{-1}$$

Repeating the same procedures implies that for all $k = 1, \dots, \lambda$ and $\omega' = 1, \dots, N$

$$\mu_{\omega'k}^\omega = \mu_{\omega'\lambda}^\omega \left[\frac{\delta F^\omega(G^{\omega'})}{\sum_{\omega'' \neq \omega'} F^\omega(G^{\omega''})} \right]^{-(\lambda-k)} \quad (\text{A.7})$$

and

$$\mu_{\omega'k}^\omega = \mu_0^\omega \left[\frac{\delta F^\omega(G^{\omega'})}{\sum_{\omega'' \neq \omega'} F^\omega(G^{\omega''})} \right]^k \quad (\text{A.8})$$

As $\sum_{\omega'=1}^N \sum_{k=1}^\lambda \mu_{\omega'k}^\omega + \mu_0^\omega = 1$, and denote $\frac{\delta F^\omega(G^{\omega'})}{\sum_{\omega'' \neq \omega'} F^\omega(G^{\omega''})}$ by $r^{\omega\omega'}$, we have

$$\mu_{\omega\lambda}^\omega \sum_{k=1}^\lambda [r^{\omega\omega}]^{-(\lambda-k)} + \mu_{\omega\lambda}^\omega [r^{\omega\omega}]^{-\lambda} + \mu_{\omega\lambda}^\omega [r^{\omega\omega}]^{-\lambda} \sum_{\omega' \neq \omega} \sum_{k=1}^\lambda [r^{\omega\omega'}]^k = 1$$

$$\Rightarrow \mu_{\omega\lambda}^\omega = \frac{1}{\sum_{k=1}^\lambda [r^{\omega\omega}]^{-(\lambda-k)} + [r^{\omega\omega}]^{-\lambda} + [r^{\omega\omega}]^{-\lambda} \sum_{\omega' \neq \omega} \sum_{k=1}^\lambda [r^{\omega\omega'}]^k}$$

and the probability of choosing actions in $\Omega \setminus \omega$ in state ω is smaller than $\sum_{m \notin \{\omega_1, \dots, \omega_\lambda\}} \mu_m^\omega$ which is as follows:

$$\frac{[r^{\omega\omega}]^{-\lambda} + [r^{\omega\omega}]^{-\lambda} \sum_{\omega' \neq \omega} \sum_{k=1}^\lambda [r^{\omega\omega'}]^k}{\sum_{k=1}^\lambda [r^{\omega\omega}]^{-(\lambda-k)} + [r^{\omega\omega}]^{-\lambda} + [r^{\omega\omega}]^{-\lambda} \sum_{\omega' \neq \omega} \sum_{k=1}^\lambda [r^{\omega\omega'}]^k}$$

$$< [r^{\omega\omega}]^{-\lambda} + [r^{\omega\omega}]^{-\lambda} \sum_{\omega' \neq \omega} \sum_{k=1}^\lambda [r^{\omega\omega'}]^k$$

$$= [r^{\omega\omega}]^{-\lambda} + \sum_{\omega' \neq \omega} \frac{r^{\omega\omega'}}{r^{\omega\omega'} - 1} \left[\left(\frac{r^{\omega\omega'}}{r^{\omega\omega}} \right)^\lambda - [r^{\omega\omega}]^{-\lambda} \right]$$

$$< [r^{\omega\omega}]^{-\lambda} + \sum_{\omega' \neq \omega} \frac{r^{\omega\omega'}}{r^{\omega\omega'} - 1} \left[\left(\frac{r^{\omega\omega'}}{r^{\omega\omega}} \right)^\lambda \right] \quad (\text{A.9})$$

$$\begin{aligned}
&< [r^{\omega\omega}]^{-\lambda} + (N-1) \frac{\max_{\omega' \neq \omega} r^{\omega\omega'}}{\max_{\omega' \neq \omega} r^{\omega\omega'} - 1} \left[\max_{\omega' \neq \omega} \left(\frac{r^{\omega\omega'}}{r^{\omega\omega}} \right)^\lambda \right] \\
&< \left[(N-1) \frac{\max_{\omega' \neq \omega} r^{\omega\omega'}}{\max_{\omega' \neq \omega} r^{\omega\omega'} - 1} + 1 \right] \max \left\{ [r^{\omega\omega}]^{-\lambda}, \max_{\omega' \neq \omega} \left(\frac{r^{\omega\omega'}}{r^{\omega\omega}} \right)^\lambda \right\}
\end{aligned}$$

The first inequality of Equation (A.9) is implied by the fact that the denominator is strictly greater than 1. Now, using Equation (A.9) and denoting $K^\omega = \left[(N-1) \frac{\max_{\omega' \neq \omega} r^{\omega\omega'}}{\max_{\omega' \neq \omega} r^{\omega\omega'} - 1} + 1 \right]$, we can compute the upper bound of the utility loss $L(m_1, \mathcal{T}, d)$

$$\begin{aligned}
L(m_1, \mathcal{T}, d) &\leq \sum_{\omega=1}^N p^\omega \max_{a \neq \omega} \{u(\omega, \omega) - u(a, \omega)\} \left[\sum_{m \notin \{\omega 1, \dots, \omega \lambda\}} \mu_m^\omega \right] \\
&< \sum_{\omega=1}^N p^\omega \max_{a \neq \omega} \{u(\omega, \omega) - u(a, \omega)\} K^\omega \max \left\{ [r^{\omega\omega}]^{-\lambda}, \max_{\omega' \neq \omega} \left(\frac{r^{\omega\omega'}}{r^{\omega\omega}} \right)^\lambda \right\} \quad (\text{A.10}) \\
&< \max_{\omega} \left[\max_{a \neq \omega} \{u(\omega, \omega) - u(a, \omega)\} K^\omega \max \left\{ [r^{\omega\omega}]^{-\lambda}, \max_{\omega' \neq \omega} \left(\frac{r^{\omega\omega'}}{r^{\omega\omega}} \right)^\lambda \right\} \right]
\end{aligned}$$

As $[r^{\omega\omega}]^{-1}$ and $\max_{\omega' \neq \omega} \left(\frac{r^{\omega\omega'}}{r^{\omega\omega}} \right)$ are strictly smaller than 1 for all ω , the exponential rate bound in Proposition 1(i) follows. Moreover, the right-hand side of Equation (A.9) converges to 0 as λ goes to ∞ . This proves Proposition 1(ii). $\square$

A.2 Proof of Proposition 2 and Proposition 5

Proof. Fix a sequence $\left\{ (\mathbf{u}_{N(k)}^\omega, p_{N(k)}^\omega, f_{N(k)}^\omega)_{\omega=1}^{N(k)}, M(k) \right\}_{k=1}^\infty$ satisfying Assumptions (Limit-FS), (Limit-NRO), and (Limit-ID), where $N(k) = \Theta(M(k)^h)$. I first prove that $\lim_{k \rightarrow \infty} L_k^* = 0$ if $h < 1$, then proceed to the case where $h \geq 1$ and prove Proposition 5. Lastly, I prove that Proposition 5 implies $\lim_{k \rightarrow \infty} L_k^* = 1 - \bar{p}_0$.

1) $h < 1$. I show that for all $\epsilon > 0$, there exists a sequence of updating mechanisms with utility losses that converge to smaller than ϵ as $k \rightarrow \infty$. First, by Assumption (Limit-ID), for all $\frac{\epsilon}{u} > 0$, there exists an $\xi > 0$ and a sequence of subset of states N_ξ such that $\lim_{N \rightarrow \infty} \left(\sum_{\omega \in N_\xi} p_N^\omega \right) > 1 - \frac{\epsilon}{u}$ and

$$\lim_{N \rightarrow \infty} \inf_{\omega, \omega' \in N_\xi; \omega' \neq \omega} \left\{ -\log \frac{\int f_N^\omega(s) f_N^{\omega'}(s) ds}{\sqrt{\int (f_N^\omega(s))^2 ds} \sqrt{\int (f_N^{\omega'}(s))^2 ds}} \right\} > \xi$$

Before proving the Proposition, I first prove the following Lemma for $(G_N^\omega(s))_{\omega=1}^N$ defined as:

$$G_N^\omega(s) = \frac{f_N^\omega(s)}{\sqrt{\int (f_N^\omega(s))^2 ds}}.$$

Lemma A.2. For $\frac{\epsilon}{u} > 0$, there exists an $\tilde{\xi} > 0$, a sequence of $\delta_N > 1$ and a sequence of subset of states N_ξ with $\lim_{N \rightarrow \infty} \left(\sum_{\omega \in N_\xi} p_N^\omega \right) > 1 - \frac{\epsilon}{u}$ such that

$$\begin{aligned} \frac{\delta_N F_N^\omega(G_N^{\omega'})}{\sum_{\omega'' \neq \omega'} F_N^\omega(G_N^{\omega''})} &> 2 \text{ for all } \omega, \omega' \in \Omega \text{ and } N; \\ \lim_{N \rightarrow \infty} \inf_{\omega, \omega' \in N_\xi; \omega' \neq \omega} \frac{\frac{\delta_N F_N^\omega(G_N^\omega)}{\sum_{\omega'' \neq \omega} F_N^\omega(G_N^{\omega''})}}{\frac{\delta_N F_N^\omega(G_N^{\omega'})}{\sum_{\omega'' \neq \omega'} F_N^\omega(G_N^{\omega''})}} &> 1 + \tilde{\xi}. \end{aligned} \quad (\text{A.11})$$

Proof. As $F_N^\omega(G_N^{\omega'}) > 0$ for all ω, ω' , there always exists a big enough δ_N such that the first inequality of Equation (A.11) holds. To prove the second inequality, note that

$$\begin{aligned} &\lim_{N \rightarrow \infty} \inf_{\omega, \omega' \in N_\xi; \omega' \neq \omega} \frac{\frac{\delta_N F_N^\omega(G_N^\omega)}{\sum_{\omega'' \neq \omega} F_N^\omega(G_N^{\omega''})}}{\frac{\delta_N F_N^\omega(G_N^{\omega'})}{\sum_{\omega'' \neq \omega'} F_N^\omega(G_N^{\omega''})}} \\ &\geq \lim_{N \rightarrow \infty} \inf_{\omega, \omega' \in N_\xi; \omega' \neq \omega} \frac{F_N^\omega(G_N^\omega)}{F_N^\omega(G_N^{\omega'})} \\ &= \lim_{N \rightarrow \infty} \inf_{\omega, \omega' \in N_\xi; \omega' \neq \omega} \frac{\sqrt{\int (f_N^\omega(s))^2 ds} \sqrt{\int (f_N^{\omega'}(s))^2 ds}}{\int f_N^\omega(s) f_N^{\omega'}(s) ds} \\ &> \exp(\xi) \\ &= 1 + \tilde{\xi} \end{aligned} \quad (\text{A.12})$$

where $\tilde{\xi} = \exp(\xi) - 1$ and the first inequality of Equation (A.12) is implied by the fact that $F_N^\omega(G_N^\omega) \geq F_N^\omega(G_N^{\omega'})$ and thus $\frac{\sum_{\omega'' \neq \omega'} F_N^\omega(G_N^{\omega''})}{\sum_{\omega'' \neq \omega} F_N^\omega(G_N^{\omega''})} \geq 1$. $\square$

Now consider a sequence of updating mechanisms described in the proof of Proposition 1, but include only states in N_ξ , i.e., there are only branches that correspond to actions in N_ξ , and there exists no $m \in M$ such that $d(m) = \omega'$ for $\omega' \notin N_\xi$. As in the proof of Proposition 1, denote $\frac{\delta_N F_N^\omega(G_N^{\omega'})}{\sum_{\omega'' \neq \omega'} F_N^\omega(G_N^{\omega''})}$ as $r_N^{\omega\omega'}$, the stationary distribution under state $\omega \in N_\xi$ is as follows:

$$\mu_{N\omega\lambda}^\omega \sum_{i=1}^{\lambda} [r_N^{\omega\omega}]^{-(\lambda-i)} + \mu_{N\omega\lambda}^\omega [r_N^{\omega\omega}]^{-\lambda} + \mu_{N\omega\lambda}^\omega [r_N^{\omega\omega}]^{-\lambda} \sum_{\omega' \in N_\xi \setminus \{\omega\}} \sum_{i=1}^{\lambda} [r_N^{\omega\omega'}]^{(\lambda-i)} = 1 \quad (\text{A.13})$$

where μ_{Nm}^ω is the stationary probability in memory state m under state ω when state space

has size N . Given the first inequality of Lemma A.2, $r_N^{\omega\omega} > 2$ for all ω and N and therefore $\lim_{k \rightarrow \infty} \mu_{N(k)\omega\lambda}^\omega \left[r_{N(k)}^{\omega\omega} \right]^{-\lambda(k)} = 0$. On the other hand,

$$\begin{aligned} [r_N^{\omega\omega}]^{-\lambda} \sum_{\omega' \in N_\xi \setminus \{\omega\}} \sum_{i=1}^{\lambda} [r_N^{\omega\omega'}]^{(\lambda-i)} &= [r_N^{\omega\omega}]^{-\lambda} \sum_{\omega' \in N_\xi \setminus \{\omega\}} \frac{[r_N^{\omega\omega'}]^\lambda - 1}{r_N^{\omega\omega'} - 1} \\ &\leq \sum_{\omega' \in N_\xi \setminus \{\omega\}} \left[\left[\frac{r_N^{\omega\omega'}}{r_N^{\omega\omega}} \right]^\lambda - [r_N^{\omega\omega'}]^{-\lambda} \right] \\ &\leq \sum_{\omega' \in N_\xi \setminus \{\omega\}} (1 + \tilde{\xi})^{-\lambda} \leq N(1 + \tilde{\xi})^{-\lambda} \end{aligned}$$

where the first inequality is implied by the first inequality in Lemma A.2, and the second inequality is implied by the second inequality in Lemma A.2 and that $[r_N^{\omega\omega'}]^{-\lambda} > 0$. As $(1 + \tilde{\xi})^{-\lambda}$ converges to 0 exponentially and N converges to infinity linearly, $\lim_{k \rightarrow \infty} N(k)(1 + \tilde{\xi})^{-\lambda(k)} = 0$. To see it formally, note that $N(k) = \Theta(M(k)^h)$ with $h < 1$ and $\lambda(k) = \frac{M(k)}{N(k)}$ imply that $\lambda(k) = \Theta(M(k)^{1-h})$. By the definition of Θ , we have for some constant K and some q big enough such that

$$\begin{aligned} \frac{N(k)}{(1 + \tilde{\xi})^{\lambda(k)}} &\leq K \frac{M(k)^h}{(1 + \tilde{\xi})^{M(k)(1-h)}} = K \left[\frac{M(k)}{(1 + \tilde{\xi})^{M(k)\frac{1-h}{h}}} \right]^h \\ &= K \left[\frac{M(k)}{(1 + \tilde{\xi})^{M(k)}} \right]^h \\ &< K \left[\frac{M(k)}{\binom{M(k)}{2} \tilde{\xi}^2} \right]^h \\ &= K \left[\frac{2}{(M(k) - 1) \tilde{\xi}^2} \right]^h \end{aligned}$$

where $1 + \tilde{\xi} = (1 + \tilde{\xi})^{\frac{1-h}{h}}$ and the second inequality is implied by the Binomial Theorem. Thus,

$$\begin{aligned} \lim_{k \rightarrow \infty} [r_N^{\omega\omega}]^{-\lambda} \sum_{\omega' \in N_\xi \setminus \{\omega\}} \sum_{i=1}^{\lambda} [r_N^{\omega\omega'}]^{(\lambda-i)} &\leq \lim_{k \rightarrow \infty} [r_N^{\omega\omega}]^{-\lambda} N(1 + \tilde{\xi})^{-\lambda(k)} \\ &< \lim_{k \rightarrow \infty} [r_N^{\omega\omega}]^{-\lambda} K \left[\frac{2}{(M(k) - 1) \tilde{\xi}^2} \right]^h = 0. \end{aligned}$$

It implies that $\lim_{k \rightarrow \infty} [r_N^{\omega\omega}]^{-\lambda} \sum_{\omega' \in N_\xi \setminus \{\omega\}} \sum_{i=1}^{\lambda} [r_N^{\omega\omega'}]^{(\lambda-i)} = 0$ and

$$\lim_{k \rightarrow \infty} \sum_{i=1}^{\lambda} \mu_{N\omega i}^\omega = \lim_{k \rightarrow \infty} \mu_{N\omega\lambda}^\omega \sum_{i=1}^{\lambda} [r_N^{\omega\omega}]^{-(\lambda-i)} = 1 \quad (\text{A.14})$$

for all $\omega \in N_\xi$. In state $\omega \in N_\xi$, the DM chooses action ω with asymptotic probability equal to 1. As $\lim_{N \rightarrow \infty} \sum_{\omega \in N_\xi} p_N^\omega > 1 - \frac{\epsilon}{\bar{u}}$, the asymptotic utility loss of the proposed non-ignorant updating mechanism is strictly bounded above by $\bar{u} \times \frac{\epsilon}{\bar{u}} = \epsilon$.

2) $h \geq 1$.

Proof of Proposition 5. The proof uses the following lemma.

Lemma A.3. Fix a sequence satisfying Assumptions (Limit-FS), (Limit-NRO), and (Limit-ID), where $N(k) = \Theta(M(k)^h)$ with $h \geq 1$. First, $\lim_{k \rightarrow \infty} L_k^* > 0$. Moreover, for any sequence of updating mechanisms, there does not exist a sequence of subsets $\hat{\Omega}^*(k) \subseteq \{1, \dots, N(k)\}$ with $\lim_{k \rightarrow \infty} |\hat{\Omega}^*(k)|/M(k) > 0$ such that $\lim_{k \rightarrow \infty} \max_m \mu_{km}^\omega / \mu_{km}^{\omega'} = \infty$ for all $\omega, \omega' \in \hat{\Omega}^*(k)$ with $\omega \neq \omega'$.

Proof of Lemma A.3. First, note that if in contrast $\lim_{k \rightarrow \infty} L_k^* = 0$, denote $M_k^\omega = \{m : d_k(m) = \omega\}$, we must have for some $\{m_{1k}, \mathcal{T}_k, d_k\}_{k=1}^\infty$,

$$\lim_{k \rightarrow \infty} \sum_{m \in M_k^\omega} \mu_{N(k)m}^\omega = 1$$

for almost all ω measured in prior probability, i.e., there must exist a sequence of subsets of states $\hat{\Omega}(k) \in \{1, \dots, N(k)\}$ where $\lim_{k \rightarrow \infty} \sum_{\omega \in \hat{\Omega}(k)} p_{N(k)}^\omega = 1$ and

$$\lim_{k \rightarrow \infty} \sum_{m \in M_k^\omega} \mu_{N(k)m}^\omega = 1.$$

for all $\omega \in \hat{\Omega}(k)$. Additionally, Assumption (Limit-FS) implies that there must exist a sequence of subsets of states $\hat{\Omega}(k) \in \{1, \dots, N(k)\}$ where $\lim_{k \rightarrow \infty} \frac{|\hat{\Omega}(k)|}{N(k)} = 1$ and

$$\lim_{k \rightarrow \infty} \sum_{m \in M_k^\omega} \mu_{N(k)m}^\omega = 1,$$

for all $\omega \in \hat{\Omega}(k)$. That is, the DM chooses the optimal action in almost all states, measured in both prior probability and in fraction w.r.t. N . It implies that for all ω in $\hat{\Omega}(k)$, there must exist a set of memory state $\hat{M}_k^\omega \subseteq M_k^\omega$ such that

$$\begin{aligned} \lim_{k \rightarrow \infty} \sum_{m \in \hat{M}_k^\omega} \mu_{N(k)m}^\omega &= 1; \\ \lim_{k \rightarrow \infty} \sum_{m \in \hat{M}_k^\omega} \mu_{N(k)m}^{\omega'} &= 0 \text{ for all } \omega' \in \hat{\Omega}(k) \setminus \{\omega\}; \\ \lim_{k \rightarrow \infty} \max_{m \in \hat{M}_k^\omega} \frac{\mu_{N(k)m}^\omega}{\mu_{N(k)m}^{\omega'}} &= \infty \text{ for all } \omega' \in \hat{\Omega}(k) \setminus \{\omega\}. \end{aligned} \tag{A.15}$$

In the following I prove that for Equation (A.15) to hold, $\frac{M}{N}$ has to go to ∞ . The same iterated counting argument in below also proves the second statement of the Lemma: it does not actually use the assumption $\lim L_k^* = 0$, but only the hypothesis that there exists a sequence $\hat{\Omega}^*(k)$ with $\lim_{k \rightarrow \infty} |\hat{\Omega}^*(k)|/M(k) > 0$ exhibiting unbounded likelihood ratios on all pairs within. Eq. (A.15) (which follows from $L_k^* = 0$) supplies exactly this hypothesis with $\hat{\Omega}^*(k) = \hat{\Omega}(k)$, and the same contradiction yields both statements of the Lemma.

First consider an irreducible automaton $\mathcal{T}_k$. Fix a $\omega' \in \hat{\Omega}(k) \setminus \{\omega\}$, without loss of generality rearrange the memory states such that $\frac{\mu_{N(k)m}^\omega}{\mu_{N(k)m}^{\omega'}}$ is weakly decreasing in m . As argued in the proof of Lemma A.1, Lemma 2 of Hellman and Cover (1970) holds in the current setting, so that for all $m < M(k)$,

$$\begin{aligned} \frac{\mu_{N(m+1)}^\omega}{\mu_{N(m+1)}^{\omega'}} &\geq (\bar{l}_N^{\omega\omega'} \bar{l}_N^{\omega'\omega})^{-1} \frac{\mu_{Nm}^\omega}{\mu_{Nm}^{\omega'}} \\ &\geq \varsigma^2 \frac{\mu_{Nm}^\omega}{\mu_{Nm}^{\omega'}}. \end{aligned} \tag{A.16}$$

in which I omitted the argument in $N(k)$ for ease of exposition. If $\max_m \frac{\mu_{Nm}^\omega}{\mu_{Nm}^{\omega'}} = \frac{\mu_{N1}^\omega}{\mu_{N1}^{\omega'}} > K$, Equation (A.16) implies that

$$\frac{\mu_{Nm'}^\omega}{\mu_{Nm'}^{\omega'}} \geq \varsigma^{2(m'-1)} \frac{\mu_{N1}^\omega}{\mu_{N1}^{\omega'}} > \varsigma^{2(m'-1)} K.$$

As Equation (A.15) implies that there must exist some m' with $\frac{\mu_{Nm'}^\omega}{\mu_{Nm'}^{\omega'}} \leq 1$, there must exist at least $\frac{\log K}{-2 \log \varsigma} + 1$ memory states with $\frac{\mu_{Nm}^\omega}{\mu_{Nm}^{\omega'}} \geq 1$. Repeating the same analysis for other $\omega'' \in \hat{\Omega}(k) \setminus \{\omega\}$ implies that if $\max_m \frac{\mu_{Nm}^\omega}{\mu_{Nm}^{\omega''}} > K$ for all $\omega'' \in \hat{\Omega}(k) \setminus \{\omega'\}$, there must exist at least $\frac{\log K}{-2 \log \varsigma} + 1$ memory states with a likelihood ratio $\min_{\omega' \in \hat{\Omega}(k)} \frac{\mu_{Nm}^\omega}{\mu_{Nm}^{\omega'}} \geq 1$ (i.e., $\frac{\mu_{Nm}^\omega}{\mu_{Nm}^{\omega'}} \geq 1$ for all $\omega' \in \hat{\Omega}(k) \setminus \{\omega\}$).

With the same argument, if $\max_m \frac{\mu_{Nm}^{\omega''}}{\mu_{Nm}^{\omega'}} > K$ for all $\omega'' \in \hat{\Omega}(k)$ and all $\omega' \in \hat{\Omega}(k) \setminus \{\omega''\}$, there must exist at least $\frac{\log K}{-2 \log \varsigma} + 1$ memory states with a likelihood ratio $\min_{\omega' \in \hat{\Omega}(k)} \frac{\mu_{Nm}^{\omega''}}{\mu_{Nm}^{\omega'}} \geq 1$ for all $\omega'' \in \hat{\Omega}(k)$. Therefore, $\frac{M}{|\hat{\Omega}(k)|} \geq \frac{\log K}{-2 \log \varsigma} + 1$. This implies that $\frac{M}{|\hat{\Omega}(k)|}$ goes to ∞ as K goes to ∞ , which contradicts the fact that $\lim_{k \rightarrow \infty} \frac{M(k)}{|\hat{\Omega}(k)|} = \lim_{k \rightarrow \infty} \frac{M(k)}{N(k)} < \infty$ in big worlds.

For reducible automata, the same conclusion follows. Denote the recurrent communicating classes as $\mathcal{R}_{k1}, \dots, \mathcal{R}_{kr}$ and the set of transient memory states as $\mathcal{R}_{k0}$. By Lemma A.1, the global ratio decomposes via absorption probabilities: for $m \in \mathcal{R}_{kj}$,

$$\frac{\mu_{Nm}^\omega}{\mu_{Nm}^{\omega'}} \leq \varsigma^{-2} \cdot \max_{m \in \mathcal{R}_{k0}} \underbrace{\frac{\mu_{Nm}^{0,\omega}}{\mu_{Nm}^{0,\omega'}}}_{\text{within-transient ratio}} \cdot \underbrace{\frac{\mu_{Nm}^{j,\omega}}{\mu_{Nm}^{j,\omega'}}}_{\text{within-class-}\mathcal{R}_{kj} \text{ ratio}}.$$

Hence if $\max_m \mu_{Nm}^\omega / \mu_{Nm}^{\omega'} > K$, then $\varsigma^{-2} \cdot a \cdot b > K$ where a is the within-transient ratio and b is the within-recurrent-class ratio. Equivalently $ab > \varsigma^2 K$, so $\max(a, b) > \varsigma \sqrt{K}$. The

irreducible-case counting argument applied to whichever class realizes the maximum yields a chain of $\frac{\log(\varsigma\sqrt{K})}{-2\log\varsigma} + 1$ memory states with elevated ratios in that class. Repeating for all $\omega \in \hat{\Omega}(k)$ as in the irreducible case yields $\frac{M(k)}{|\hat{\Omega}(k)|} \geq \frac{\log(\varsigma\sqrt{K})}{-2\log\varsigma} + 1 \rightarrow \infty$ as $K \rightarrow \infty$, contradicting the big-world condition. $\square$

Proof of Proposition 5. Suppose to the contrary that there exists a sequence of subsets of states $\tilde{\Omega}_k$ where $\lim_{k \rightarrow \infty} \frac{|\tilde{\Omega}_k|}{M(k)} > 0$, and that for all $\omega \in \tilde{\Omega}_k$, the DM takes action ω with some strictly positive probability in some state $\omega' \in \{1, \dots, N(k)\}$ at the limit where $k \rightarrow \infty$ (where ω' could be different for different ω). Formally, this implies that for all $\omega \in \tilde{\Omega}_k$

$$\lim_{k \rightarrow \infty} \sum_{m \in M_k^\omega} \mu_{km}^{\omega'} \geq \xi$$

for some ω' and $\xi > 0$. Since $\sum_\omega |M_k^\omega| \leq M(k)$, for any $C > 0$ the number of ω with $|M_k^\omega| > C$ is at most $M(k)/C$, i.e., there are at most $M(k)/C$ states of the world with more than C memory states allocated to their associated actions. Pick a constant C with $1/C < \lim_{k \rightarrow \infty} |\tilde{\Omega}_k|/M(k)$ (such C exists since the limit is strictly positive by assumption). Then the subset $\tilde{\Omega}'_k := \{\omega \in \tilde{\Omega}_k : |M_k^\omega| \leq C\}$ satisfies $|\tilde{\Omega}'_k| \geq |\tilde{\Omega}_k| - M(k)/C$, hence $\lim_{k \rightarrow \infty} |\tilde{\Omega}'_k|/M(k) \geq \lim_{k \rightarrow \infty} |\tilde{\Omega}_k|/M(k) - 1/C > 0$. By the pigeonhole principle, for each $\omega \in \tilde{\Omega}'_k$, there exists a $m \in M_k^\omega$ with $\lim_{k \rightarrow \infty} \mu_{km}^{\omega'} \geq \xi' > 0$, where $\xi' = \xi/C$.

Now denote the set of ω' , that is, the set of states of the world where the DM picks some action $\omega \in \tilde{\Omega}'_k$ with strictly positive probability, as $\tilde{\Omega}''_k$. First, we must have $\lim_{k \rightarrow \infty} \frac{|\tilde{\Omega}''_k|}{M(k)} > 0$. Otherwise, there must exist some $\omega' \in \tilde{\Omega}''_k$ such that there are infinitely many memory states with probability bigger than ξ' , contradicting $\lim_{k \rightarrow \infty} \sum_{m=1}^M \mu_{km}^{\omega'} = 1$. Furthermore, for all $\omega'' \in \tilde{\Omega}''_k$, the number of memory states with $\mu_{km}^{\omega''} \geq \xi'$ is at most $1/\xi'$ and is thus finite (since $\sum_m \mu_{km}^{\omega''} = 1$). To briefly summarize, we have a set of tuples (ω, ω', m) , where there is a one-to-one relationship between $\omega \in \tilde{\Omega}'_k$ and m (the representative $m \in M_k^\omega$ selected by the pigeonhole step above); $\omega \in \tilde{\Omega}'_k$ (and the associated m 's) and $\omega' \in \tilde{\Omega}''_k$ have non-vanishing limit in fractions w.r.t. $M(k)$ (and thus are infinite), and for each $\omega' \in \tilde{\Omega}''_k$, only a finite number of memory states have non-vanishing probability in the limit.

I now prove that for almost all $\omega \in \tilde{\Omega}'_k$ (measured in $1/M(k)$ fraction), there exists $m \in M_k^\omega$ with $\lim_{k \rightarrow \infty} \mu_{km}^{\omega'} \geq \xi'$ for some $\omega' \in \tilde{\Omega}''_k$ and $\lim_{k \rightarrow \infty} \mu_{km}^{\omega''} = 0$ for almost all $\omega'' \in \tilde{\Omega}''_k \setminus \{\omega'\}$ (in $1/M(k)$ fraction). By the previous step, for each $\omega' \in \tilde{\Omega}''_k$, only a finite number of memory states have $\lim_{k \rightarrow \infty} \mu_{km}^{\omega'} > 0$ (since $\sum_m \mu_{km}^{\omega'} = 1$ for each k). Hence the total number of pairs (ω', m) with $\omega' \in \tilde{\Omega}''_k$ and $\lim_{k \rightarrow \infty} \mu_{km}^{\omega'} > 0$ is at most a constant multiple of $|\tilde{\Omega}''_k| = \Theta(M(k))$. Writing $X(\omega) := |\{\omega'' \in \tilde{\Omega}''_k : \mu_{km(\omega)}^{\omega''} \not\rightarrow 0\}|$ (i.e., the count of ω'' at which the long-run probability $\mu_{km(\omega)}^{\omega''}$ has a non-vanishing limit) for each $\omega \in \tilde{\Omega}'_k$, the previous bound and disjointness of $\{M_k^\omega\}$ yield $\sum_{\omega \in \tilde{\Omega}'_k} X(\omega) = O(M(k))$. By Markov's inequality (counting form: $\#\{\omega : X(\omega) > t\} \leq \sum_\omega X(\omega)/t$), applied with $t = \eta M(k)$, the count of $\omega \in \tilde{\Omega}'_k$ with $X(\omega) > \eta M(k)$ is $O(1/\eta)$; dividing by $|\tilde{\Omega}'_k| = \Theta(M(k))$, its fraction in $\tilde{\Omega}'_k$ is $O(1/(\eta M(k))) \rightarrow 0$ for any $\eta > 0$. The claim thus follows. This implies that there

exists some sequence of subsets of states $\tilde{\Omega}_k$ where $\lim_{k \rightarrow \infty} \frac{|\tilde{\Omega}_k|}{M(k)} > 0$ such that

$$\lim_{k \rightarrow \infty} \max_m \frac{\mu_{km}^{\omega'}}{\mu_{km}^{\omega''}} = \infty$$

for all $\omega' \in \tilde{\Omega}_k$ and all $\omega'' \in \tilde{\Omega}_k \setminus \{\omega'\}$ (as the numerator $\geq \xi'$ but the denominator vanishes as $k \rightarrow \infty$). The sequence $\tilde{\Omega}_k$ thus has $\lim_{k \rightarrow \infty} |\tilde{\Omega}_k|/M(k) > 0$ and exhibits unbounded likelihood ratios for all pairs within — precisely ruled out by Lemma A.3. The result thus follows. $\square$

Proposition 5 implies $\lim_{k \rightarrow \infty} L_k^* = 1 - \bar{p}_0$. I establish matching upper and lower bounds.

Lower bound: $\lim_{k \rightarrow \infty} L_k^* \geq 1 - \bar{p}_0$. Fix any sequence of updating mechanisms and let $\Omega_k^{NI} = \{\omega : \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T \Pr^{\omega'}(a_t = \omega) > 0 \text{ for some } \omega'\}$ denote the set of actions taken with strictly positive long-run probability. By the proof of Proposition 5, $\lim_{k \rightarrow \infty} |\Omega_k^{NI}|/M(k) = 0$. In every state $\omega \notin \Omega_k^{NI}$, the DM picks action ω with long-run probability 0 and earns 0 utility in such states. The long-run loss therefore satisfies

$$\begin{aligned} \lim_{k \rightarrow \infty} L_k &\geq \lim_{k \rightarrow \infty} \sum_{\omega \notin \Omega_k^{NI}} p_{N(k)}^\omega \cdot u(\omega, \omega) \\ &= 1 - \lim_{k \rightarrow \infty} \sum_{\omega \in \Omega_k^{NI}} p_{N(k)}^\omega \\ &\geq 1 - \sup \left\{ \lim_{k \rightarrow \infty} \sum_{\omega \in \Omega_k} p_N^\omega : \Omega_k \subseteq \{1, \dots, N(k)\}, \lim_{N \rightarrow \infty} \frac{|\Omega_k|}{M(k)} = 0 \right\} \\ &= 1 - \bar{p}_0. \end{aligned}$$

Hence $\lim_{k \rightarrow \infty} L_k^* \geq 1 - \bar{p}_0$.

Upper bound: $\lim_{k \rightarrow \infty} L_k^* \leq 1 - \bar{p}_0$. Fix any $\eta > 0$. By the definition of $\bar{p}_0$, there exists a sequence of subsets $\Omega_k^f \subseteq \{1, \dots, N(k)\}$ with $|\Omega_k^f|/M(k) \rightarrow 0$ and $\lim_{k \rightarrow \infty} \sum_{\omega \in \Omega_k^f} p_{N(k)}^\omega \geq \bar{p}_0 - \eta$. Apply the case-1 ($h < 1$) construction from the proof of Proposition 2 with Ω_k^f in place of $\{1, \dots, N(k)\}$: since $\lambda(k) = M(k)/|\Omega_k^f| \rightarrow \infty$, the same argument yields long-run utility converging to 1 in every state $\omega \in \Omega_k^f$ (and the DM allocates no memory states to $\omega \notin \Omega_k^f$, contributing 0 utility there). Therefore

$$\lim_{k \rightarrow \infty} L_k^* \leq 1 - \lim_{k \rightarrow \infty} \sum_{\omega \in \Omega_k^f} p_{N(k)}^\omega \leq 1 - (\bar{p}_0 - \eta).$$

Letting $\eta \rightarrow 0$ gives $\lim_{k \rightarrow \infty} L_k^* \leq 1 - \bar{p}_0$.

Combining the two bounds, $\lim_{k \rightarrow \infty} L_k^* = 1 - \bar{p}_0$. $\square$

A.3 Proof of Corollary 1 and Proposition 3

Proof of Corollary 1 (ii). First, an ignorant updating mechanism induces utility loss weakly greater than $\min_{\omega \in \Omega} \min_{a \neq \omega} [p^\omega(u(\omega, \omega) - u(a, \omega))] > 0$. On the other hand, as shown in Proposition 1, L_M^* converges to 0 as $M \rightarrow \infty$. The monotonicity of L_M^* implies there exists some big enough $\bar{M}$ such that for $M \geq \bar{M}$, $L_M^* \leq L_{\bar{M}}^* < \min_{\omega \in \Omega} \min_{a \neq \omega} [p^\omega(u(\omega, \omega) - u(a, \omega))]$. The result follows. $\square$

Proof of Corollary 1 (i) and Proposition 3. First note that Proposition 3 implies Corollary 1 (i). I first prove Proposition 3(i). Consider an updating mechanism $(m_1, \mathcal{T}, d)$ such that $\mu_m^{\tilde{\omega}} > 0$ for some $\tilde{\omega}$ and m where $d(m) = \omega$. Now I prove that there exists a different updating mechanism that yields higher asymptotic utility if p^ω is small enough.

Now using Equation (A.1), denote $u^\omega = u(\omega, \omega)$, we have

$$\begin{aligned} \min_m \frac{\mu_m^{\omega'}}{\mu_m^\omega} &\geq \varsigma^{2(M-1)} \\ \frac{u^{\omega'} p^{\omega'} \mu_m^{\omega'}}{u^\omega p^\omega \mu_m^\omega} &\geq \varsigma^{2(M-1)} \frac{u^{\omega'} p^{\omega'}}{u^\omega p^\omega} \text{ for all } m. \end{aligned}$$

Now, denote $\bar{u} = \max_{\omega \in \Omega} \min_{a \neq \omega} [u(\omega, \omega) - u(a, \omega)]$ and $\underline{u} = \min_{\omega \in \Omega} \min_{a \neq \omega} [u(\omega, \omega) - u(a, \omega)]$, as $\min \max_{\omega'} \frac{u^{\omega'} p^{\omega'}}{u^\omega p^\omega} = \frac{\underline{u}}{\bar{u}} \frac{1-p^\omega}{p^\omega}$, there exists some ω' such that

$$\frac{u^{\omega'} p^{\omega'} \mu_m^{\omega'}}{u^\omega p^\omega \mu_m^\omega} \geq \varsigma^{2(M-1)} \min_{\omega''} \max_{\omega''} \frac{u^{\omega''} p^{\omega''}}{u^\omega p^\omega} \geq \varsigma^{2(M-1)} \frac{\underline{u}}{\bar{u}} \frac{1-p^\omega}{p^\omega} \text{ for all } m \in M. \quad (\text{A.17})$$

When $p^\omega < \frac{1}{2} \varsigma^{2(M-1)} \frac{\underline{u}}{\bar{u}} \frac{1-p^\omega}{p^\omega}$, or equivalently, $\frac{p^\omega}{1-p^\omega} < \frac{1}{2} \varsigma^{2(M-1)} \frac{\underline{u}}{\bar{u}} \frac{1}{N-1}$, Equation (A.17) implies that $u^{\omega'} p^{\omega'} \mu_m^{\omega'} > 2u^\omega p^\omega \mu_m^\omega$, and the DM is better off choosing ω' than choosing ω for all $m \in M$. As $u^{\omega'} p^{\omega'} \mu_m^{\omega'} > u^\omega p^\omega \mu_m^\omega$, it violates that $d(m)$ is optimal at memory state m . This proves Proposition 3(i)(a). For Proposition 3(i)(b), the improvement in utility by choosing ω' rather than choosing ω in m is equal to

$$u^{\omega'} p^{\omega'} \mu_m^{\omega'} - u^\omega p^\omega \mu_m^\omega > u^\omega p^\omega \mu_m^\omega.$$

The updating mechanism $(m_1, \mathcal{T}, d)$ is therefore not ϵ -optimal for $\epsilon \leq u^\omega p^\omega \mu_m^\omega$. Equivalently, if $(m_1, \mathcal{T}, d)$ is ϵ -optimal, then $\mu_m^\omega < \epsilon / (u^\omega p^\omega)$ for every memory state m with $d(m) = \omega$. Summing over all such m (of which there are at most M), the long-run probability that the DM picks action ω under state ω satisfies

$$\sum_{m: d(m)=\omega} \mu_m^\omega < \frac{M\epsilon}{u^\omega p^\omega}.$$

Applying the ratio bound $\mu_m^{\omega''} \leq \varsigma^{-2(M-1)} \mu_m^\omega$ (Lemma A.1), the long-run probability of picking action ω under any state ω'' is bounded by $M\epsilon \varsigma^{-2(M-1)} / (u^\omega p^\omega) = O(\epsilon)$, which vanishes as

$\epsilon \rightarrow 0$. This proves Proposition 3(i)(b). Similar argument proves Proposition 3(ii)(a) and (ii)(b).

I now prove Proposition 3(iii)(a) and (iii)(b). Without loss of generality, assume $u^{\omega'} p^{\omega'} \geq u^{\omega} p^{\omega}$. Again, consider an updating mechanism $\mathcal{T}$ such that $\mu_m^{\tilde{\omega}} > 0$ for some $\tilde{\omega}$ and m where $d(m) = \omega$. Similar to the proof of the first part of Proposition 3,

$$\begin{aligned} \min_m \frac{\mu_m^{\omega'}}{\mu_m^{\omega}} &\geq (\bar{l}^{\omega\omega'} \bar{l}^{\omega'\omega})^{-(M-1)} \max_m \frac{\mu_m^{\omega'}}{\mu_m^{\omega}} \\ \min_m \frac{\mu_m^{\omega'}}{\mu_m^{\omega}} &\geq \bar{F}^{-(M-1)} \max_m \frac{\mu_m^{\omega'}}{\mu_m^{\omega}} \\ \min_m \frac{\mu_m^{\omega'}}{\mu_m^{\omega}} &\geq \bar{F}^{-(M-1)} \\ \frac{u^{\omega'} p^{\omega'} \mu_m^{\omega'}}{u^{\omega} p^{\omega} \mu_m^{\omega}} &\geq \bar{F}^{-(M-1)} \frac{u^{\omega'} p^{\omega'}}{u^{\omega} p^{\omega}} \text{ for all } m. \end{aligned} \tag{A.18}$$

As $u^{\omega'} p^{\omega'} \geq u^{\omega} p^{\omega}$, for $\bar{F} < \frac{1}{2} \left[\frac{u^{\omega} p^{\omega}}{u^{\omega'} p^{\omega'}} \right]^{\frac{1}{M-1}}$, Equation (A.18) implies that $u^{\omega'} p^{\omega'} \mu_m^{\omega'} > 2u^{\omega} p^{\omega} \mu_m^{\omega}$. With similar argument in the proof of Proposition 3(i)(a) and (i)(b), the result follows. $\square$

A.4 Proof of Proposition 4

Proof. As mentioned in the main text, I consider a case where $N = M$ and states of the world are a priori uniformly distributed, i.e., $p^{\omega} = \frac{1}{N}$ for all $\omega = \{1, \dots, N\}$. Moreover, without loss of generality I assume $u(\omega, \omega) = 1 > u(\omega', \omega) = 0$ for all ω and $\omega' \neq \omega$. Signals are distributed in $S = \{s^1, \dots, s^N\}$ and s^{ω} is a signal that supports state ω . The signal distribution follows:

$$F^{\omega}(s^{\omega}) = \mathcal{J} F^{\omega}(s^{\omega'}) \text{ for all } \omega \text{ and } \omega' \neq \omega \text{ where } \mathcal{J} > 1.$$

Under all states of the world, it is $\mathcal{J}$ times more likely to receive a signal that supports the true state than a signal that supports one of the other states.

First, note that if the DM ignores no state, he allocates one memory state to each action. Without loss of generality, assume he takes action ω in memory state ω , i.e., $d(m) = m$ for all $m = 1, \dots, M$. I prove the Proposition in three steps: 1) to characterize the infimum utility loss of asymmetric non-ignorant updating mechanisms, it is without loss to assume that the DM transits to memory state m only if he receives a signal s^m , i.e., a signal supporting state m ; 2) the asymptotic utility of an optimal symmetric non-ignorant mechanism equals $\frac{\mathcal{J}}{\mathcal{J} + N - 1}$; 3) a “simple ignorant updating mechanism” that ignores half of the states outperforms the optimal symmetric non-ignorant mechanism when N is large enough or when $\mathcal{J}$ is small enough.

To prove 1), note that under symmetry and $d(m) = m$, the transition function is characterized by five parameters:

- $\alpha = \Pr[\mathcal{T}(i, s^i) = i]$ (stay on signal supporting the current action);
- $\beta = \Pr[\mathcal{T}(i, s^k) = i]$ for $k \neq i$ (stay on non-matching signal);
- $\gamma = \Pr[\mathcal{T}(i, s^i) = j]$ for $j \neq i$ (transit on own-state signal);
- $\delta = \Pr[\mathcal{T}(i, s^j) = j]$ for $j \neq i$ (transit on the target's matching signal);
- $\eta = \Pr[\mathcal{T}(i, s^k) = j]$ for i, j, k all distinct (transit on a cross signal),

with $\alpha + (N - 1)\gamma = 1$ and $\beta + \delta + (N - 2)\eta = 1$. Part 1's claim — transitions only on matching signals — corresponds to $\gamma = \eta = 0$.

By symmetry, the long-run distribution under any state ω takes the form $\mu_\omega^\omega = \mu^*$ and $\mu_m^\omega = \nu$ for $m \neq \omega$, with $\mu^* + (N - 1)\nu = 1$ and μ^* the same across ω . In the following I focus on state 1 and characterize μ^* . Let $A := q_{11}^1$ denote the self-transition probability of memory state 1 in state 1, and $B := q_{j1}^1$ for $j \neq 1$ the transition probability from any other memory state into memory state 1 in state 1 (well-defined because, by symmetry, q_{j1}^1 is the same for all $j \neq 1$). Using $F^1(s^1) = \mathcal{J}/(\mathcal{J} + N - 1)$ and $F^1(s^k) = 1/(\mathcal{J} + N - 1)$ for $k \neq 1$,

$$A = F^1(s^1)\alpha + \sum_{k \neq 1} F^1(s^k)\beta = \frac{\mathcal{J}\alpha + (N - 1)\beta}{\mathcal{J} + N - 1},$$

$$B = F^1(s^1)\delta + F^1(s^j)\gamma + \sum_{k \neq 1, k \neq j} F^1(s^k)\eta = \frac{\mathcal{J}\delta + \gamma + (N - 2)\eta}{\mathcal{J} + N - 1}.$$

As the long-run probabilities of transiting in and out of memory state 1 are equal, we have $\mu^*(1 - A) = (N - 1)\nu B$. Combined with $\nu = (1 - \mu^*)/(N - 1)$ gives

$$\mu^* = \frac{B}{1 - A + B}.$$

Maximizing μ^* is equivalent to minimizing $(1 - A)/B$. Substituting the normalizations $\alpha = 1 - (N - 1)\gamma$ and $\beta = 1 - \delta - (N - 2)\eta$ into the formula for A above gives $1 - A = (N - 1)[\mathcal{J}\gamma + \delta + (N - 2)\eta]/(\mathcal{J} + N - 1)$, and therefore

$$\frac{1 - A}{B} = (N - 1) \frac{\mathcal{J}\gamma + \delta + (N - 2)\eta}{\mathcal{J}\delta + \gamma + (N - 2)\eta}.$$

Differentiating the ratio with respect to γ gives

$$\frac{\partial}{\partial \gamma} \frac{\mathcal{J}\gamma + \delta + (N - 2)\eta}{\mathcal{J}\delta + \gamma + (N - 2)\eta} = \frac{(\mathcal{J} - 1)[(\mathcal{J} + 1)\delta + (N - 2)\eta]}{(\mathcal{J}\delta + \gamma + (N - 2)\eta)^2} \geq 0,$$

so the ratio is minimized at $\gamma = 0$. Differentiating with respect to η ,

$$\frac{\partial}{\partial \eta} \frac{\mathcal{J}\gamma + \delta + (N - 2)\eta}{\mathcal{J}\delta + \gamma + (N - 2)\eta} = \frac{(N - 2)(\mathcal{J} - 1)(\delta - \gamma)}{(\mathcal{J}\delta + \gamma + (N - 2)\eta)^2}.$$

Given $\gamma = 0$ and $\delta > 0$, this is non-negative, so the ratio is further minimized at $\eta = 0$. At $\gamma = \eta = 0$ (with any $\delta > 0$),

$$\frac{1 - A}{B} = \frac{(N - 1)\delta}{\mathcal{J}\delta} = \frac{N - 1}{\mathcal{J}}, \quad \mu^* = \frac{\mathcal{J}}{\mathcal{J} + N - 1}.$$

Hence (one of) the optimal symmetric non-ignorant mechanism has $\gamma = \eta = 0$, i.e., the DM transits to a different memory state j only upon receiving the matching signal s^j . This proves Part 1 within the symmetric class; the maximum value $\mu^* = \mathcal{J}/(\mathcal{J} + N - 1)$ also delivers Part 2 directly.

I now prove 3). Consider an ignorant mechanism that follows the similar idea of the non-ignorant updating mechanism illustrated in Figure 9, but ignores half of the states of the world. For simplicity, assume that N is even. The ignorant mechanism is illustrated in Figure 10, with an example of $N = 4$. By ignoring half of the states, the DM allocates two memory states to each action that he does not ignore. Without loss of generality, assume that the DM takes action ω in memory states $2\omega - 1$ and 2ω for $\omega \leq \frac{N}{2}$, and ignores all actions $\omega' > \frac{N}{2}$. In the “more confident” memory state 2ω , upon receiving a signal supporting state $\omega' \neq \omega$ where $\omega' \leq \frac{N}{2}$, the DM transits to state $2\omega - 1$ with probability $\delta < 1$ and stays in his current memory state otherwise. In the “less confident” memory state $2\omega - 1$, upon receiving a signal that supports state ω , he transits to the “more confident” memory state 2ω ; upon receiving a signal that supports state $\omega' \neq \omega$ where $\omega' \leq \frac{N}{2}$, the DM transits to state $2\omega' - 1$ with probability $\delta < 1$ and stays in his current memory state otherwise.

Formally, the transition function is as follows:

$$\begin{aligned} \mathcal{T}(2\omega, s^\omega) &= 2\omega && \text{for all } \omega \leq \frac{N}{2}, \\ \mathcal{T}(2\omega, s^{\omega'}) &= \begin{cases} 2\omega - 1 \text{ with probability } \delta \\ 2\omega \text{ with probability } 1 - \delta \end{cases} && \text{for all } \omega \leq \frac{N}{2}, \omega' \neq \omega \text{ and } \omega' \leq \frac{N}{2}, \\ \mathcal{T}(2\omega - 1, s^\omega) &= 2\omega && \text{for all } \omega \leq \frac{N}{2}, \\ \mathcal{T}(2\omega - 1, s^{\omega'}) &= \begin{cases} 2\omega' - 1 \text{ with probability } \delta \\ 2\omega \text{ with probability } 1 - \delta \end{cases} && \text{for all } \omega \leq \frac{N}{2}, \omega' \neq \omega \text{ and } \omega' \leq \frac{N}{2}, \\ \mathcal{T}(m, s^{\omega'}) &= m && \text{for all } \omega > \frac{N}{2}. \end{aligned}$$

Suppose the true state is 1, In the stationary distribution, we must have

$$\begin{aligned} \delta \mu_{2\omega}^1 \sum_{\omega' \neq \omega, \omega' \leq \frac{N}{2}} F^1(s^{\omega'}) &= \mu_{2\omega-1}^1 F^1(s^\omega) \text{ for all } \omega \leq \frac{N}{2}, \\ \delta \mu_{2\omega-1}^1 F^1(s^{\omega'}) &= \delta \mu_{2\omega'-1}^1 F^1(s^\omega) \text{ for all } \omega, \omega' \leq \frac{N}{2} \text{ and } \omega \neq \omega'. \end{aligned} \tag{A.19}$$

From the first Equation of Equation (A.19), we know that when δ is close to 0, $\mu_{2\omega-1}^1$ is close

to 0 for all $\omega \leq \frac{N}{2}$. Moreover, we have

$$\begin{aligned}\mu_2^1 &= \frac{F^1(s^1)}{\delta \sum_{\omega' \neq 1, \omega' \leq \frac{N}{2}} F^1(s^{\omega'})} \times \frac{\delta F^1(s^1)}{\delta F^1(s^\omega)} \times \frac{\delta \sum_{\omega' \neq \omega, \omega' \leq \frac{N}{2}} F^1(s^{\omega'})}{F^1(s^\omega)} \times \mu_{2\omega}^1 \\ &= \frac{\mathcal{J}}{\frac{N}{2} - 1} \times \mathcal{J} \times \left(\frac{N}{2} - 2 + \mathcal{J} \right) \times \mu_{2\omega}^1 \\ &= \frac{\mathcal{J}^2(N + 2\mathcal{J} - 4)}{N - 2} \mu_{2\omega}^1\end{aligned}$$

for all $\omega \neq 1$ and $\omega \leq \frac{N}{2}$. We have

$$\begin{aligned}\mu_2^1 + \sum_{\omega \neq 1, \omega \leq \frac{N}{2}} \mu_{2\omega}^1 &= 1 \\ \mu_2^1 + \frac{N - 2}{\mathcal{J}^2(N + 2\mathcal{J} - 4)} \left(\frac{N}{2} - 1 \right) \mu_2^1 &= 1 \\ \mu_2^1 &= \frac{2\mathcal{J}^2(N + 2\mathcal{J} - 4)}{2\mathcal{J}^2(N + 2\mathcal{J} - 4) + (N - 2)^2}.\end{aligned}$$

By repeating the same computation for all $\omega \leq \frac{N}{2}$, the asymptotic utility equals:

$$\sum_{\omega=1}^{\frac{N}{2}} \frac{2\mathcal{J}^2(N + 2\mathcal{J} - 4)}{2\mathcal{J}^2(N + 2\mathcal{J} - 4) + (N - 2)^2} \times \frac{1}{N} = \frac{\mathcal{J}^2(N + 2\mathcal{J} - 4)}{2\mathcal{J}^2(N + 2\mathcal{J} - 4) + (N - 2)^2} \quad (\text{A.20})$$

which is larger than the utility of an optimal ignorant mechanism if and only if:

$$\frac{\mathcal{J}^2(N + 2\mathcal{J} - 4)}{2\mathcal{J}^2(N + 2\mathcal{J} - 4) + (N - 2)^2} \geq \frac{\mathcal{J}}{\mathcal{J} + N - 1}$$

or equivalently:

$$\mathcal{J}(N + 2\mathcal{J} - 4)(\mathcal{J} + N - 1) - 2\mathcal{J}^2(N + 2\mathcal{J} - 4) - (N - 2)^2 \geq 0$$

Simple algebra shows that the Inequality holds for $\mathcal{J} \in [1, \frac{1}{4}(\sqrt{9N^2 - 32N + 32} + N)]$, where the segment increases in N .

□

A.5 Proof of Corollary 2

Proof. Suppose to the contrary that the sequence of updating mechanisms ignores a set of states with total prior probability strictly greater than ϵ at the limit. That is, denoting $\tilde{N}(k)$ as the set of states ignored by the sequence at the limit (in the sense of Definition 2), we

have, for every $\omega \in \tilde{N}(k)$ and every state ω' ,

$$\lim_{k \rightarrow \infty} \lim_{T \rightarrow \infty} E_{\omega'} \left[\frac{\sum_{t=1}^T \mathbb{1}_{a_t=\omega}}{T} \right] = 0,$$

and $\lim_{k \rightarrow \infty} \sum_{\omega \in \tilde{N}(k)} p_{N(k)}^\omega > \epsilon$. The utility loss as $k \rightarrow \infty$ is

$$\begin{aligned} \lim_{k \rightarrow \infty} L_k &\geq \lim_{k \rightarrow \infty} \sum_{\omega \in \tilde{N}(k)} \left[u_{N(k)}(\omega, \omega) - \max_{\omega' \neq \omega} u_{N(k)}(\omega', \omega) \right] p_{N(k)}^\omega \\ &= \lim_{k \rightarrow \infty} \sum_{\omega \in \tilde{N}(k)} p_{N(k)}^\omega > \epsilon = \lim_{k \rightarrow \infty} L_k^* + \epsilon \end{aligned}$$

where the equality uses $u(\omega, \omega) = 1$ and $u(\omega', \omega) = 0$ for $\omega' \neq \omega$. This contradicts ϵ -optimality of the sequence at the limit, proving the result. $\square$

A.6 Proof of Corollary 3

Proof. First, (i) is directly implied by Proposition 3. If p_A^ω is small enough for all $\omega \in N_A \subset N$ and p_B^ω is small enough for all $\omega \in N_B = N \setminus N_A$, individual A (almost) never picks action ω for all $\omega \in N_A$ and individual B (almost) never picks action ω for all $\omega \in N \setminus N_A$. Therefore, they must disagree with each other.

To prove (ii), note that by Proposition 1, we know that if an individual adopts an ϵ -optimal updating mechanism, his utility loss is bounded above by $Kr^{\lfloor \frac{M-1}{N} \rfloor} + \epsilon$ for some $K > 0$ and $r < 1$. Thus, for all ω , the probability that the individual chooses a “wrong” action $\omega' \neq \omega$ is strictly smaller than $\left[\frac{Kr^{\lfloor \frac{M-1}{N} \rfloor} + \epsilon}{\min_\omega p^\omega} \right]$. For all ω , the probability that both individuals choose action ω and thus agree with each other is greater than

$$\begin{aligned} &\left[1 - \left[\frac{K_A r_A^{\lfloor \frac{M_A-1}{N} \rfloor} + \epsilon}{\min_\omega p_A^\omega} \right] \right] \left[1 - \left[\frac{K_B r_B^{\lfloor \frac{M_B-1}{N} \rfloor} + \epsilon}{\min_\omega p_B^\omega} \right] \right] \\ &> 1 - \left[\frac{1}{\min_\omega p_A^\omega} + \frac{1}{\min_\omega p_B^\omega} \right] (\max\{K_A, K_B\} (\max\{r_A, r_B\})^{\lfloor \frac{\min\{M_A, M_B\}-1}{N} \rfloor} + \epsilon) \end{aligned}$$

for all ϵ -optimal updating mechanisms of individual A and B . The result follows. $\square$

A.7 Proof of Corollary 4

Proof. I prove the Corollary using the following simple example. Consider a setting with $N = 3$ and two individuals, A and B , who share the same prior belief and the same objective

signal structure:

$$\begin{aligned} p^1 &= \frac{1}{3} + 2\nu \\ p^2 &= p^3 = \frac{1}{3} - \nu \end{aligned} \tag{A.21}$$

$$\begin{aligned} \sup_s \frac{f^1(s)}{f^n(s)} &= \sup_s \frac{f^n(s)}{f^1(s)} = \sqrt{1+\tau} \text{ for } n = 2, 3 \\ \sup_s \frac{f^2(s)}{f^3(s)} &= \sup_s \frac{f^3(s)}{f^2(s)} = \sqrt{1+\Psi} \text{ where } \Psi > \tau. \end{aligned} \tag{A.22}$$

with $1 + \tau \geq \frac{\frac{1}{3} + 2\nu}{\frac{1}{3} - \nu}$.³² Moreover, assume that $u(1, 1) = u(2, 2) = u(3, 3) = 1$ and $u(\omega, \omega') = 0$ for all $\omega \neq \omega'$. The only difference the two individuals have is their levels of cognitive ability, where $M_A = 1$ and $M_B = 2$.

First, as $M = 1$ for individual A , his action is constant in all periods for all signal realizations and thus $d(1) = 1$. Individual A always takes action 1. Now I characterize the ϵ -optimal updating mechanism of individual B . With some abuse of notations, denote $L_2^*(nn')$ as optimal utility loss where the DM chooses action n in memory state 1 and action n' in memory state 2. We have, obviously,

$$\begin{aligned} L_2^*(11) &= \frac{2}{3} - 2\nu, \\ L_2^*(22) &= L_2^*(33) = \frac{2}{3} + \nu. \end{aligned}$$

Moreover, as argued in the proof of Corollary 1 (i), suppose $d(1) = n$ and $d(2) = n' \neq n$, we have

$$\frac{\mu_1^n}{1 - \mu_2^n} / \frac{1 - \mu_1^{n'}}{\mu_2^{n'}} = \frac{\mu_1^n}{\mu_1^{n'}} / \frac{\mu_2^n}{\mu_2^{n'}} \leq \sup_s \frac{f^n(s)}{f^{n'}(s)} \sup_s \frac{f^{n'}(s)}{f^n(s)}.$$

The upper bound of $L^*(nn')$ is given by the following minimization problem:

$$\begin{aligned} \min_{\mu_1^n, \mu_2^{n'}} & 1 - p^n - p^{n'} + p^n(1 - \mu_1^n) + p^{n'}(1 - \mu_2^{n'}) \\ \text{given} & \frac{\mu_1^n}{1 - \mu_2^{n'}} / \frac{1 - \mu_1^n}{\mu_2^{n'}} \leq \sup_s \frac{f^n(s)}{f^{n'}(s)} \sup_s \frac{f^{n'}(s)}{f^n(s)}, \end{aligned}$$

Hellman and Cover (1970) shows that the upper bound is tight with the updating mechanism where $\mathcal{T}(1, s') = 2$ if and only if $\frac{f^{n'}(s')}{f^n(s')}$ is close to $\sup_s \frac{f^{n'}(s)}{f^n(s)}$, and $\mathcal{T}(2, s') = 1$ if and only if

³²It ensures that if $M \geq 2$, the DM never chooses action 1 with probability 1 and he can achieve a strictly lower utility loss compared to the benchmark of no information.

$\frac{f^n(s')}{f^{n'}(s')}$ is close to $\sup_s \frac{f^n(s)}{f^{n'}(s)}$. Therefore,

$$\begin{aligned} L_2^*(12) = L_2^*(13) &= \frac{1}{3} - \nu + \frac{2\sqrt{(1+\tau)(\frac{1}{3}+2\nu)(\frac{1}{3}-\nu)} - (\frac{2}{3}-\nu)}{\tau}, \\ L_2^*(23) &= \frac{1}{3} + 2\nu + \frac{2(\frac{1}{3}-\nu)\sqrt{1+\Psi} - (\frac{2}{3}-2\nu)}{\Psi}. \end{aligned}$$

where $1+\tau \geq \frac{\frac{1}{3}+2\nu}{\frac{1}{3}-\nu}$ implies $L_2^*(22) = L_2^*(33) > L_2^*(11) \geq L_2^*(12) = L_2^*(13)$. In the following, I prove $L_2^*(12) > L_2^*(23)$ if ν is small enough which implies that $a_t^B = 2$ or 3 for all t in all ϵ -optimal updating mechanism when ϵ is smaller than $L_2^*(12) - L_2^*(23)$. Now, $L_2^*(12) > L_2^*(23)$ if and only if

$$L_2^*(12) - L_2^*(23) = 3\nu + \frac{2(\frac{1}{3}-\nu)\sqrt{1+\Psi}}{\Psi} - \frac{\frac{2}{3}-2\nu}{\Psi} - \frac{2\sqrt{(1+\tau)(\frac{1}{3}+2\nu)(\frac{1}{3}-\nu)}}{\tau} + \frac{\frac{2}{3}+\nu}{\tau} < 0.$$

When $\nu = 0$,

$$L_2^*(12) - L_2^*(23) = \frac{2}{3} \left(\frac{\sqrt{1+\Psi}}{\Psi} - \frac{\sqrt{1+\tau}}{\tau} \right) - \frac{2}{3} \left(\frac{1}{\Psi} - \frac{1}{\tau} \right).$$

As both $\frac{\sqrt{1+x}}{x}$ and $\frac{1}{x}$ decreases in x , $\Psi > \tau$ implies that $L_2^*(12) - L_2^*(23) < 0$ when $\nu = 0$. The result follows by continuity. $\square$

A.8 Proof of Corollary 5

Proof. With some abuse of notation, denote $\mathcal{J}_{\mathbf{k}i}^\epsilon$ as the set of ϵ -optimal updating mechanisms of individual $i = A, B$; for ease of notation, I write p_{Ni}^ω for $p_{N(k)i}^\omega$ throughout, suppressing the k -dependence. By Proposition 2, $\lim_{k \rightarrow \infty} L_k^* = 0$ for $h < 1$. Moreover, by Assumption (Limit-FS), any set of states with non-vanishing fraction w.r.t. N carries non-vanishing prior probability. Together these imply that the prior probability on states where the long-run probability of choosing the correct action tends to 1 must itself tend to 1; formally,

$$\lim_{k \rightarrow \infty} \lim_{\epsilon \rightarrow 0} \sup_{(m_{1ki}, \mathcal{J}_{ki}, d_{ki}) \in \mathcal{J}_{\mathbf{k}i}^\epsilon} \sum_{\omega} \left[\mathbb{1} \left\{ \lim_{T \rightarrow \infty} E_{\omega, \mathcal{J}_{ki}} \frac{\sum_{t=1}^T \mathbb{1}_{a_t^i = \omega}}{T} = 1 \right\} p_{Ni}^\omega \right] = 1$$

which is equivalent to

$$\lim_{k \rightarrow \infty} \lim_{\epsilon \rightarrow 0} \sup_{(m_{1ki}, \mathcal{J}_{ki}, d_{ki}) \in \mathcal{J}_{\mathbf{k}i}^\epsilon} \sum_{\omega} \left[\mathbb{1} \left\{ \lim_{T \rightarrow \infty} E_{\omega, \mathcal{J}_{ki}} \frac{\sum_{t=1}^T \mathbb{1}_{a_t^i = \omega}}{T} < 1 \right\} p_{Ni}^\omega \right] = 0,$$

i.e., individual i would only take sub-optimal actions in probability 0 events measured by $(p_{Ni}^\omega)_{\omega=1}^N$. As $\lim_{N \rightarrow \infty} p_{NA}^\omega > 0$ if and only if $\lim_{N \rightarrow \infty} p_{NB}^\omega > 0$ for all ω , combined with

Assumption (Limit-FS), this implies that

$$\lim_{k \rightarrow \infty} \sum_{\omega \in \tilde{\Omega}(k)} p_{NA}^\omega = 0 \text{ if and only if } \lim_{k \rightarrow \infty} \sum_{\omega \in \tilde{\Omega}(k)} p_{NB}^\omega = 0$$

for all $\tilde{\Omega}(k) \in \{1, \dots, N(k)\}$. Thus,

$$\lim_{k \rightarrow \infty} \lim_{\epsilon \rightarrow 0} \sup_{(m_{1ki}, \mathcal{T}_{ki}, d_{ki}) \in \mathcal{T}_{\mathbf{k}i}^\epsilon} \sum_{\omega} \left[\mathbb{1} \left\{ \lim_{T \rightarrow \infty} E_{\omega, \mathcal{T}_{ki}} \frac{\sum_{t=1}^T \mathbb{1}_{a_t^i = \omega}}{T} < 1 \right\} p_{NI}^\omega \right] = 0 \text{ for all } i, I \in \{A, B\},$$

i.e., even when individual i 's mechanism is evaluated under the other individual's prior, the prior mass on suboptimal outcomes vanishes in the limit. Bounding disagreement by a union bound over the two individuals' suboptimal events,

$$\begin{aligned} & \lim_{k \rightarrow \infty} \lim_{\epsilon \rightarrow 0} \sup_{(m_{1kA}, \mathcal{T}_{kA}, d_{kA}) \in \mathcal{T}_{\mathbf{k}A}^\epsilon, (m_{1kB}, \mathcal{T}_{kB}, d_{kB}) \in \mathcal{T}_{\mathbf{k}B}^\epsilon} \sum_{\omega} \left[\mathbb{1} \left\{ \lim_{T \rightarrow \infty} E_{\omega, \mathcal{T}_{kA}, \mathcal{T}_{kB}} \frac{\sum_{t=1}^T \mathbb{1}_{a_t^A = a_t^B = \omega}}{T} < 1 \right\} p_{Ni}^\omega \right] \\ & \leq \lim_{k \rightarrow \infty} \lim_{\epsilon \rightarrow 0} \sup_{(m_{1kA}, \mathcal{T}_{kA}, d_{kA}) \in \mathcal{T}_{\mathbf{k}A}^\epsilon} \sum_{\omega} \left[\mathbb{1} \left\{ \lim_{T \rightarrow \infty} E_{\omega, \mathcal{T}_{kA}} \frac{\sum_{t=1}^T \mathbb{1}_{a_t^A = \omega}}{T} < 1 \right\} p_{Ni}^\omega \right] \\ & \quad + \lim_{k \rightarrow \infty} \lim_{\epsilon \rightarrow 0} \sup_{(m_{1kB}, \mathcal{T}_{kB}, d_{kB}) \in \mathcal{T}_{\mathbf{k}B}^\epsilon} \sum_{\omega} \left[\mathbb{1} \left\{ \lim_{T \rightarrow \infty} E_{\omega, \mathcal{T}_{kB}} \frac{\sum_{t=1}^T \mathbb{1}_{a_t^B = \omega}}{T} < 1 \right\} p_{Ni}^\omega \right] \\ & = 0 \end{aligned}$$

for $i = A, B$. The result follows. $\square$

A.9 Proof of Corollary 6

Proof. For ease of notation, I write p_N^ω and u_N for $p_{N(k)}^\omega$ and $u_{N(k)}$, suppressing the k -dependence. Consider an example where $\lim_{k \rightarrow \infty} p_N^\omega = 0$ for all ω (so that a set of states has non-vanishing prior probability if and only if it has a non-vanishing fraction w.r.t. N). As shown in Proposition 5, all sequences of updating mechanisms must ignore almost all actions when k goes to infinity. Thus, $\lim_{k \rightarrow \infty} L_k^* = \lim_{k \rightarrow \infty} \sum_{\omega} u_N(\omega, \omega) p_N^\omega$ and in the limit, every updating mechanism is ϵ -optimal. Thus, if individual A adopts an updating mechanism with $d_k(m) = 1$ for all m and k , while individual B adopts an updating mechanism with $d_k(m) = 2$ for all m and all k such that $N(k) \geq 2$, then they must disagree with each other under all ω even though they adopt a sequence of updating mechanisms that is ϵ -optimal in the limit. $\square$

B Discounted future utility

In this section, I consider the alternative formulation where future utility is discounted. More specifically, consider the following updating mechanism

$$(m_1^{*D}(\rho), \mathcal{T}^{*D}(\rho), d^{*D}(\rho)) = \arg \max_{m_1, \mathcal{T}, d} E \left[(1 - \rho) \sum_{t=1}^{\infty} \rho^t u(a_t, \omega) \right],$$

for some discount factor $\rho < 1$. We have the following result.

Proposition B.1. *For all $\epsilon > 0$, there exists some $\bar{\rho} < 1$ such that $L(m_1^{*D}(\rho), \mathcal{T}^{*D}(\rho), d^{*D}(\rho)) < L_M^* + \epsilon$ for all $\rho > \bar{\rho}$.*

Proof. Suppose to the contrary that $\lim_{\rho \rightarrow 1} L(m_1^{*D}(\rho), \mathcal{T}^{*D}(\rho), d^{*D}(\rho)) > L_M^* + \epsilon$ for some $\epsilon > 0$. It implies that there exists some updating mechanism $(m_1, \mathcal{T}, d)$ that yields strictly higher asymptotic utility than $(m_1^{*D}(\rho), \mathcal{T}^{*D}(\rho), d^{*D}(\rho))$ for all ρ close to 1. By a standard Abelian-type result (Hardy, 1949) — for any bounded sequence $\{u_t\}$ with a well-defined long-run average $\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=1}^T u_t$, the discounted average $(1 - \rho) \sum_{t=1}^{\infty} \rho^t u_t$ converges to the same value as $\rho \rightarrow 1$ — $(m_1, \mathcal{T}, d)$ also yields strictly higher discounted sum of utility than $(m_1^{*D}(\rho), \mathcal{T}^{*D}(\rho), d^{*D}(\rho))$ for all ρ close to 1. This contradicts that $(m_1^{*D}(\rho), \mathcal{T}^{*D}(\rho), d^{*D}(\rho))$ is the discounted-utility maximizer, proving the result. $\square$

In other words, the updating mechanism that maximizes the discounted sum of utility, $(m_1^{*D}(\rho), \mathcal{T}^{*D}(\rho), d^{*D}(\rho))$, is ϵ -optimal for any $\epsilon > 0$ given ρ close enough to 1. This implies that the main results extend to the discounted case as $\rho \rightarrow 1$. To see why, consider the results on ignorance (Corollary 1 and Proposition 5): ϵ -optimal mechanisms with sufficiently small ϵ must ignore some states in some big-world environments (Corollary 1) and ignore almost all states in any big-world environment at the limit (Proposition 5). Since the discount-optimal mechanism is ϵ -optimal for small enough ϵ when ρ is close to 1, it must also ignore some states. Conversely, consider the results on non-ignorance (Corollary 1(ii) and Corollary 2): every ϵ -optimal mechanism with sufficiently small ϵ is non-ignorant in the small world. Since the discount-optimal mechanism falls in this class for ρ close to 1, it is also non-ignorant.